Finding virtue in open science?

Biological scientists' constructions of openness in historical, advocacy and policy contexts

---

Rosalind Attenborough

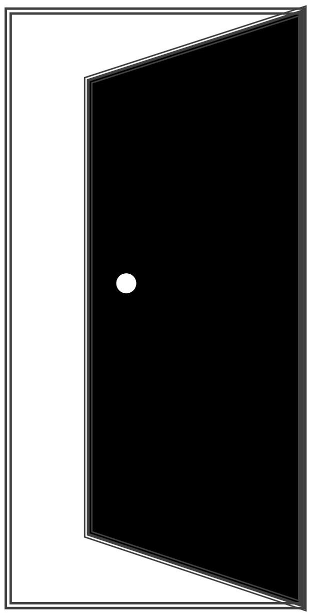
PhD

Science and Technology Studies

The University of Edinburgh

2020

For my mentors and role models in science, in STS, and amongst my family and friends.
I have been very fortunate.



Abstract

Science has a special relationship with the term "open" and its connotations. A traditional story about scientific openness goes as follows: if scientists share their findings, scientific communities can collectively build upon these findings and a progressive corpus of knowledge emerges. But since the turn of the twenty-first century, a distinctive, online "open science" has rapidly gained global salience, incorporating practices from open access publishing and open research data, to open preprints, open peer review, and open notebook science. Movements towards such practices have often been led from within scientific communities – by scientist-activists and entrepreneurs. Such actors see the Internet as an unprecedented opportunity to "open" science, and fix seemingly broken aspects of the scientific system: inaccessibility, opacity, irreproducibility. More recently, the "open" imperative is also top-down, as funding and research organisations increasingly treat open practices as desirable or mandatory.

This work focuses on academic, biological scientists in the UK and Australia whose professional and epistemic worlds are undergoing transformation in this open science "revolution" – in whose communities openness may have long-standing meaning, but wherein "open science" may have risen from obscurity to salience in the space of only 15 or 20 years. While some scientists are the leaders of open movements, many are said to be ambivalent and slow to adopt open practices, forming a "cultural" barrier to openness that is rarely explored in systematic empirical studies. This disparity has a moral dimension, as openness is positioned a quality of good science and scientists. My research questions consider how scientists' constructions sit in relation to historical, advocacy and policy framings; why scientists may be disengaged from contemporary open science movements; and the extent to which they construct and internalise openness as an epistemic virtue: a moralised truth-making quality.

The thesis begins with an analysis of openness in science as a concept and practice with historical depth as well as contemporary salience; I then explore its contemporary framing in advocacy and policy contexts through document analysis. In both I consider how "open" (or "openness") is being constructed, and the significance of its flexibility and expansiveness. These analyses set the scene for the empirical core of the work: 40 in-depth, semi-structured interviews with biologists, purposively sampled for disciplinary, generational, gender, and attitudinal diversity. For context and counterpoint, I conducted 14 similar interviews with open science advocates and policymakers. Through these interviews, I attend to how a broad population of scientists, as well as advocates and policymakers, construct "open" in science.

My findings focus on the three most common categories of scientific openness emerging from interviews with biologists: open access, data openness, and interpersonal openness. The first two of these have close connections with policy and advocacy movements, whereas the third appears to be anchored only in scientists' experiences and implicit conceptualisations. Nonetheless, interpersonal openness is constructed in consistent ways and with conviction: it refers to the practice and principle of "talking freely" about unpublished ideas and data in small-scale interpersonal situations, or the contextual withholding of such information. I characterise scientists' constructions of each of these three categories as indicators of how scientists encounter and enact top-down and bottom-up forms of scientific openness.

In my discussion and conclusion, I bring these three categories of scientific openness into conversation, using them to theorise the variety of relationships that scientists form with scientific openness under the contemporary open science "revolution" - including whether and how different forms of openness are internalised as epistemic virtues. In turn, this allows commentary about: the apparent disconnectedness of scientists from certain agendas of openness; the significance of generational differences; and interactions between openness, privilege and inequality in pressured scientific systems.

Lay summary

In the last twenty years, “open” has become a prominent buzzword in science. Many movements supporting open science have arisen within communities of researchers and overlapping groups interested in Web technology and research communication. Open science has also become a top-down policy agenda. Many governments, research funders, universities, and publishers worldwide now ask or insist that researchers adopt “open” practices. Open practices and concepts are diverse. Two of the most prominent are open access and open data, which entail free, unrestricted online access to published research findings, and to underlying datasets, respectively. More ambitious visions of open science evoke “revolution”, and press for online transparency of entire research processes including plans, methods, and real-time results. Open science has other interesting features: it resonates with an older, traditional notion that openness is essential to science. It is also curiously difficult to define, as its scope continues to expand. It is sometimes characterised simply as “good science”. Yet, many scientists do not seem to understand open science in the same way as its advocates, and seem slow or ambivalent in their uptake of open practices.

In this research, I investigate these interesting social features of open science. I begin in historical contexts, tracing traditional ideals of scientific openness, and the roots of today’s open science movements. I also examine present-day policy and advocacy documents, to make sense of current meanings of “open”. This provides context for an interview study with 40 biological scientists and 14 advocates or policymakers about their understandings of openness in science. These were mainly based in the UK and Australia. The most prominent “openness” topics for scientists were open access, data openness, and interpersonal openness. The last of these topics is not part of today’s open science movements – and is not an online practice – but has special importance to many of the scientists interviewed. I provide in-depth qualitative analysis of how scientist interviewees understood each topic. This allows me to reflect in a nuanced way about whether – and how – scientists see open science as good science, and why they may adopt openness enthusiastically, or keep it at arm’s length.

Acknowledgements

Like many PhD projects, this work is a labour of love – and a struggle – that has been supported by the care, generosity, and expertise of many people. First I wish to thank the 54 scientists and open science advocates/policymakers whose contributions are at the core of this thesis. What I have written is in many ways for you, and I hope I have done justice representing and empathising with the diverse complexity of your experiences. Interviewing was one of the most enjoyable parts of this research; thank you for your time, your candour, and your insights. Thank you to my named interviewees: Paul Ayris, Gregory P. Copenhaver, Enrico Crema, Stephanie Dawson, Jenny Graves, David Hume, Danny Kingsley, Catriona MacCallum, Jenny Molloy, Peter Murray-Rust, Cameron Neylon, Mark Patterson, Jason Gellis, Jon Tennant, Stuart Taylor, Mark Tester, and Matthew Todd. Thank you equally to those of you who remain anonymous. This work was funded by a 1+3 studentship from the Economic and Social Research Council.

Heartfelt thanks to my supervisors, Niki Vermeulen and James Mittra, who proactively supported this research idea from its early stages. Niki, thank you for so empathetically seeing what I was trying to achieve and why I care about it – and being there to remind me of this, and to exchange ideas, all the way through the journey. James, thank you for your unwavering support throughout, your insights, and your practical focus that helped me make it to submission day. Thank you both for helping me maintain confidence and a sense of joy, humour, and proportion. I could not have done a PhD without such caring mentorship. Many thanks to Jane Calvert and Sabina Leonelli for their expert thesis examination: it was a true privilege to have your input.

I owe thanks to many other members of the Science, Technology and Innovation Studies community at the University of Edinburgh. Emma Frow was leaving STIS as I joined, but helped me find my path. She and Steve Sturdy examined my first year board paper and provided valuable encouragement and critique. Thank you to Robin Williams and Cathie Lyall who have supported me as Postgraduate Advisors. Thank you to the many STIS staff and students who have taught me, provided opportunities, or been part of my community. Thank you to Margaret Acton, Géraldine Debard, Anne Valentine, and Issy, for underpinning this community. Many thanks to the wonderful early career researchers at STIS, including Debby Scott, Faye Wade, Farah Huzair, Dominic Berry, and James Lowe. To Thoko Kamwendo, thanks for the cathartic chats and for illuminating the path forwards. Thanks to Rob Smith for helping me to begin post-PhD life, and creating valuable spaces for thinking. Thank you to my long-time office mates, Justyna Bandola-Gill, Anna Kuslits, and Yazmin Morlet Corti. Thank you to those further ahead in their PhDs who shared wisdom, including Valeri Wiegel, Hung Nguyen, Michael Kattirtzi, Daniel Thorpe, Valentina Marcheselli, Mason Robbins, Matjaz Vidmar, Leah Gilman, and Chihwei Yeh. Thank you to many fellow PhD travellers as well, especially to Mareike Lührs.

I am indebted to many other friends, colleagues and teachers. Imogen Bevan, Claudia Benham, Julia Brown, Guy Leedon, Liz Ing-Simmons, Chris Mellingwood, Erica Mason, Naomi Penfold, Sigrid Robinson, Sarah Schönbauer, Zeba Talkhani, Toni Velikova, and Sam Vilkins: thank you variously for PhD wisdom, comments, insight, and practical help. I am grateful to my PLOS friends and colleagues; I was lucky to learn about open science in the most collegial of workplaces. Thanks especially to Sam Moore, who led the way from PLOS to "open" PhD and provided valuable comments. Thanks to UCL STS staff including Brian Balmer and Martin Savransky who shaped my early ideas on this topic. Thank you to Adam Mosley for directing me towards a wider set of historical perspectives. Deep gratitude to Joan Leach, Sujatha Raman and everyone at the Australian National Centre for the Public Awareness of Science, who hosted me on a PhD visit. Thanks to Ginny Barbour for talking to me about open access in Australia. Thank you to Claudia and Peter for so generously hosting and supporting me on this visit. To Lindy Orthia, and to other early academic mentors, some of whom are interviewees: thanks for teaching me, and for showing me different (good) ways to be a researcher.

Thank you to all my family and friends who have been so supportive and patient during these years. My sister and I completed our PhDs – hers in biology, mine about biologists – side by side. Thank you Tess for the writing retreats, the memes, the pep talks, the remote yoga, and most of all for your company all this time. So many thanks to my parents, Robert and Margherita, for their unconditional support of two daughters on doctoral journeys. Thank you both for the uplifting video chats, for your incredible patience and dedication reading drafts, and for being interested in what I am doing (and letting me ramble on about it). Finally, to Bryan: I don't have enough words. You have loved and supported me through every stage of this work, including in this final year when Covid-19 made everything so difficult. Thank you for being a daily reminder of what really matters.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADRC	Australian Research Data Commons
ANDS	Australian National Data Service
AOASG	Australasian Open Access Strategy Group
APC	Article Processing Charge
ARC	Australian Research Council
ARPANET	Advanced Research Projects Agency Network
BAME	Black, Asian and Minority Ethnic
BMC	BioMed Central (open access publisher)
BOAI	Budapest Open Access Initiative
CC BY	Creative Commons Attribution Licence
DMP	Data Management Plan
DNA	Deoxyribonucleic acid
DOAJ	Directory of Open Access Journals
DORA	San Francisco Declaration on Research Assessment
EC	European Commission
EMBL	European Molecular Biology Laboratory
ESRC	Economic and Social Research Council (UK)
FAIR	Findable, Accessible, Interoperable, Reusable
FOSTER	Facilitate Open Science Training for European Research
HEFCE	Higher Education Funding Council for England
HGP	Human Genome Project
IEG	Information Exchange Group

LERU	League of European Research Universities
MRC	Medical Research Council (UK)
NCBI	National Center for Biotechnology Information (US)
NHMRC	National Health and Medical Research Council (Australia)
NIH	National Institutes of Health (US)
OA	Open Access
ODI	Open Data Institute
OECD	Organisation for Economic Co-operation and Development
OKF	Open Knowledge Foundation
PDB	Protein Data Bank
PI	Principal Investigator
PLOS	Public Library of Science (open access publisher)
PMC	PubMed Central (open access digital repository)
QUT	Queensland University of Technology
RDA	Research Data Alliance
RDM	Research Data Management
RNA	Ribonucleic acid
ROAC	Radical Open Access Collective
REF	Research Excellence Framework
RCUK	Research Councils UK
SHERPA	Securing a Hybrid Environment for Research Preservation and Access
SSAC	Scottish Science Advisory Council
STS	Science and Technology Studies
UKRI	UK Research and Innovation

Chapter 1 | Introduction

Science has a special relationship with the term “open” and its connotations. I first encountered “open” in science in 2011, as a new biology graduate, when I became an employee of an open access publisher that sought to change the world: specifically, to make scientific knowledge free to anyone, online. I spent three years happily representing this vision of a better, more open science – mostly to scientists. Over time, I became curious about the weight of meaning carried by “open”: it seemed to speak for itself, to hold myriad hopes, and to be irreproachable. And yet, opening science was a battle: even as open science policies gained traction, scientists themselves often seemed to hold back. This research is the result of a curiosity about openness in science that got out of hand.

To introduce the topic of analysis, I will present six observations about openness in science that captured my interest early on, and convinced me that it was worthy of study. The first observation is **salience**: “open” discourses have rapidly risen to prominence in twenty-first century science, both bottom-up through scientific communities, and top-down through organisations that govern research. As these discourses lead to policy, they become salient to progressively larger populations of scientists. Secondly, I saw that open science carries **big claims and promises**: we are said to be in the midst of – or shortly expecting – an “open science revolution” comparable with the birth of modern science (Bartling and Friesike 2014b; Nielsen 2011; The Royal Society 2012), that transforms practices and cultures of science. And, as open science looks forward, it also looks back: my third observation is of **deep historical resonance**. Interpretations of historical sources and narratives suggest that science is defined by openness, and contemporary movements seek to restore and elevate this archetypal state. My fourth observation was of an absence: despite its salience, and the great promises it makes, open science is **hard to pin down**. It seems to shift enigmatically in meaning, to be more than the sum of its parts, and to resist definition. Whatever open science may be, however, it is taken to be **good science**: a virtuous epistemology. This is my fifth observation. Sixth and finally, there seemed to be a **gap in perspective** between advocates of open science – those invested in its virtue and promise, including scientist–activists – and a broader population of scientists. This constitutes a "cultural problem" for the imminent "revolution" towards scientific openness.

Having established my reasons for studying this topic, I will detail the **research questions** that become the common threads structuring this work. I will then outline my **theoretical position and methods**, and the **positioning of this research** in the literature**.** Finally, I explain the **thesis structure**.

1.1 The contemporary salience of "open" in science

"Open science" and related "open" discourses have risen from obscurity to become salient for researchers, policymakers, governments and (some) citizens worldwide in the space of only 15 to 20 years. Movement has often been, initially, bottom-up: from research, tech, and scholarly communication (BOAI 2002; Murray-Rust et al. 2010; OKF n.d.; PLOS 2000). It is now significantly top-down: from research funders, universities, learned societies, big publishers, and international organisations such as the European Commission, OECD, and G7 (EC 2016; G7 2017; HEFCE et al. 2016; LERU 2018; OECD 2014; The Royal Society 2012; Wellcome Trust n.d.). This rise in salience, which seems slow to many open science advocates, but is rapid in historical terms, is my first reason for considering it worthy of study. Here I sketch key developments marking the ascent of open science.

Established early within these discourses and ever more prominent is **open access** (OA): the concept and practice of making scientific and other scholarly research – especially in the form of research articles – free to access online. Liberal reuse rights are often considered an essential part of this definition (Eve 2014; Suber 2012). Consensus grew around this idea the early 2000s through a series of international agreements (Berlin Declaration 2003; Bethesda Statement 2003; BOAI 2002). This period of time was also marked in the global north by the establishment of PubMed Central – “a free archive of biomedical and life sciences journal literature”– by the US National Institutes of Health (NCBI n.d.); the mass signing by scientists of a petition supporting free access, and opposing restrictions by publishers (PLOS 2000); and the related founding of two internationally influential open access publishers, BioMed Central (BMC), and PLOS. Universities were also beginning to set up repositories for, and even to mandate, freely accessible research (e.g. QUT n.d.; Steele 2013). Grassroots activism continued: in 2011, Kazakhstani graduate student Alexandra Elbakyan launched of Sci-Hub, a “pirate” website that now provides free access to most scholarly literature (Himmelstein et al. 2018). The following year, UK mathematician Timothy Gowers started the Cost of Knowledge protest against Elsevier (Gowers 2012). And in 2013, the suicide of US computer programmer Aaron Swartz – who was facing harsh criminal charges brought by the US Attorney’s office for having systematically downloaded JSTOR articles for public release - brought tragic significance to the cause of open access (Greenwald 2013).

From the mid-2000s into the 2010s, major research funders and governments in some parts of the world began encouraging, and then insisting upon, open access for the work they fund (e.g. Finch 2012; NHMRC 2018; RCUK 2013; Wellcome Trust 2012). Big commercial publishers, including Springer and Elsevier, began to adapting to this landscape – and assimilating themselves into an “open” discourse – by providing options for open access amongst their otherwise subscription-access content (Björk 2012). Landmark open access policies continue to radiate and intensify. On a national scale, the UK’s next national research assessment – REF2021 – has required researchers to meet strict open access requirements since 2016 (HEFCE 2014). And on a global scale, “Plan S” – backed by an international alliance of research funders, including the World Health Organization, European Commission, Wellcome Trust, and UK Research and Innovation (UKRI) – is due, from 2021, to mandate immediate, liberally licensed open access on an unprecedented scale (European Science Foundation n.d.; Van Noorden 2020). Collectively, these developments will have significantly raised awareness of open access amongst researchers.

The idea of **open data** has also risen rapidly in salience in research contexts over this time period (Murray-Rust 2008; The Royal Society 2012). "Open" again refers to free online access with liberal reuse rights, and applies in this case to data underlying reported results. Conceptually, 'open data' is not completely separate from long-held practices of publishing evidence to support scientific results (Leonelli 2013); however, it is associated with relatively new technological capabilities to produce larger volumes of data, together with digital and online capability to store and communicate those data. The Human Genome Project (HGP), spanning the turn of the twenty-first century, can be positioned as a close forerunner of the contemporary open data phenomenon in research: it established in genomics a practice of making large-scale datasets publicly available online for reuse by a wider community (HGP Information Archive 1997). Momentum towards open data has built through bottom-up and top-down movements, as for open access. Funders and journals increasingly encourage or require researchers to make their underlying data openly available. In the UK, this stance is articulated in the *Concordat on Open Research Data* (HEFCE et al. 2016).

Open access and open data are two of the most prominent instantiations of "open", but do not encompass the whole phenomenon of **open science**. Fecher and Friesike (2014) characterise open science as an "umbrella term encompassing a multitude of assumptions about the future of knowledge creation and dissemination" (p. 17) and divide it into five schools of thought (infrastructure, public, measurement, democratic, pragmatic). Following a systematic literature review on the topic, Vicente-Saez and Martinez-Fuentes (2018) define open science as "transparent and accessible knowledge that is shared and developed through collaborative networks" (p. 1). These schemes and definitions include open access and open data, as well as movements and practices as diverse as: citizen science; science blogging; open peer review; open code; pre-registration; alternative metrics or reputation systems; networked collaborative tools and infrastructures; and open notebook science. Preprints – research manuscripts posted openly prior to publication – are also commonly associated with open science (Desjardins-Proulx et al. 2013). These tend to be grassroots and entrepreneurial movements, but the European Commission (2016) and other major institutional actors have played a role in solidifying and advocating the language and concept of "open science".

1.2 The "open science revolution": promises of a transformed future

"Open" discourses have not only become highly salient in the last two decades; they also tend contain grand, expansive claims about the future of science. Texts from both individuals and institutions conjure visions of an "open science" future that entail vast transformations to research practice and culture – transformations that are currently under way, or impending. One of the clearest instances appears in *Reinventing Discovery* (2011) by Michael Nielsen, who writes about a "second open science revolution extending and completing the first open science revolution, of the seventeenth and eighteenth centuries" (p. 184). The influence of this idea and language can be seen in its adoption by the Royal Society the following year, in their report *Science as an open enterprise* (2012): "The internet [...] may pave the way for a second open science revolution, as great as that triggered by the creation of the first scientific journals." (p. 7). Nielsen's language of revolution also appears in Bartling and Friesike's edited volume *Opening Science* (2014a), that begins with a chapter titled "Towards Another Scientific Revolution" (p. 3). Open science is thus envisioned as a transformation comparable in scale and profundity to the birth of modern science itself: an upending of centuries' worth of tradition.

This transformation is depicted not only as historically significant, but ambitious in scope: its trajectory is towards complete, or extreme, scientific openness. The Royal Society (2012) states: "We are now on the brink of an achievable aim: for all science literature to be online, for all of the data to be online and for the two to be interoperable" (p. 7). Even this superlative ambition is dwarfed by statements that extend beyond open access and open data. A 2018 report from the European Commission's Open Science Policy Platform included the following expectation:

> *"For Open Science to be successful, it must become embedded at every level and in every aspect of the scientific endeavour [...] Open Science requires a systemic shift in current practices to bring transparency across the system..."* (2018:4)

Years earlier, Nielsen (2009) had already conjured a vision of systematic scientific openness, which for him was an online, networked, techno-utopia:

> *"We should aim to create an open scientific culture where as much information as possible is moved out of people's heads and labs, onto the network and into tools that can help us structure and filter the information. This means everything – data, scientific opinions, questions, ideas, folk knowledge, workflows and everything else." (p. 32)*

In the mid-2000s, chemist Jean-Claude Bradley pioneered an approach of this kind: open notebook science "expos[es] a researcher's complete record of progress to the public in near real time" (Bradley, Owens, and Williams 2008:1). Bartling and Friesike (2014a) encourage their readers to imagine a world in which approaches like these have transformed science into a rapid, seamless, universal flow of knowledge:

> *"Picture a situation in which scientists would be able to publish all their thoughts, results, conclusions, data, and such as they occur, openly and widely available to everybody. [...] Imagine the potential for interactions between researchers. Knowledge could flow quickly, regardless of institutions and personal networks. Research results could be published as they occur. There would be no need to wait until results are complete enough to support a full paper." (p. 8)*

These enthusiastic renderings of completely open, revolutionised scientific futures add weight and intensity to the rising salience of open science – and make critical study of the topic particularly important.

1.3 The long history of scientific openness

My third observation was that contemporary open science discourse is oriented both to the future, and to the past. Nielsen (2011) characterises centuries worth of science as essentially open by describing the so-called scientific revolution – the birth of modern Western science – as the "first open science revolution" (p. 184). I found that this view was a common one, anchored in traditional narratives about the workings of science: if researchers collectively share their results, they can build upon the findings of others and a progressive, self-organising corpus of knowledge emerges (Polanyi 1962). Openness is thus understood as essential to – and definitive of – the function of science. The early sociologist of science Robert Merton became an influential source of this narrative when he identified “communism” (or “communalism”, Vanderstraeten and Eykens 2018) as a characteristic value in the ethos of science: “the substantive findings of science are a product of social collaboration and are assigned to the community” (Merton 1973 [1942]:274). The same idea is embedded in Newton’s aphorism, “If I have seen further, it is by standing on the shoulders of giants” (Newton 1675). This form of openness is recognisable to a wide audience. Polanyi (1962) writes that much of what he has to say “will be common knowledge among scientists” (p. 54). The Royal Society (2012) present this type of openness as a foundation for their contemporary advocacy of open science.

This observation suggested that openness resonates deeply through the history of science, piquing my curiosity about the relationship between historical openness and the rapid, enthusiastic embrace of contemporary open science. Chapters 3 and 4 became an in-depth exploration of this relationship, in which I draw both upon historians’ accounts of scientific openness, and upon broader socio-technological histories that contextualise the emergence of contemporary open science. Historical accounts (e.g. David 2008; Eamon 1985) show that openness was not an inevitable cultural achievement for science, but contingent on a particular history. For me, accounts like these also raised questions about the assumed meaning of openness in past science, and its significance in a contemporary context. This social–historical dimension was my reason for studying openness specifically in relation to the natural sciences, even though its relevance is not exclusive to scientific fields.

1.4 The elusive, expansive definition of open science

Despite its remarkably deep history, contemporary salience, and future promise, I found that it was difficult to summarise open science. The definitions I have cited (Section 1.1) are laudable, but have a nebulous, elusive quality. Another, from the EU-funded FOSTER Open Science project, is perhaps more grounded, but loses some scope as a result, despite its breadth: "Open science is the movement to make scientific research, data and dissemination accessible to all levels of an inquiring society" (FOSTER n.d.). Bartling and Friesike (2014b) mention their struggle to define open science and related "umbrella terms", and at one point collapse into circularity, describing open science as: "a scientific culture that is characterized by its openness. Scientists share results almost immediately and with a very wide audience" (p. 10). A later document that sets out an international strategy for open scholarship – a discipline-inclusive version of open science – eschews definition altogether, instead "recognis[ing] that it is a holistic term that encompasses many disciplines, practices, and principles" (Tennant et al. 2019:1).

To some degree, anchoring can be achieved by looking under the open science "umbrella", to concepts such as open access and open data – as I have done above. However, this is insufficient as a summary, since the umbrella continually expands to include additional "open" concepts, movements and practices. As Catriona MacCallum, Director of Open Science at the publisher Hindawi, put it in interview: "…open science is the box, and it's very hard to define what goes in that box, because I think the goalposts are continually shifting". The contents of the "open" box do have a concrete feature in common, however: they tend to be enabled by digital and internet technologies. My early observations of open science suggested it was also valued for abstract qualities that made it more than the sum of its parts. The PLOS website at one time described "open" as "a mindset that represents the best scientific values" (PLOS 2017). Tennant et al. (2019) proposed an open scholarship strategy based on shared values and principles, despite acknowledging "numerous competing, parallel, or overlapping definitions of what Open Scholarship comprises in terms of both research principles and practice" (p. 14).

This elusive, expansive, holistic quality of open science made it interesting to me as a topic for social research. The meaning of open science seems to be actively under construction and negotiation through its rapid rise to salience, as more and more social groups take part in shaping it. As Fecher and Friesike (2014) note, "each peer group discussing the term has a different understanding of the meaning and application of Open Science" (p. 18). I reasoned that we are at an illuminating historical time for the study of scientific openness and its construction by different groups, as consensus has not been reached: it remains messy, unsettled, out in the "open". It also occurred to me that the elusive expansiveness of open science may perform and reflect a social function: it leaves room for new meaning, as specific definitions quickly become obsolete. However, my fifth observation suggested that "open" in science was not an empty signifier.

1.5 "Open science is just good science!"

Through my involvement with open publishing, I observed that "open" tended to signal – without explanation, at least in some social settings – the worth of a knowledge practice, and the virtue of those pursuing it. This has been made explicit in some advocacy communication, including the PLOS' "mindset" statement in the previous section (PLOS 2017). Similarly, Jon Tennant, an influential advocate of open science, used the phrases "open science is just good science" and "open science: just science done right" to make his case on social media and in talks (e.g. Tennant 2018). This may have been inspired by bioinformatician Mick Watson's (2015) article titled *When will 'open science' become simply 'science'*, in which he concludes, "open science isn't a movement, it's just (good) science" (p. 2). Screenshots of these three illustrative cases are shown in **Figure 1**.

In these examples, "open" has moral overtones because it considered the "best" or "right" way to do science. For me, this evoked the work of historians of science Daston and Galison (2007), who use the term "epistemic virtue" to capture a quality in which knowledge-making and morality are inseparable: a truth-making quality that comes, in social and historical context, to be taken for granted. Daston and Galison conceptualise objectivity as an epistemic virtue, and trace its emergence in the nineteenth century. I thought it would be particularly illuminating – and novel – to explore openness through the lens of epistemic virtue: to consider, that is, whether we are currently witnessing a cultural shift that redefines good science as open science. Epistemic virtue as rendered by Daston and Galison also places emphasis on the role of the scientific self, showing that such cultural shifts alter not only a social consensus about qualities of good science – but also collective, internalised understandings of the virtues that make *good scientists*.

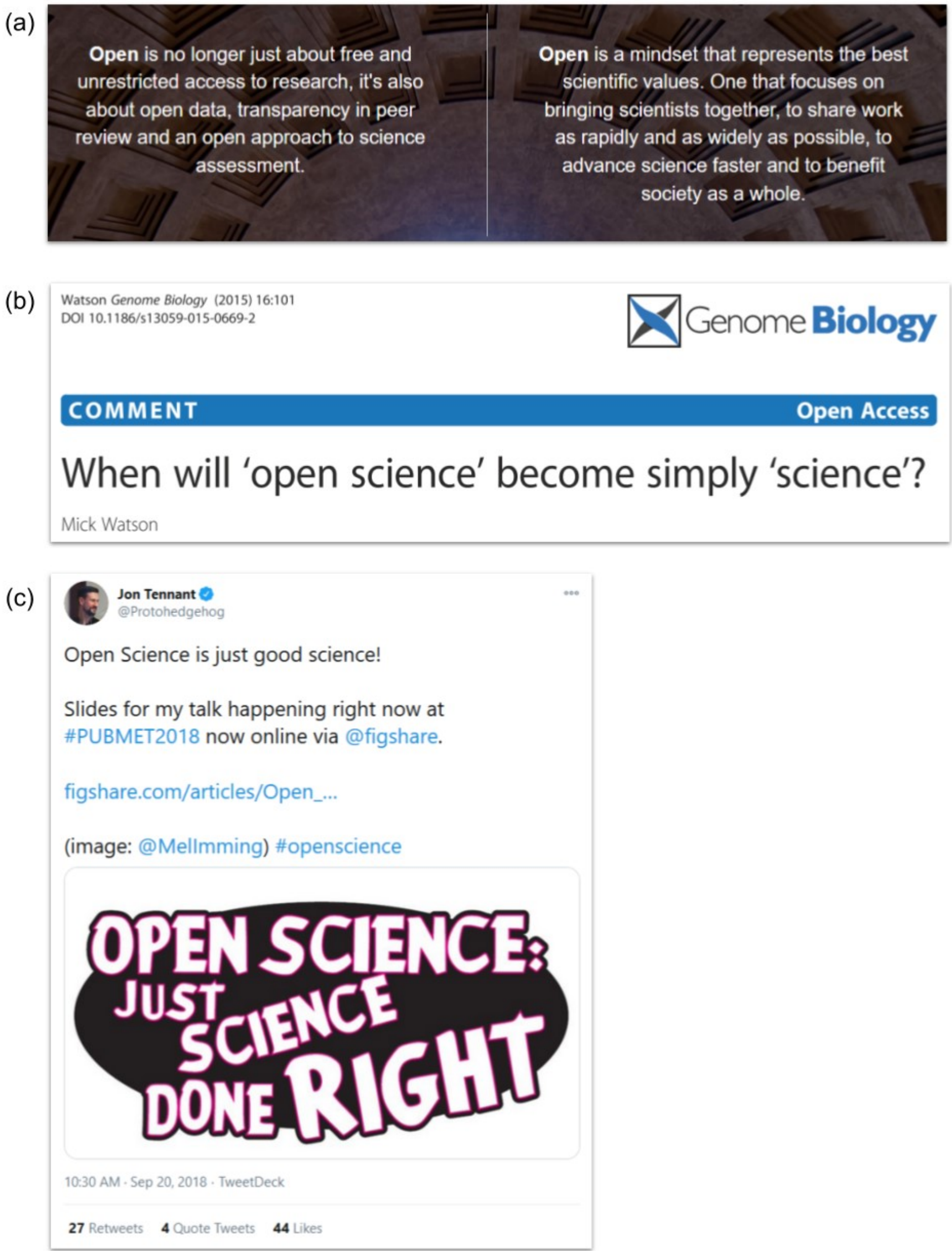


Figure 1 | Examples of discourse that frame open science as good science – a part or extension of doing science well, "right", or in "the best" way. (a) A screenshot from *Who We Are* PLOS website several years

ago (PLOS 2017); (b) The start of Mick Watson's (2015) article in which he argues that open science is "just (good) science" (p. 2); (c) Open science advocate Jon Tennant's use of a similar framing.

1.6 A gap between advocates and scientists

This final observation was my human and pragmatic motivation for conducting this study. I noted, again from my involvement in open publishing, that scientists were among those most enthusiastically leading open science movements. I also noted that a many other scientists – including those with whom I interacted in the course of managing an open access journal – were less invested in "open", and seemed disconnected from its promises of a better future for science. I am far from alone in making this observation: it seemed to be a common refrain amongst open science advocates[1], who routinely encounter ambivalence or resistance from researchers. For example, in their international open scholarship strategy, Tennant et al. (2019) have a section on "research awareness and apathy", in which they note that "awareness of Open Scholarship is still often very low among certain research communities" (p. 30). This is in line with a study by Proctor et al. (2010) which, nearly a decade earlier, found that "only a small group of enthusiastic open researchers – 5% of our respondents – publish their outputs and their work in progress openly, using blogs and other tools" (p. 6). One of many related comments that I heard in the course of this research (at an open access event) was, from an advocate: "researchers need constant and repetitive advocacy of the way the world is changing"[2].

This observation is connected with a widespread advocacy and policy discourse in which progress towards open science is framed as a "cultural" issue. Nielsen (2009) spoke of "major cultural barriers preventing scientists from getting involved" (p. 31), and much more recently, the League of European Research Universities produced a

[1] "Advocates" come from many overlapping domains including (scientific) research, library, publishing, open source, digital innovation, and research funding/policy communities. Experiences of open science advocacy are diverse, but seem to have some common threads. For more detail about the advocates/policymakers I interviewed, see Section 2.3.4.7.
[2] Recorded in my notes from the event.

"roadmap for cultural change" (LERU 2018) to guide universities on practical steps towards embracing open science. It includes the statement:

> *"There are real dangers in trying to introduce new practices without carrying the academic community with the leaders of those changes [...] In many ways, cultural change is the most difficult outcome to achieve in embracing Open Science..." (p. 21)*

These discourses oscillate between identifying researchers themselves as cultural barriers, and laying blame upon a system of incentives governing researcher behaviour, sometimes combining the two arguments. For instance, Ali-Khan et al. (2017) describe the provision of appropriate incentives for individual researchers as the "key-rate limiting step" in successful implementation of open science (p. 1), and Tennant et al. (2019) at one point comment that much is "dependent on the willingness of researchers themselves to contribute to scholarly research in an open, collaborative, and collective manner", which is "largely down to whether such researchers perceive this process as being advantageous to them in some way" (p. 23). However, open advocate Cameron Neylon has pointed out that despite the common "desire for cultural change" from a policy perspective, "there is little engagement with the concept of culture" (Neylon 2017:2).

I thus considered it especially important to study open science from a social science perspective that might shed further light on these "cultural" issues. Moreover, it seemed to me that the gap in perspective between advocates and (many) scientists represented a gap in knowledge about scientists' experiences and constructions of scientific openness. To me, the assumption that scientists' ambivalence or resistance stems from a lack of appropriate awareness, knowledge, or incentives was reminiscent of assumptions made by scientists about the public in the 1980s as part of the UK's public understanding of science (PUS) movement: that if only the public understood science and technology, they would be inclined to trust and accept it (Yearley 2005:102–14). Sociologists of science countered that public disengagement might not reflect an *absence* of understanding, but the presence of other understandings, such as knowledge – and distrust – of institutions representing science; and forms of public knowledge with particular relevance in local contexts (ibid.). Open science turns the tables: (many) scientists are disengaged, but have local, contextual knowledge; and open advocates are technical experts recommending "revolution" of scientists' professions and lives. The politics of this analogy might be debated, but it suggests that scientists' absence of engagement could be productively understood as a complex, context-informed position, mediated by trust in institutions, with its own understandings of openness in science. I reasoned that a broad-minded exploration of scientists' positions on this topic might "open" up new and fruitful narratives and conversations between scientists, advocates, and policymakers.

1.7 The research questions

My six observations above show that open science is an interesting, timely, and important topic, on which significant insight might be gained from a social studies of science perspective. I distilled these observations into a series of research questions that that guide my investigations and shape this thesis. I have one broad, overarching research question that addresses the common factor in my observations: that the meaning of openness in science is actively under construction at present through multiple discourses and interactions. This question allows that issue to be explored in an open-ended way and has the capacity to receive unexpected responses:

1. How is the meaning of "open" (or "openness") being constructed in the context of science?

My attention is then further guided and structured by a series of sub-questions:

a) How do contemporary "open science" discourses relate to older, historical discourses about the essential value of openness in science?
This question guides exploration of my third observation (p. 6), that openness in science has deep historical resonances that appear in contemporary discourses.

b) How has "open" in science recently gained such salience and discursive power, despite its multiple meanings and lack of specificity?
This question notes my first and second observations (pp. 2, 5) about the rapidly rising salience and future promise of open science, and guides exploration of my

fourth observation (p. 7) – of elusive, expansive qualities – in that context.

c) To what extent is "open" (or "openness") being constructed as an epistemic virtue?
This question is sensitised by my fifth observation (p. 9) that the rise to salience of open science may represent a shift in assumed qualities of "good science".

d) Why does it appear that many academic scientists are disengaged from "open" discourses and practices?
This question guides exploration of my sixth observation (p. 11), of a gap in perception between advocates and (many) scientists on this topic.

1.8 Theoretical approach and methods

A detailed explanation of my approach is given in Chapter 2. The research is broadly situated by the field of science and technology studies (STS), according to which science – its institutions, ways of working, and knowledge – is considered to be "irredeemably human" (Edge 1995:5). Beyond this, my study is *grounded* in the sense that it emphasises data exploration over the application of any detailed, prescribed theoretical framework. My analytical attention is guided, however, by the observations and research questions above, and in particular the idea of epistemic virtue (Daston and Galison 2007), which I adopt as a sensitising concept.

In order to explore meaning in a broad way, I adopted two main qualitative methods:

1. **Analysis of academic, advocacy and policy documents** that have shaped the meaning of scientific openness both recently and over a longer history.
2. **Semi-structured interviews** with 40 biological scientists and 14 open science advocates and policymakers.

The first of these methods produces some novel analysis, blended with a more traditional literature review process (Chapters 3-5), and the second method produces novel data and forms the empirical core of the thesis (Chapters 6-8). The interview study was largely set the UK and Australia, and thus had most relevance to open science movements and research settings in these countries, with some relevance to an international – but particularly global north and Anglophone – context. Biological scientist interviewees were based at research-focused, prestige-oriented academic institutions. Selection of scientists was purposive, aiming to include diversity in sub-discipline, gender, race and ethnicity, research group size, and extent of commercial affiliation. This was with the intention of representing a broad range of perspectives on scientific openness, within the study scope. Interviews with advocates and policymakers added context from fields including scholarly communication and publishing; freelance research, advocacy, and consultancy; university leadership, administration and libraries; and research funding and policy.

1.9 Positioning this research in the literature

A very wide range of literatures are conceivably relevant to my analysis. There are new and fast-growing academic and policy literatures on different aspects of the topic, from fields as diverse as economics; life sciences; science communication; history of science; library, management, communication, and education studies; and many interdisciplinary fields. This is in part because open science is concerned with a transformation of research. As a result, scholars in many fields are increasingly advocating and/or analysing this idea through media familiar to them: journal articles and books. Only a small proportion of this literature, however, takes a step back to question broad social and historical aspects of open science as a holistic phenomenon, as I hope to do. The intersection between STS and studies of open science in particular is sparsely populated. However, there are an increasing number of studies that occupy these critical spaces, and they have proven very valuable as I have worked towards my own contribution.

Studies from several fields have analysed open science as a holistic, contemporary phenomenon, and worked towards defining or classifying it, usually with normative aims (e.g. Bowman and Keene 2018; Fecher and Friesike 2014; Vicente-Saez and Martinez-Fuentes 2018). Some authors have also contributed holistic social, historical and political analyses of the whole phenomenon (David 2008, 2014; e.g. Delfanti 2010, 2013; Strasser 2019; Willinsky 2005). Kelty (2012) and Grand et al. (2012) illuminate open science by bringing it into conversation with concepts familiar in STS: moral economies, and trust technologies, respectively. There are also, increasingly, wide-ranging critical analyses on particular open movements, especially open access and its history in a publishing context (Eve 2014; Fyfe et al. 2017; Guédon 2001; Lawson 2019; Moore 2017, 2019a), and open data (Gabrielsen 2020; Leonelli 2013). Šimukovič (2020) argues that STS is well-suited to address the topic of open access, but that so far, scholars in the field have shown only limited interest. Where many – or most – critical analyses come from an advocacy position, Mirowski (2018) instead delivers a scathing critique of open science as an "artifact of the current neoliberal regime of science" (p. 172). He makes valuable STS-relevant points, confirming many of my impressions: about the ambiguity and multiplicity of open science discourse; its tendency to technological determinism; and capacity to reinforce power asymmetries.

I aim to contribute to this growing literature an analysis of open science is sensitised to both breadth (its holistic, expansive qualities) and depth (its historical, social nuances), and brings with it an STS sensibility (an alertness to human entanglements of science and technology). I also contribute a perspective that is not primarily normative or aimed towards open science advocacy – although of course, I am not a neutral analyst. Nor is my perspective shaped by an antipathy towards open movements as is shown by Mirowski (2018). Instead, have come to this research through close professional and personal experience with both biological science and open science (see Reflexivity, Section 2.3.10) and my aim has been to care for both of these worlds through my work (Puig de la Bellacasa 2011).

An empathetic approach is particularly important because the main empirical contribution I will make is on (biological) scientists' understandings of openness in science. In this area, my work is aimed towards a clear gap. There are a wealth of studies that characterise or measure researchers' practices and attitudes in relation to particular categories of openness (e.g. Cragin et al. 2010; Piwowar 2011; Wallis, Rolando, and Borgman 2013; Weckowska et al. 2017), but very little prior research that: i) explores what "open" or "openness" in science means to researchers in a broad way; ii) captures and prioritises qualitative depth. In recent years there have been some in-depth qualitative studies of the kind I believe to be important, most of which have been published after I set out on this PhD: those I am aware of are by Scheliga and Friesike (2014), Levin and Leonelli (2016), Levin et al. (2016), Van den Eynden et al. (2016; commissioned by the Wellcome Trust), and Ali-Khan et al. (2017). These studies provide valuable, relevant findings, many of which resonate with my own, through qualitative interviews or focus groups. Several of these studies approached scientific openness in an exploratory way, allowing participants to frame the issue, and each study brings contextual insights.

1.9.1 Position in relation to similar studies

Levin et al. (2016) in particular employed a framing and methodology similar to my interview study: through semi-structured interviews they asked 22 principal investigators (PIs) in biomedical fields about "their understanding of 'openness' in science and experiences with Open Access, Open Data, and Open Innovation" (p. 3). The article reports key themes from these interviews; additionally, Levin and Leonelli (2016) analysed specific elements of these interviews in more depth. The latter article conceptualises openness "as a *mode of valuing* the research process and its outputs" (p.1, original emphasis) in which policies codify the value of certain "open" outputs and practices, often in tension with the more varied, nuanced and intangible ways that researchers value research as a process. These papers are an excellent comparison point for my study findings; I explore the similarities and differences in Section 9.6.

However, it should be noted that this study was conceived independently from my own and there are key differences in theoretical lens and method as well as findings. Where Levin and Leonelli (ibid.) write about *value* in relation to openness, I use the concept of *epistemic virtue* (Daston and Galison 2007). Both theoretical lenses are clearly relevant to the context and add insight from different perspectives. They are also consistent with one another – epistemic virtue is a particular lens through which to understand how value in knowledge-making is assigned and performed. I chose epistemic virtue as a sensitising lens because it adds extra depth to my understanding of scientists' relationships with – and subjective experiences of – openness, due to its foregrounding of the scientific self and accordingly, identity. Thus, my research explores openness not only as a mode of *valuing* science and its outputs, but a mode of *being* in science, in which the person doing the "valuing" comes into focus. From this perspective, scientists are valuing not only aspects of research process and output, but valuing and cultivating aspects of their scientific self; their idea and feeling of being a good scientist. The lens of epistemic virtue also aids understanding of the moralised quality of "open" discourse, and contextualises ways of valuing openness within a long social history of science, in which other knowledge-making "virtues" co-exist and arise contingently in different times and places (for more on epistemic virtue, see Section 2.2.1).

Other comparable studies (e.g. Ali-Khan et al. 2017; Van den Eynden et al. 2016; Scheliga and Friesike 2014) tend to have a narrower framing methodologically, initially asking participants about "open science" or "open research" thus evoking established concepts rather than adopting the broader framing of "openness", which makes a key difference to responses[3]. Additionally, comparable studies including Levin et al. (2016), either i) speak only to participants with an existing connection to open science, or ii) are conducted from the advocacy perspective of a funder or institution in relation to their open science policies. In contrast, I contribute a study with a degree of distance from institutional open science agendas that includes participants without existing connections to open science. This allowed me to elicit and depict aspects of researchers' stories that tend to be obscured by a focus on incentivising open science.

[3] See Section 8.3 for an analysis of how scientist interviewees understood "open science" differently from "openness in science".

Approaches to analysis also differ across this set of similar studies: all employ some kind of qualitative, thematic approach, but the degree of conceptualisation around and between these themes varies, as does the theoretical lens. My analysis adds in-depth conceptualisation through a focus on three overarching themes (open access, data openness, and interpersonal openness). By sensitising to epistemic virtue, with an attendant focus on scientific self and openness *as framed and understood by scientists,* I contribute an analysis sensitive to diversity, complexity and affect in scientists' relationships with (different types of) openness. This allowed unearthing of systemic factors – beyond immediate "drivers" or "obstacles" – that condition engagement with open practices. My study framing, choice of participants and analytical approach also led to the identification of a novel category of openness (interpersonal openness, Chapter 8) that was important to scientists but not salient either in advocacy or policy framings, or in other similar studies.

1.10 Chapter outlines

Following this Introduction, the thesis comprises eight chapters. Chapter 2 details my theoretical approach and methods, which are relevant to all of the subsequent chapters. It includes an explanation of my sensitising concept – epistemic virtue – its relevance in the study context, and its positioning within a broader STS perspective. The rest of the chapter explains how I chose and carried out the methods: most attention is given to the semi-structured interviews, which generate new data, but I also address document analysis and an incidental role for ethnographic observation. Detail is provided on the sampling and characteristics of interviewees, the interview procedure, ethical considerations, analysis process, limitations, and reflexivity.

Chapters 3-5 provide in-depth context on the study topic that situates the interview study. These chapters begin to address directly the research questions and establish findings, by blending literature review with an analysis of documents as empirical materials. In **Chapter 3** I consider the meaning of scientific openness in a centuries-long historical context. This addresses research sub-question (a) about the relationship between contemporary "open science" and older, historical discourses about the essential openness of science. In **Chapter 4** I take a closer look at developments of the twentieth century as background for the emergence of contemporary "open science" in the 2000s. This allows me to address research sub-question (b), which asks how "open" in science has recently gained such salience and discursive power, despite its multiple meanings and lack of specificity. In **Chapter 5**, I move to a narrower focus on the twenty-first century, and the way in which open access, open data, and overarching open science movements have unfolded in the last two decades. I focus particularly on these topics owing to their relevance in the interview study, and I focus on UK and Australian contexts. I illustrate the unfolding of these movements through nine case studies of advocacy and policy documents, which I analyse in relation to all of the research sub-questions.

Chapters 6-8 present the main empirical research findings, structured according to the three most salient categories of scientific openness raised by biological scientist interviewees: open access (**Chapter 6**), data openness (**Chapter 7**), and interpersonal openness (**Chapter 8**). The first two categories are closely linked with policy and advocacy discourse; the third emerges entirely from scientists' reports of their experiences and their implicit conceptualisations of "openness in science". I characterise the main themes and patterns in scientists' constructions of each of these openness categories. In particular I consider the range of relationships that scientists seem to have with each category; the attitudes, emotions and experiences of scientific self that accompany these relationships; and the contextual factors that seem to mediate these relationships.

Chapter 9 is a discussion and conclusion in which I bring the three categories of scientific openness considered in Chapters 6-8 into conversation with one another. This sheds light on differences and similarities between kinds of scientific openness as experienced by scientist interviewees. I draw on my sensitising concept, epistemic virtue, as a way of theorising these differences and similarities, and consider whether openness may encompass a variety of distinct epistemic virtues that have different roles in the present "open science revolution". In this chapter – though foreshadowed in those leading up to it – I directly address research sub-questions (c) and (d) that consider the extent to which openness is constructed as an epistemic virtue, and the reasons why scientists may often be disengaged from "open" discourses.

Chapter 2 | "Views from somewhere"
Theoretical position and methods

2.1 Introduction

This thesis is primarily composed of chapters that analyse and synthesise qualitative data about scientific openness, whether these data are in the form of academic literature, policy and advocacy documents, or interview responses (Chapters 3-9). This chapter explains my theoretical and practical approaches to these upcoming chapters. The first part of the chapter addresses my **theoretical position**, which is primarily a *grounded* one, meaning that I do not bring an established theoretical framework to this research against which results are tested or interpreted. Instead this work is exploratory, and aims to be receptive to new ways of thinking on a topic that has not typically been explored in a broad, qualitative, "open" way. My research questions are designed to be receptive to the unexpected, whilst also focusing attention on areas I know to be interesting, as introduced in Chapter 1. I also introduce a *sensitising concept*: not a fixed theoretical position, but an idea to that can be brought into fruitful conversation with all of the research data. This concept is *epistemic virtue* as conceived by Daston and Galison (2007). Earlier parts of this chapter explain this concept and its relevance to scientific openness.

The second, larger part of this chapter is a guide to my **methodological choices** and **methods**. This section addresses document analysis, as well as the empirical core of the thesis, an interview study. I explain how and why I recruited certain interviewees, and summarise features of the interviewee group. I draw attention to methodological choices that I made and their intended effects, describe my interviewing procedure, and address ethical considerations as well as – reflexively – my own "openness". I also explain my approach to transcription and qualitative analysis.

The chapter finishes with **reflexivity** section in which I lay out the perspective from which I conceived and carried out this research. It shows how – from a feminist STS perspective – the knowledge I produce is not a universal truth, but a "view from somewhere" (Haraway 1988:590). Haraway proposes a feminist objectivity in which knowledge can be understood as *situated*, using vision as a metaphor:

> *"Feminist objectivity is about limited location and situated knowledge, not about transcendence and splitting of subject and object. It allows us to become answerable for what we learn to see [...] There is no unmediated photograph or passive camera obscura in scientific accounts of bodies and machines; there are only highly specific visual possibilities, each with a wonderfully detailed, active, partial way of organizing worlds." (p. 583)*

It is this accountability for what I have learned to see that I seek to cultivate through reflexivity and other forms of care in my research.

2.2 Theoretical position

My study is broadly situated in the interdisciplinary field of science and technology studies (STS), which began to take form in the 1960s, and is diverse both in its disciplinary origins – which include history and philosophy of science, sociology, and anthropology – and in the range of theoretical and methodological approaches with which it is associated (Edge 1995). There is a common focus in STS, however, upon the analysis of science as a social phenomenon:

> *"...STS holds out the "new" view of science and technology as essentially and irredeemably* human *(and hence social) enterprises— both in the context that nourishes, supports, and directs them and in their inner character. And this is a triumphant, positive humanism: not the miserable confession that "scientists are only human" because you can catch them making mistakes, getting angry, being secretive and fraudulent." (ibid., p. 5, original emphasis)*

An acknowledgement of scientists' humanity extends to an understanding in STS that knowledge is constructed through social processes. This means challenging traditional boundaries: between the social and the technical; science and society, nature and culture; facts and values; and objective and subjective. For any readers who are scientists – an identity that I and many STS scholars have also occupied – this position can be provocative, but need not be so. It does not mean that all knowledge is relative or flawed, or that scientists are irredeemably biased, but that knowledge can be better understood as entangled with the circumstances of its creation: as *situated* (Haraway 1988), in the way that I conceive of my own research.

I take this broad positioning from STS into my work while not aligning with any specific theory, so as not to bring a prescribed worldview to my study. Instead I intended this study as a *grounded* one: guided by exploration, situated analysis, and the building up of concepts from empirical data, rather than by deduction or explanation from theory. The other reason why I do not adopt a particular STS theory is that understanding *knowledge* is not the purpose of my research, although its socio-technical quality is assumed and adjacent. Instead, my interest is in overtly social aspects of science: scientists; research communities, cultures, institutions and funders; and the scientific norms, ethos, and practices prevalent in these contexts. This is arguably an old-fashioned approach reminiscent of work by pre-STS sociologist of science Robert Merton (1973 [1942]): I will argue in due course why a return to Merton's topics of interest, with a new STS sensibility, is warranted in the case of scientific openness.

Although it can be difficult to grasp how *knowledge* is socially constructed, this is not the case for a concept like "open science" that clearly – as I show in Chapter 1 – is value-laden, actively evolving, shaped by a multiplicity of specific historical and cultural influences. Part of my theoretical position is to acknowledge that I am not attempting to detect an essential nature of "open" in science – or to dictate what it *should* mean – but to study its meaning and significance, in a particular context, to those who shape and are impacted by it. Many truths can be told about openness in science: I aim to tell several that I believe to be useful and underrepresented. Whilst I am not working within a specific theoretical framework, I do adopt a *sensitising concept* (Blumer 1954) that "suggest[s] directions along which to look" rather than "providing prescriptions of what to see" (p. 7): *epistemic virtue*.

2.2.1 Sensitising concept: epistemic virtue

Epistemic virtue is one of many potential framings for scientific openness, and I adopted it early on, more as a provocation than a solution. I found that it added depth and insight to my interpretations of scientists' perspectives. I refer to epistemic virtue as characterised by Daston and Galison (2007): an ethos implicit within a way of knowing. The concept proposes that epistemology is value-laden; that it is not possible to construct knowledge without a moral compass, which itself is socially constructed. It therefore by definition relates knowledge-making closely to the person producing that knowledge – the "knower".

Daston and Galison's virtue of interest is objectivity: the removal of oneself and one's biases from a scientific process in the belief that a better, more truthful kind of knowledge will result. The authors' detailed study of visual representation in the history of science locate an origin for this virtue in the mid-nineteenth century, when they argue that a new relationship emerged between knower and knowledge. This emergence related to broader socio-technical changes but could be traced in the intimate domain of the self. It was a shift from a passive self, flooded with sensations, that could be corrected by virtues of focus and selectivity ("truth-to-nature"); to an active, integrated self that required self-restraint ("mechanical objectivity"):

> *"The true savant was a 'genius of observation' whose directed and critical exercise of attention could extract truth-to-nature from numerous impressions, as the smelter extracts pure metal from ore. In contrast, the subjective self of nineteenth-century scientists was viewed as overactive and prone to impose its preconceptions and pet hypotheses on data. Therefore, these scientists strove for a self-denying passivity…" (p. 203)*

Daston and Galison give a visual example of a droplet of liquid falling onto a surface, depicted in the 1870s as perfectly symmetrical, and in the 1890s – with the aid of photography – as irregular and varied. These images were produced by the same scientist, British physicist Arthur Worthington, who viewed each as the most truthful depiction in its time (ibid., pp. 11-16).

Epistemic virtue is a subtle, powerful idea because it can conceive of an inward striving by scientists towards methods and attitudes of "good" science, in combination with a performative rhetorical endorsement of these principles (the latter is often noted in STS, e.g. Gieryn 1983). Daston and Galison (2007) write that "what had originally struck us as an oddly moralizing tone [...] now made sense" (p. 39), because the creation of knowledge was simultaneously a technical matter and a responsibility. This is an important acknowledgement because – ironically – of objectivity's cultural legacy: the best science is often assumed to be a "view from nowhere" in Haraway's (1988) terms: valueless, disembodied, amoral. To contend that the opposite is true assigns appropriate credit to scientists, who knowingly or unknowingly labour under the weight of techno-moral expectation. It also exposes the moral framework of science to appropriate scrutiny. For as Daston and Galison make clear, epistemic virtues are not immutable: they change and accrue over time, and are socio-technologically entangled.

Ideas related to epistemic virtue are not without precedent, even in the earliest social study of science. Merton (1973 [1942]) famously identified four norms underpinning an "ethos of science", and wrote that "they are moral as well as technical prescriptions" (p. 270). However, many analysts later took Merton's account to be a false empirical description of science (Etzkowitz and Leydesdorff 2000:116), or argued that he was crafting an unimpeachable image for science out of a much more ambivalent reality (Mulkay 1976). Studies of norms and ethos fell out of favour in STS, and were treated as "performative utterances of a rather old fashioned scientific culture that was disrobed in the 1970s and 1980s..." (Holden 2014:4). Merton did have a strong motive for presenting science in a way that promoted its purity and asserted its autonomy: he was writing in the context of Nazi science and Lysenkoism (Etzkowitz and Leydesdorff 2000). However, with the dismissal of his normative framework, the idea of a morality embedded in epistemology – and its role in scientists' own images of their work – fell into neglect (Holden 2014).

By overcoming some key limitations, Daston and Galison revive valuable aspects of Merton's thesis in a manner that is compatible with present-day STS scholarship.

Firstly, they remove any hint of essentialism. While Merton implied that “the” ethos of modern science existed, and could be fulfilled to greater or lesser extent by institutional conditions, Daston and Galison show that just as scientific knowledge is constructed in a social context, so too is its moral framework. The authors’ rich theoretical language can speak to a contemporary STS audience. Secondly – although this is arguably not Merton's failing – they make it clear that epistemic virtues are not descriptions of scientific practice, nor are they straightforward behavioural outcomes of rules and reinforcement. A final important difference is the theoretical emphasis that Daston and Galison place on the epistemic subject: the scientist.

2.2.2 Theorising the scientific self

The concept of epistemic virtue centres the “knower” who strives towards the virtue in question for themselves and the knowledge they produce. This is a valuable contribution of Daston and Galison's concept, because the person of the scientist and their internal experience is under-theorised in STS. Merton (1973 [1942]) did not neglect this area despite his focus on institutional structure and function. He made the following telling comment about the mechanism by which his norms operate:

> *“These imperatives, transmitted by precept and example and reinforced by sanctions are in varying degrees internalized by the scientist, thus fashioning his scientific conscience...” (p. 269)*

Post-Merton, STS has largely turned its attention to the construction of scientific knowledge, rather than those producing it. Actor-network theory, for example, has made a point of a radical theoretical symmetry that breaks down the social and treats human and non-human “actants” similarly in explanations of scientific knowledge production (Latour 2005). It is only more recently that scientists’ subjective experiences have come back into focus, this time as dynamic constructions alongside scientific knowledge:

> *“The primary concern of most science studies accounts is to understand how such apparatus shapes the kind of scientific knowledge that is produced. Making the production of the epistemic subject the focus of empirical investigation highlights how ‘machineries of knowing’* (Knorr Cetina 1999) *function as disciplining*

> *technologies that constitute both the subjects and objects of epistemic work."* (Bulpin and Molyneux-Hodgson 2013:93)

Scientific self in Daston and Galison's account refers to a way of being a scientist that is enabled and constrained by history and context – "the character and conduct of the scientist as a recognisable human type" (p. 198).

According to Daston and Galison (2007) and Bulpin and Molyneux-Hodgson (2013), the construction of the self is a daily, micro-act in the lives of scientists as everyday discourses and practices of science continuously shape inner life and bodily habit. Both sets of authors draw upon Foucauldian concepts to theorise this process:

> *"We pay close attention to what the historian-philosopher Michel Foucault called "technologies of the self": practices of the mind and body (more often the two in tandem) that mold and maintain a certain kind of self."* (Daston and Galison 2007:198–99)

> *"…institutional routines, discursive and material practices of particular epistemic settings are not only performed by scientists in the everyday work of knowledge production, but also operate as mechanisms of power. These apparatus organise, regulate and discipline the ways in which scientists think, behave and act."* (Bulpin and Molyneux-Hodgson 2013:93)

If a scientific self is characterised by certain ways of thinking, feeling and acting, constructed by immediate, day-to-day experiences and filtered through wider social influences, then the question arises of how this entity relates to wider epistemic cultures and institutional settings. These dynamics are not fully explicated by Daston and Galison, who focus upon "intrinsic" explanations, in which "explanans and explanandum" are "on the same level" (i.e. the self: pp. 197, 205). However, the professionalization and institutionalisation of science was only just beginning around the time at which Daston and Galison place the construction of objectivity, in the nineteenth century (Pandora and Rader 2008), whereas now multiple institutions – including universities, national research assessors, funding bodies and publishers – as well as larger-scale collaborative dynamics, govern the daily and career-long experiences of academic scientists (Felt 2009; Vermeulen, Parker, and Penders 2013).

Bulpin and Molyneux-Hodgson (2013) provide an example of how the practices, values and identities of science students are "disciplined" in larger community and institutional settings, in the case of the International Genetically Engineered Machine (iGEM) competition. This is not one-way, deterministic process: these authors document occasional resistance to incentives to standardize biological parts, where students instead prioritise the local needs of their project (pp. 100-101). This shows how the relationship between community/institutional structures and scientific self is plural and negotiated. Consistent with this, Savransky (2014) argues that "the process of subjectification is more complex and unstable that the analytics of [Foucauldian] governmentality suggest" (p. 106), and that a messiness and multiplicity of governing influences creates space for resistance, agency and perhaps systemic change.

In the current context, this allows one to conceive of scientific selves (and the epistemic virtues they uphold) both as products of social systems – e.g. laboratory groups, scientific communities, or universities – and as agents of change who may play a role in reconfiguring them. The sociologist Margaret Archer (2003) also argues that subjects, while influenced by the opportunities and constraints of social structures, possess agency to resist or change those structures. Further, she argues that it is their subjectivity – or "internal conversation" – that mediates this relationship between structure and agency. Archer's perspective is useful for understanding of how the scientific self might be both personally meaningful, and an explanatory factor in continuities or transitions in science's moral framework.

2.2.3 Openness as a candidate epistemic virtue

I have introduced epistemic virtue as a sensitising concept because of the striking ways in which scientific openness is currently framed (introduced in Chapter 1, and analysed in Chapter 5). One such framing is the phrase "open science is just good science", popularised by open science advocate Jon Tennant (2018). Such rhetoric is direct in its claim that "open" is, or must become, a taken-for-granted element of good science. In this context, "good" may not be intended to carry moral weight beyond technical competence, but the epistemic virtue concept invites us to consider the two as part of one another. The moralising tone of this rhetoric is also reminiscent of Daston and Galison's account of objectivity. Beyond explicit rhetoric about the inherent or inevitable openness of (good) science, the moral and truth-making qualities of "open" are implicit in a much wider range of discourse, as subsequent chapters will show. To the best of my knowledge there is no precedent in the scholarly literature for conceptualising openness as an epistemic virtue in Daston and Galison's terms. Freese and Peterson (2018) think along similar lines, however, and demonstrate the relevance of the concept to contemporary research. They describe "statistical objectivity" as an emerging epistemic virtue related to reproducibility concerns in science, which are in turn related to "open" movements.

A significant challenge to adopting epistemic virtue as a sensitising concept was its operationalisation. What did it mean to observe openness being constructed as an epistemic virtue? This question was both easy and difficult to answer. In many cases it was readily apparent: Tennant's (2018) meaning above, and many others, were unambiguous. But for the most part, epistemic virtue operated precisely as a sensitising concept, "suggest[ing] directions along which to look" (Blumer 1954:7) in a grounded exploration of qualitative data, without knowing precisely where this would lead. In a process of sensitising to epistemic virtue I found the following useful guides for my attention: i) a "moralizing tone" (Daston and Galison 2007:39) in relation to openness; ii) implications that science is improved fundamentally – in its truthfulness – by openness; and especially iii) implications that scientists themselves are improved by openness; adopt open identities; internalise openness; and/or report moralised "internal conversations" (Archer 2003) on the topic.

2.3 Methodology and methods

Key features of this research make it best suited for a qualitative methodology. Firstly, it is exploratory and invites the emergence of perspectives that may be unanticipated both to me as the researcher, and to advocacy and policy communities. To this end my overarching research question was intentionally broad: *How is the meaning of "open" (or "openness") being constructed in the context of science?* The research sub-questions (Section 1.7) provided focus and structure, but did not place limits on this breadth of inquiry. Qualitative methods are suited to this kind of exploration because they do not require parameters of measurement to be set out in advance. They instead ask what these parameters might be and why, allowing input from participants that may challenge the researcher's assumptions and reframe the study. Secondly, the research explores social construction of meaning, asking "how" questions about discourses, practices, values, and virtues. These questions are best addressed with qualitative data, which allows the capture of nuanced, "thick" descriptions that are sensitive to the "contextual understanding of social behaviour" (Bryman 2012:401). Thirdly, a key purpose of this study is to understand scientific openness "through the eyes of the people being studied" (ibid., pp. 399-401), in this case from the perspective of scientists. Qualitative research allows these perspectives to be recorded in participants' own words, and to be studied with empathy, context sensitivity, and an acknowledgement that people reflect upon and shape the worlds they inhabit. This is important given there seems to be a gap in understanding between many scientists and those who advocate open practices (see Section 1.6).

Qualitative methods suited to generating these in-depth, exploratory, social data include interviewing, ethnographic observation, and focus groups; documents are also a valuable data source for qualitative analysis (Bryman 2012:383). I chose the semi-structured interview as my primary empirical method, accompanied by document analysis to provide contextual depth. Occasional ethnographic observations also contributed to my interpretations. The following sections give more detail on my choice of methods.

2.3.1 Document analysis

Documents were important sources of qualitative data for this study. In Chapter 3, which would traditionally be a literature review, relevant texts were treated both as sources of information, and as social objects that reflect and shape the research context. I explore long histories of scientific openness in this chapter, drawing on academic history and social science articles, as well as occasional policy texts, in an attempt to understand established knowledge on the topic. I then consider how scientific openness is constructed by these documents, and begin to view them as products of historical context in their own right, that have shaped discourse. I particularly consider framings, narratives, and intertextual dynamics. Chapter 4 is analytical in a different way: it brings together texts from a wide variety to contexts to synthesise an understanding the breadth, salience, and apparently rapid emergence of contemporary open science at the start of the twenty-first century.

Most instances of document analysis appear in Chapter 5, which narrates twenty-first century developments in open science advocacy and policy. To add qualitative depth and exemplify these developments, I include analyses of nine advocacy and/or policy documents from the period 2000-2019. These documents indicate key characteristics and patterns in the contemporary construction of open science, addressing my research aims and situating the interview study. I consider the themes of the research sub-questions in relation to these documents. Qualitative patterns across the documents add depth to a narrative about the rise in salience, in turn, of open access, open data, and the overarching category of open science.

2.3.2 Semi-structured interviews

Semi-structured interviewing involves a flexible set of questions (Bryman 2012:471). Adjustments can be made to phrasing and order, and spontaneous questions can be added. This flexibility allows the researcher to explore responses beyond surface level, following up potentially illuminating responses. The semi-structured format also fosters a responsive, conversational interaction with potential for rapport and mutual trust. It contrasts with a structured interview, in which the same questions are asked in the same order, and with an unstructured interview. Structure is thought make responses of different interviewees more comparable (ibid., p. 210). However, especially in an exploratory context, it would be misleading to imagine that the context of the interview can be standardised. Thus, I have embraced the flexibility and the richness of an uncontrolled research setting. However, a degree of structure in the form of prepared questions aided organisation and focus. It also allowed a degree of context-mediated comparison.

Calvert (2001, 2006) was able to effectively explore the construction of “basic” science through semi-structured interviews with scientists and policymakers. I considered a similar approach to be promising for “open” because both descriptors of science are ambiguous and value-laden. They also each have particular significance in research and policy settings, and are each conditioned by long and complex social histories. Interviews do not, of course, reveal how open scientists “really” are, nor do they directly convey experiences of openness. What goes on is a type of social performance: a “dance of expectations” (Dingwall 1997:56) in which the respondent is cast in a role and “put on notice to talk” (ibid., p. 58) in a manner appropriate to the setting and topic. This does not make the interview a false account, however, only a situated one. Moreover, I do not expect “openness in science” to exist in scientists’ lives in any single, stable way. In Chapter 3, I show how scientific openness is relational, subjective, and conditioned by intention: thus, key dimensions of the topic can *only* be captured through scientists’ first-person accounts. I consider performative dynamics in my interpretations, including e.g. the possibility that interviewees expect openness to be more acceptable than closure. However, I also observed that interviewees seemed relatively unconstrained by their allocated role, perhaps due to critical/independent academic sensibilities. They often raised topics that did not conform to dominant discourses on scientific openness (one of which features in Chapter 8). And my initial studied concern about “leading” questions began to feel out of place; interviewees seemed resistant to being “led”.

It was not immediately clear how to engage with epistemic virtue, with its implied links to the scientific self, through interviews. The most obvious challenge is that our “inner lives” are “knowable only by our first-person selves” (Archer 2003:22). Nonetheless, Archer was able to study participants’ inner lives using semi-structured interviews. She acknowledged that both interviewee and interviewer must make theory-laden interpretations, but argued that "imperfectly successful" communication of this kind underpins all interpretations in the social sciences: "Are not the opinions and attitudes that we venture in any interview also 'digests' of a long and imperfectly self-analysed process of their formation?" (p. 155). Archer's argument and empirical work indicate that there is no in-principle problem with studying the internal life of the subject via interview.

2.3.3 Ethnographic observations

My knowledge and interpretations were shaped to a smaller degree by participant–observer experiences. Some of these occurred before I conceived of the research: I have previously worked in biological research and open access publishing, and my observations from these contexts were informative (see *Reflexivity*, Section 2.3.10). I also made observations in relevant contexts that arose parallel to the research, such as conferences or events related to open science; conversations; and online interactions, for example on Twitter. This was not part of my data-gathering method, but I made notes for myself as circumstances allowed, and reflections on these experiences shaped my interpretative mind-set. I did not adopt ethnographic observation as a primary method because it was not clear, at the outset, how "openness" – or lack thereof – would look in practice, or in which settings, if any, it would be meaningfully observable. Research planning would have involved assumptions about these parameters, whereas I intended to prioritise participants' framings of openness. Ethnographic observation would be a suitable method for follow-up studies on specific types of openness, e.g. interpersonal openness at scientific conferences.

2.3.4 Research participants

I recruited interview participants from two partially overlapping domains: academic, biological research science (40 interviewees); and open science advocacy and/or policymaking (14 interviewees). The emphasis was on the first of these domains –

biological science – as guided by research sub-question (d) that notes apparent widespread disengagement from "open" discourses and practices amongst academic scientists. Accordingly, I aimed to include scientists with a wide range of relationships to open science, some of whom had no prior connection with the topic.

Advocate and/or policymaker interviewees were working in fields including scholarly communication and publishing; university leadership, administration and libraries; and research funding and policy. Some were working freelance or as consultants on community-led open science advocacy projects; one had additionally become a humanities scholar. Some had engaged in several of these domains. These interviews were primarily intended to add context rather than forming the core empirical material of the study, but in reality took on a varied role, contributing both supplemental and core empirical data. They added breadth and depth to my knowledge of policy and advocacy perspectives; disclosed and reflected the latest developments in these areas; and provided informal, frank reflections on the policy and advocacy landscape that would be unlikely to appear in documents.

The boundary between these two categories of interviewees was not entirely clear. Many advocates and policymakers were former, retired, or freelance scientists whose careers had come to be defined by open science advocacy. Additionally, a small number of those I categorised primarily as biological scientists could be considered open science advocates. It was not necessary to draw a strict boundary during the interviewing process: I adjusted my script of questions for each individual to explore scientist and advocate roles as and where they applied. When it came to counting interviewees in each category, I considered interviewees' ways of presenting themselves professionally, and the capacity in which they came to my attention.

2.3.4.1 *Why academic scientists?*

I interviewed scientists working primarily in academic rather than commercial or government settings, while acknowledging that the three often overlap (see concepts such as "Mode 2" and "Triple Helix" science: Etzkowitz and Leydesdorff 2000). This was primarily to maintain a realistic scope. It would also be interesting to explore e.g. commercial scientists' constructions of scientific openness. However, varied data could be gathered in academic contexts. Several participants had commercial collaborations, or previous/current experience in industry. Moreover, academia, with its historical links to norms of communal, "public" knowledge-sharing (see Chapter 3) was a pertinent setting from which to begin questioning meanings of openness. More tangibly, academics are more likely than scientists in the private sector to be governed by recent open science policies, as they are often funded by public or charity money, which comes with an obligation to be "open" and accountable to wider society.

2.3.4.2 *Why biological scientists?*

Biological science was a manageably bounded, but sub-disciplinarily diverse, setting in which to explore relationships between openness and epistemic culture. My background in biology (see Section 2.3.10) was more than a convenience: it increased my access to potential interviewees through professional and personal networks, and contributed to my understanding of interviewees' work, and thus the rapport and meaning I was able to build in interview. It also added to my empathy and purpose in conducting the study.

Biology also has a distinctive position in relation to open science. Historically, biological fields have been shaped by cultural tensions and transformations related to openness. These are detailed in Chapter 4, and include the development of communal moral economies in twentieth century model organism biology (Kohler 1999); the commercialisation of molecular biology in the 1970s and 1980s (Wright 1986); and the Human Genome Project (HGP) in the 1990s and 2000s (Chow-White and Garcia-

Sancho 2011; Maxson Jones, Ankeny, and Cook-Deegan 2018). Perhaps owing to these growing tensions and transformations, open science movements of the early 2000s had a particular association with biomedical fields. Academic scientists established in these fields are thus likely to have encountered open science trends as a distinct change in the last 10-15 years, and are likely to have witnessed or participated in relatively rapid changes in discourse and practice within their field. UK biological scientists are also particularly likely to have encountered a moving frontier of open science policies governing their research output during this timeframe.

2.3.4.3 National contexts

Open science movements have become salient at a global scale and thus there is no uniquely informative national or institutional context for this study. For the same reason, any national or institutional context sheds some light on the phenomenon as a whole as well as situated features. This study was primarily conducted in the UK and secondarily in Australia. The UK is home to particularly vocal, developed and influential open science advocacy and policy movements, including those influentially led by the Wellcome Trust. The UK has also been influenced by the embrace of an ambitious open science policy platform in the EU. Australia provides an interesting comparison as its open science movements have been shaped by global trends, but it has taken a less top-down, regulated approach than the UK, and a less directive position than that of the EU. As an indicator of the contrast between Australia and the UK, an international analysis by Robinson-Garcia et al. (2020) shows that:

> *"British universities have by far the largest share of OA [open access] publications (median = 74%) [...] Except for the United States (median = 51%) and Brazil (median = 47%), all countries above world median are European. Asian countries, as well as Canada and Australia show OA shares below the world median." (p. 6)*

This is only one mode and measure of openness, but it could be expected to reflect contrasting contexts encountered by scientists in these countries.

Choice of national context is also related to my background, and again it is more than a matter of convenience. My research questions were inspired by professional involvement with open science movements in the UK, and thus have particular relevance to that context (see Section 2.3.10). Moreover, this past experience made it easier to recruit open science advocates in the UK than elsewhere. The comparative dimension arose from my dual nationality and undergraduate training, through which I was connected with a biologists in Australia. This allowed me to conduct a pilot interview phase with familiar individuals, easing recruitment and rapport, and creating a flexible space in which to develop interview structure and technique. Moreover, my connection to these biologists arose before my involvement with open science and thus represented a privileged point of access to scientists ambivalent towards the topic, who were harder to recruit – but particularly important in the study context.

It became clear that interviewees, particularly scientists, did not have a simple relationship with national context. While they were largely based in the UK or Australia, many had done scientific training or research elsewhere. There was nothing homogeneous about the national experiences of a typical interviewee: both across their careers and in any given moment their experience blended national research cultures. National comparisons were thus primarily a means to increase the diversity of my sample and to retain awareness how contextual specificity interacts with experiences of a global movement. Notably however, this heterogeneity of experience was largely limited to the global north and to English-speaking contexts.

2.3.4.4 Institutional and funding contexts

Institutional and funding contexts in many cases had a more direct influence upon experiences of open science policy. The perspective of this study is partial in this regard: across national contexts, scientist interviewees worked at relatively well-funded institutions that cultivated prestigious research reputations. Interviewees tended to have received grants from various sources including national research councils, and in the UK, the Wellcome Trust and European Commission, as well as any commercial collaborators. Each interviewee was subject to a mixture of different open science policies, from both their funders and their institution.

2.3.4.5 Sampling strategy

Amongst biological scientists, I sampled purposively for diversity (Bryman 2012:418) primarily according to sub-discipline, career length, gender, and race and ethnicity. Secondarily I aimed to include scientists in research groups of different sizes, and variably related to industry. This approach could also be called “theoretical sampling”, a strategy of representing categories that are theoretically meaningful, as opposed to “statistical sampling”, which samples a population randomly (Glaser and Strauss 1967:62). This strategy was intended to increase the diversity of stories, experiences, and reactions that I would record in interview, and to increase the visibility of perspectives that might be marginalised using a proportional approach. The factors that I considered in the approach were those which I suspected would have a relationship with scientific openness, as well as those which relate to social diversity. I expected these factors, and others I had not anticipated, to intersect in complex ways for each interviewee. My goal was not to disentangle and measure the influence of these factors, but to create a situation in which rich, varied, contextual relationships with openness could be observed. These factors are listed and expanded in **Table 1**. Advocates and policymakers were selected based on their roles in open science advocacy/policy. I mapped relevant organisations and communities based on past experience and policy documents.

Table 1 | Factors used to guide purposive sampling of scientist populations.

Primary factors guiding purposive sampling
Biological sub-discipline: I referred to classifications of these sub-disciplines (e.g. subject divisions in universities; biorxiv.org list of subject areas) during recruitment and targeted new invitations to scientists in under-represented areas.
Length of career: Established scientists may have different orientations to openness from early career scientists, perhaps linked to the recent rise of open science. I recruited from different academic "generations", gauged according to decade of doctoral training.
Gender: Balancing gender is a vital part of any study seeking to represent scientists' experiences. Women face well-known systemic inequalities in science (Etzkowitz, Kemelgor, and Uzzi 2000), and non-binary or transgender scientists may face greater barriers.
Race and ethnicity: Black, Asian, and minority ethnic scientists also face systemic inequalities (Wadman 2012), so I aimed to over-represent these groups compared to the study population (with limited success: see Section 2.3.4.7)
Secondary factors guiding sampling
Research group size: I sought to include scientists from both small and large groups, and those who typically publish with smaller or larger groups of co-authors.
Commercial links: I sought to include scientists both with and without commercial links, as this may be related to scientific openness in multiple ways.

2.3.4.6 Recruitment strategy

I recruited interviewees via (i) "cold" emailing, i.e. contacting individuals without a prior connection; (ii) my professional and personal networks; and (iii) snowballing (Bryman 2012:202–3) – asking for recommendations from prior interviewees. I employed all three strategies to recruit scientists, which ensured that a variety of motivations to participate (not all based on an interest in openness). Cold emailing was better for purposive sampling, as diverse characteristics of scientists could be researched in advance. However, it was less effective for recruiting junior scientists, who less often had up-to-date profiles on their institution or laboratory websites. The response rate to cold emails was also relatively low (approx. 20%; higher for men than women). I anticipated this challenge from preliminary research (Attenborough 2015) in which individuals from disadvantaged/minority groups were less likely to accept a cold email invitation, for myriad possible reasons: e.g. limited time due to direct or indirect burdens of minority status and systematic inequality. I therefore increased the proportion of invitations to individuals in these groups in the current research, with more success in relation to gender than race and ethnic minority status (see Section 2.3.4.7). Networking and snowballing were more effective for recruitment – nearly all such invitations were accepted – so I employed these as well, and attempted to steer them towards gaps in my purposive sampling, albeit imperfectly. These were also my main strategies for recruiting advocates and policymakers, where it was important to speak to specific individuals.

2.3.4.7 Interviewee characteristics

Scientist interviewees were working in a wide range of biological fields, listed with relevant details in **Table 2**. I did not question interviewees about their gender, racial or ethnic identities in association with the interview as this seemed needlessly intrusive for the study purpose (see *Limitations*, Section 2.3.11). I gauged and attempted to balance demographic diversity according interviewees' professional self-presentations, including e.g. pronouns on university profiles, according to which 17 scientist interviewees were women (~42%) and 23 were men (~58%). Although relatively balanced overall, gender representation was uneven across the career length category (decade of PhD completion). I tended to be able to recruit more established men and more early-career women, reflecting population characteristics that likely result from both historical gender disparities and a "leaky pipeline". At present, women tend to be highly represented at doctoral training stages in biological fields, but are less represented thereafter (Shaw and Stanton 2012). Genders are not specified in **Table 2**, but are indicated by pseudonyms or names, as well as pronoun use where quotations are reported[4]. **Figure 2** visualises the characteristics of the scientist group by gender, career length, and location.

[4] One female interviewee opted for a traditionally male pseudonym for a variety of reasons other than gender identity, but is recorded (anonymously) as a woman in Figure 2.

Between four and six (10-15%) scientist interviewees were visible[5] to me as individuals from Black, Asian, or racial or ethnic minority backgrounds (the majority were men and non-Black). This is roughly in line with the representation of racial and ethnic minority academics in UK higher education in 2018/19 (approx. 15% BAME, including 2% Black, 9% Asian, 2% mixed, 2% other; Higher Education Statistics Agency n.d.). Therefore although this study included some racial and ethnic diversity, it did not over-represent BAME individuals compared with the study population, which would have been better for representing diverse perspectives. This shortcoming is related to extremely low representation of BAME scientists in the institutions from which I recruited. It proved insufficient to send more invitations to underrepresented groups, because it was possible to exhaust the pool of possible BAME interviewees at a given institution. These individuals were likely overburdened and had no duty to participate in exercises – such as this research – that reached for diversity in deeply unequal settings. Addressing this issue in the future would require a study design with a focused and prioritised effort at racial and ethnic minority inclusion.

Eight (20%) scientist interviewees were based in Australia, with the majority, 29 (~72%), based in the UK (Scotland and England). Networking and snowballing recruitment techniques also led to the inclusion of two interviewees based in the US, and one based in Saudi Arabia (the latter with institutional and personal ties to Australia and the UK). I decided it would be beneficial to include these interviewees given I had the privilege of access, even though their national settings were not a primary research focus. UK interviewees were from seven different institutions (19 interviewees were from one of these). Australian interviewees were from mainly from one institution (two overall). The specific institutions are not named to increase interviewee anonymity.

Advocate and policymaker interviewees are listed in **Table 3**. Most of these individuals had a professional base in the UK, but all had international networks, and

[5] Visible minority status is a partial and problematic view but preferable to "colour-blindness" (Song 2020).

two were Australians with professional knowledge of Australian open science contexts. The remaining two advocates were based in Germany, one having begun his advocacy career in the UK. Eight of these interviewees were men and six were women. The group was not racially or ethnically diverse, likely reflecting features of open science movements in the global north, as well as the tendency for snowballing network-based recruitment methods to homogenise representation.

Table 2 | List of scientist interviewees and contextual information in alphabetical order by pseudonym/real name. Country corresponds to their institutional base at the time of interview. Disciplinary keywords are those used by interviewees in relation to their professional identities.

Pseudonym/ *real name	**PhD/DPhil completion**	**Country**	**Disciplinary keywords**	**Month of interview**
Adam	1980s	Australia	Parasitology, biochemistry, molecular biology, physiology	Jan 2017
Andrea	PhD current	UK	Immunology, microbiology	Oct 2018
Arthur	2010s	UK	Analytical chemistry, cell biology	Jul 2018
Ben	1990s	UK	Applied conservation genetics	Jan 2018
Cedric	1960s	UK	Agricultural botany, physiology, molecular biology	Feb 2018
David **David Hume*	1970s	UK	Comparative genome function, immunology	Dec 2017
Elliot	2010s	UK	Molecular, computational biology, bacterial genomes	Jan 2019
Enrico **Enrico Crema*	2010s	UK	Computation archaeology, biological anthropology	Nov 2018
Erin	2000s	UK	Palaeoanthropology, biological anthropology, biomechanics	Aug 2018
Ernie	1970s	UK	Cell biology, chromosome structure	Oct 2017
Gavin	1990s	Australia	Computational evolutionary biology	Jan 2017
Greg **Gregory P. Copenhaver*	1990s	US	Genetics	Feb 2018
Henry	2000s	UK	Developmental immunology, animal studies	Dec 2017
Ian	1980s	Australia	Human genetics, genomics, policies and ethics	Sep 2017
Jacqueline	1980s	Australia	Ecophysiology, climate	Jan 2017
Jason **Jason Gellis*	PhD current	UK	Biological anthropology, palaeoanthropology	Nov 2018
Jenny **Jenny Graves*	1970s	Australia	Animal (comparative) genetics and genomics	Nov 2017

Jude	1970s	Australia	Zoology, developmental molecular biology, genomics	Jan 2017
Julia	2000s	UK	Immunology, evolutionary biology, ageing research	Jun 2018
Kate	2000s	UK	Synthetic biology, interdisciplinary	Jun 2018
Lara	PhD current	UK	Psychology, neuroscience, human-robot interaction	Dec 2018
Lena	2000s	UK	Biological physics, microbiology	Apr 2018
Luke	2010s	UK	Evolutionary biology, ageing research, behavioural ecology	Nov 2018
Madison	PhD current	UK	Disease ecology, immunology, parasitology	Jan 2019
Mark **Mark Tester*	1980s	Saudi Arabia	Plant physiology, genetics, bioinformatics, genomics	Feb 2017
Mat **Matthew Todd*	1990s	UK	Organic chemistry, open source medicine development	Apr 2019
Melanie	2000s	UK	Ecology, parasitology, immunology	Jul 2018
Michelle	1980s	Australia	Biochemistry, genetics, education	Jan 2017
Miriam	2000s	UK	Quantitative genetics, data science	Jan 2018
Molly	PhD current	UK	Behavioural ecology, evolutionary biology	Nov 2018
Neil	2000s	UK	Structural, computational biology	Oct 2017
Nicole	PhD current	UK	Neuroscience, clinical psychology, dementia	Oct 2017
Olivia	PhD current	UK	Epidemiology, data science	Jul 2018
Oscar	2010s	UK	Biological engineering, synthetic biology	Dec 2017
Richard	1980s	US	Developmental, evolutionary genetics, genomic medicine	Mar 2018
Roger	1980s	Australia	Molecular, developmental, evolutionary biology	Jan 2017
Rory	2010s	UK	Biomedicine, diagnostic clinical genetics	Jan 2019
Steve	1970s	UK	Plant biochemistry	Nov 2017
Thomas	1990s	UK	Systems biology, plant science, data management	Oct 2017
Yvonne	1970s	UK	Human and animal genetics, gene expression	Nov 2017

Figure 2 | Characteristics of the scientist interviewee group by gender, career length (PhD/DPhil completion decade), and location of home institution at the time of interview. The distribution of location-colours reflects (within categories) the order in which interviews were carried out.

Table 3 | List of advocate and policymaker interviewees and contextual information in alphabetical order by pseudonym/real name. Underlined country indicates their location at the time of interview, other countries listed are influencing contexts.

Pseudonym/ *real name	**Domain**	**Professional role / organisation** (at time of interview)	**Country**	**Month of interview**
Cameron **Cameron Neylon*	Advocacy and research	Prof of Research Communications, influential open science advocate and scholar, trained as a scientist	UK, Australia	Feb 2018
Catriona **Catriona MacCallum*	Scholarly communication, publishing	Director of Open Science at publisher Hindawi, former PLOS Advocacy Director, trained biological scientist	UK	Jan 2018
Danny **Danny Kingsley*	University policy & admin, scholarly communication	Head of Scholarly Communication, University of Cambridge, PhD on open access in Australian context	UK, Australia	Feb 2018
Donald	Research, research policy	Scientist with contextual knowledge of the Royal Society's *Science as an Open Enterprise* report (2012)	UK	Jan 2018
Graeme	Research funding and policy	Open science policymaker and advocate at major UK research funder	UK	Feb 2018
Jenny **Jenny Molloy*	Advocacy and research	Researcher and high profile advocate of open approaches, trained biological scientist	UK	Dec 2018
Jon[†] **Jon Tennant*	Freelance advocacy / research	High profile freelance open science advocate, early career palaeontologist	Germany, UK	Jan 2018
Laura	University administration, scholarly communications	Administrator supporting and advocating compliance with open access, university biological science context	UK	Jul 2018
Mark **Mark Patterson*	Scholarly communication, publishing	Leader in pioneering open access publishing contexts (Director at PLOS and eLife), trained biological scientist	UK	Nov 2018
Paul **Paul Ayris*	University leadership & policy, libraries	Pro-Vice-Provost UCL (Library Services, Office for Open Science & Scholarship), European policy connections	UK	Feb 2018
Peter **Peter Murray-Rust*	Freelance advocacy / research	Computational chemist, pioneering advocate for open source approaches to science, particularly open data	UK	Feb 2018
Sarah	Scholarly communication, publishing	Senior journal editor, trained biological scientist	UK	Jan 2018

[†] I acknowledge with sadness the death of Dr Jonathan P. Tennant. Jon was not yet 32 years old when he died in April 2020 in a motorcycle accident (European Geosciences Union, 2020).

Stephanie **Stephanie Dawson*	Scholarly communication, publishing	CEO of the platform ScienceOpen, academic publishing background, training in biological science	Germany, US	Jan 2018
Stuart **Stuart Taylor*	Scholarly communication, publishing	Publishing Director at the Royal Society and advocate of open science, trained scientist	UK	Feb 2018

2.3.5 Interview procedure

Interviews were semi-structured and guided by a schedule (see *Appendix B: Interview schedule*, and **Figure 3**). The majority were conducted in person, at a location chosen or guided by the interviewee: typically their office or a meeting room. A significant minority [16, ~30%] were conducted via online video chat. Most interviews lasted around one hour. All but one were audio recorded, which allowed me to take minimal notes, focusing my attention on listening and remaining visibly attentive. This promoted social connection, as well as the presence of mind that I needed to formulate follow-up questions and shape the interview as it unfolded.

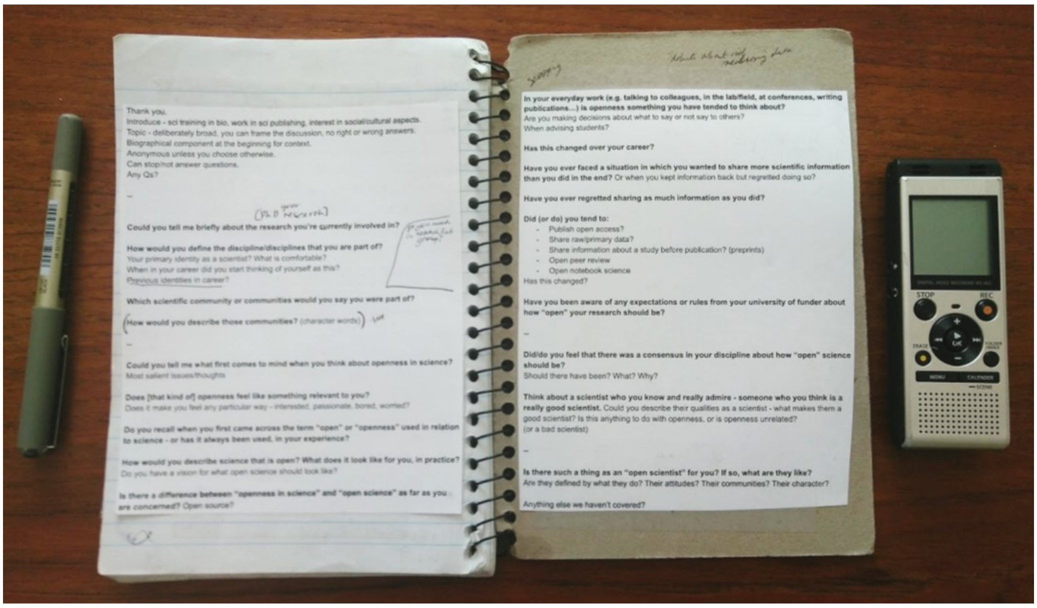


Figure 3 | Interview materials. The notepad shows my typical interview schedule, visible across one page spread to maximise my focus on the interview, with main questions in bold. Legible detail is available in *Appendix B: Interview schedule*.

Full details of the interview procedure are given in *Appendix B: Interview schedule*, but the main elements were as follows:

- Preamble: I introduced myself, the interview premise, informed consent information, and made space for questions. I sought consent, and with approval, began the audio recording and interview.
- Biographical questions, to develop context and social ease: The first question was about interviewee's research or current role. For scientists, I asked about disciplinary identity and experiences of community.
- **Core question**: *"What first comes to mind when you think about openness in science?"*, designed to encourage a wide range of responses including those not anticipated from policy and advocacy. This proved so: it also shaped my findings in distinctive and worthwhile ways (see Section 8.3). Follow-up was flexible, aiming to explore the response.
- Other exploratory, "opening up" questions: *"Does openness feel like something relevant to you?"*, followed by, *"Does it make you feel any particular way? e.g. interested, passionate, bored, worried…". "Do you recall when you first came across the term 'open' or 'openness' used in relation to science – or has it always been used, in your experience?". "How would you describe science that is open?"* and *"What does it look like for you, in practice?"*. Each approached from a different angle, encouraging expression of complex or conflicting meanings and experiences. Part-way through interviews, I added: *"Is there a difference between "openness in science" and "open science" as far as you are concerned?"*.
- Targeted questions: I asked about types of scientific openness established in policy and advocacy contexts. This gauged awareness, attitude, and practice around e.g. open access, open data, open preprints, open peer review, and open notebook science, if these had not already arisen.
- Early on, I found that some questions evoked "interpersonal openness" (see Chapter 8): *"In your everyday work (e.g. talking to colleagues at conferences) is openness something you have tended to think about?" and "Are you making decisions about what to say or not say to others?"*. Once the category had emerged in earlier interviews, I used these as prompts for interviewees who had not raised it initially.

- Concluding question, sensitised to epistemic virtue: *"Think about a scientist who you know and really admire, who you think is a really good scientist,"* then describe their "good" characteristics. This question was added after the pilot.
- Advocates and policymakers: Many of the same questions were useful, allowing a degree of comparison. Otherwise, I asked role-specific questions, and about changes they wished to see in science and society; their specific policy or advocacy activities; perceived successes or challenges of engaging with scientists; and whether/how they envisaged an "open scientist".

2.3.6 Research ethics

I obtained ethics approval from the School of Social and Political Sciences, University of Edinburgh. All participants gave informed consent (see *Appendix C: Ethics approval*; *Appendix D: Sample information sheets and consent forms*). The study was of minimal risk to participants. Interviewees were adults with a full capacity to consent; they were also highly educated and were relatively empowered both societally and in the interview context. They ranged from being my peers – fellow PhD researchers – to being professional elites: professors, department heads, and project or university leaders. I could reasonably expect interviewees to understand explanations of the study and their role, and to ask questions or withdraw if concerned. One area of ongoing ethical reflection was reflexive "openness" around anonymity and data archiving, which became a methodological consideration in its own right.

2.3.7 Methodological reflexivity: openness and closure

I developed a reflexive alertness to "openness" in my own research, without *assuming* that openness would be good for this particular form of knowledge-making. With this in mind, I considered the appropriateness of participant anonymity, which is often treated as a cornerstone of ethical practice in the social sciences (e.g. Bryman 2012:146). Anonymity is protective and is not only a form of closure: it enables the participation of individuals who might otherwise feel uncomfortable, leading to the representation of a broader set of views. It also enables a certain freedom in the interview space (a form of interpersonal openness with parallels in Chapter 8). However, using participants' real names enables acknowledgement of their contributions, and is a form of methodological transparency that may help readers to understand the data. Moreover, it enables participants with certain "open" principles to contribute in a way that is consistent with their own ethics and epistemology.

To navigate this tension, I offered interviewees a choice between anonymity (by default) and use of their real name. The choice was signposted before the interview but offered afterwards, together with an opportunity to view the interview transcript. This led to a mixture of anonymous and attributed responses: either condition could be interpreted as more "open" or credible. To represent these two categories evenly, I include only first names, for scientists, in the results Chapters 6-8 – whether they are real first names or pseudonyms.

For similar reasons, I offered a choice about the interview transcript: by default it remained confidential, but interviewees could opt for secure online archiving (access for research purposes) or public online archiving. This acknowledges the importance of open data principles to some interviewees: at least one participant viewed their transcript as a potential "historic artefact". I also value such principles – conditionally and experimentally – and recognise their value to my PhD funder:

> *"Whilst not compulsory, ESRC-funded students are strongly encouraged to offer copies of data created or repurposed during their PhD for deposit at the UK Data Service as it is considered good research practice."* (ESRC n.d.)

Offering a choice – with confidentiality as a default – was important because of the rich, personal, deeply contextual qualities of interview data, and the permanent, public, discoverable qualities of online archiving. The interview is a social interaction that occurs in a private space; making it "open" had the potential to be discomfiting if not risky to interviewees. One interviewee commented, upon receiving the transcript, that he had been somewhat unguarded as I had "got him to relax", whereas immediately after the interview he had spoken of standing by his responses publicly.

This underlined the importance of providing the transcript and allowing an opportunity for reflection before enacting data (transcript) openness.

Some interviewees opted for secure or public transcript archiving. Some also wished to go by their real names. This degree of transparency – which compounded the risks and benefits of both forms of openness – prompted a particular ethical attention on my part. I sought written assurance from these interviewees that they had considered the implications of their comments being widely and indefinitely attributable online, and I examined transcripts for sensitive content. It was challenging to negotiate this aspect without unintentionally signalling that I view either openness or closure as problematic. Significant labour was involved in realising this reflexive engagement with openness and closure, but I felt that it was worthwhile. Two transcripts that have been approved for public sharing are included in the Appendices both as data and as examples of method (see *Appendix E: Interview transcripts*).

There are an almost unlimited number of additional avenues for reflexive experimentation with open research practices. One obvious consideration is open access to this PhD thesis, and to any scholarly publications that arises from it. I intend to arrange this, while acknowledging that open access does not render my work automatically understandable, receptive, or ethical. It does, however, increase the ability of interested parties to find and read this work, to discuss it with me or others, and use it in their own work. This is especially important as I hope my work will be of interest in open science communities who value and sometimes rely on online accessibility. Subsequent sections include reflexive engagements with openness during my analytical process, during which it becomes clear that methodological openness for grounded, situated social research may need to be different from that advocated in the natural sciences (e.g. open notebook science).

2.3.8 Interview transcription

I transcribed most of the interviews myself within the software package *NVivo*, which took 6-8 hours per interview. This established my familiarity with the data, which was part of the analytical process. For time efficiency, I used my studentship funding allocation to pay for professional transcription of the final 17 interviews; I re-familiarised myself with these interviews at later stages of analysis.

2.3.9 Interview analysis

I took a grounded approach to interview analysis, with an emphasis on exploring data rather than applying a top-down theoretical framework. A prevalent approach of this kind is grounded theory (Glaser and Strauss 1967) which formalises a progressive building up of "theory" from close, iterative examinations, categorisations ("coding"), and conceptualisations of qualitative data. I took some guidance from this approach – of which there are several versions – but did not adopt it wholesale. My reasons for avoiding prescribed qualitative methods align with those of Thomas and James (2006). They argue that grounded theory "misses the best": that in "hankering after order", it fractures and relegates the original voice of both participant and researcher, "constraining and distorting qualitative inquiry" (pp. 767, 790). This results in a "mirage of some kind of reliable knowing" (p. 791), which has been key to the success of grounded theory and its continued popularity:

> *"Grounded theory, and other techniques of analysis in qualitative inquiry, are bound to be popular, because they meet a need. For while qualitative inquiry is absolutely valid, it is difficult to do [...] Such ways of doing research can lead to a floating feeling, a lack of direction. What does one do with one's data? Surely one can't just talk about it." (p. 768)*

After grappling with qualitative data, I found that Thomas and James' argument rang true. My attempts to apply any one structure – especially one that breaks down narratives in order to build up concepts – led to feeling that I was missing the forest for the trees: losing context, subtext, and connections between micro and macro textual elements. I turned away from prescribed analytical methods primarily out of concern that they were suppressing both obvious and subtle creative insights.

What is the alternative to prescription? Thomas and James (2006) assert that "the argument for qualitative research now stands in its own right" (p. 790) and that although qualitative analysis requires systematic effort, imagination, and skill, its core processes need not be elevated above everyday reasoning:

> *"…the tacit and spoken tools of normal sense-making, or of what Schatzman (1991, p. 304) calls 'common interpretative acts' – of review, rehearsal, of talking about it with friends, of employing practical syllogism, recognition, evaluation, coming to a conclusion." (pp. 788-789)*

This position is arguably one of methodological openness (and the authors refer to it in passing as "open qualitative inquiry", p. 770). It requires an admission that there is no formula for truth-telling about social worlds.

My eventual interview analysis arose from combination of techniques and reflections that evolved over the research period (for detail and visualisations, see *Appendix F: Approach to interview analysis*). My approach was anchored by close attention to the same data at multiple stages, allowing insight to emerge progressively both within and across a growing set of interviews. Analysis began in the interview interaction itself and immediately afterwards through attention, exploratory questioning, and reflective note-taking. It continued during transcription, which allowed me to attend precisely to each question-and-response while taking interpretative notes. I then experimented with several techniques for transcript analysis, including close reading with line-by-line annotation; mind-mapping (using Coggle.it); and coding in *NVivo*. These techniques helped me achieve an initial depth of insight with each interview: familiarity not only with content, but with tone and context; patterns of narration; and the relationship of the interviewee with the topic and interview setting.

NVivo coding brought breadth as well as depth, as this process allows systematic comparison across interviews. I went through several phases of expansive coding that evidenced the breadth of the topic and interviewees' associated experiences. It became manifest at this point that the interviews contained multitudes: there were many valid, valuable routes to a final analysis. No single route would emerge through coding and refinement of codes: an overarching decision was required about the focus and purpose of the analysis. To do this, I reflected systematically on my analyses in relation to my overarching research question: *How is the meaning of "open" (or "openness") being constructed in the context of science?* I noted that three stand-out categories characterised scientists' constructions from early on: (1) open access, (2) data openness, and (3) interpersonal openness. These were not the only categories of "openness" constructed by scientists, but they gave an accurate indication of where interviews were focused and offered broad, cross-cutting entry points for more in-depth analysis. They were "obvious" categories: the first two were well-established categories in "open science" discourse, and all three arose with exceptional frequency in scientists' responses.

From this point I structured my analysis according to these three categories, which eventually became Chapter 6-8. Once I could focus on each category separately, the common and varied characteristics of scientists' *relationships* with these types of openness became clearer. I analysed each category in a different way, to best illustrate these relationships. For example, I found that relationships with open access were best captured through identification of common framings across scientists' experiences (e.g. a financial framings). Reactions to data openness were instead extremely varied and more emotional: this prompted me to conceptualise the circumstances and internalised experiences that mediate relationships with data openness, as distinct from open access. Interpersonal openness was a novel category and thus a novel analysis, but had features in common with data openness, which added depth to both analyses. These analytical choices: my focus on the *meaning* of openness to interviewees; my decision to *break down* the analysis according to three common scientist-identified categories, and my focus on *relationships* with openness, all shaped my findings. The third choice in particular was honed by my sensitising concept, epistemic virtue, which draws attention to openness in relation to experiences of scientific self.

The final phase of analysis involved several techniques that helped me achieve both depth and a systematic overview. This included tabulating the salience of each category of openness in each scientist's account; re-reading transcripts; and writing summaries to characterise individuals' relationships with each type of openness. I completed this analysis through drafting and finalising Chapters 6-8, because the writing itself prompted reflection, revisiting of transcripts, and higher-level synthesis with each revision. This process is represented in Chapter 9 as well, which contains analysis bringing together the previous three chapters, and conceptualisation that transcends the three "obvious" analytical categories.

2.3.10 Reflexivity

As a social researcher I am attempting to make sense of meanings, attitudes, concepts and cultures through the lens of my own social worlds: reflexivity about my own position is thus particularly important. As Malterud (2001) writes:

> *"A researcher's background and position will affect what they choose to investigate, the angle of investigation, the methods judged most adequate for this purpose, the findings considered most appropriate, and the framing and communication of conclusions." (p. 483-484)*

This research is strongly shaped by my broader life experience. This makes it distinctive, partial, enriched, and insightful in ways that I will attempt to capture.

2.3.10.1 Conceiving of the research

This research brought together at least three threads of my past experience. The first was my identity as a biologist, which I developed as an undergraduate with the opportunity to conduct primary research, and as a laboratory technical assistant prior to this. The second was a fascination with *the social* in science that I did not understand as such during my undergraduate degree: biological theories seemingly tinged with social ideals; trust and care in experimental setups; the necessity and power of publishing (well). This led me an entry-level job in scientific publishing: an opportunity to make sense of behind-the-scenes work that legitimised knowledge. I was hired by PLOS, a small non-profit publisher focusing on biomedical sciences that – then eight years old as a business – had recently consolidated its financial and reputational success. This serendipitously added a third thread to my experience: PLOS, an open access publisher, was one of the most significant and symbolic actors in a growing movement promoting openness of scientific knowledge. I stepped into this world naive about "open" but enthused by the embrace of a moral mission in science.

Working at PLOS for three years honed my knowledge of – and care for – both "open" movements, and biological science communities. My job was to support and manage a "community" journal: it was run by and for practising scientists from a particular biological discipline. This meant daily communication with biologists as authors, reviewers, and editors. As a representative of an "open" organisation I witnessed, in everyday interactions, a meeting between open ideals and the daily priorities of scientists. I observed what many advocates of open science have noted, which is that scientists who actively pursue "open" practices and values seem to be small minority in a population for whom such concerns are at most secondary. This tension between open advocates and a wider scientific community was one that I identified with – on both sides. I was also fascinated by the advocacy discourse that gave "open" meaning; the moral imperative that "open" seemed to hold; and the way that PLOS' mission shifted from open access towards data openness and other "open" practices during my time there (2011-2014). By the time I left PLOS it was with a sense that I had been part of a fascinating social–scientific phenomenon that merited systematic reflection. I became aware of STS, and embarked on postgraduate study that, by the time of my PhD application, led me back to scientific openness. My experiences at PLOS were the most immediate motivators of my research questions.

2.3.10.2 My position as a researcher

For the reasons above, I have more than a theoretical relationship with open science movements and biological science communities. In this research I step back from open science discourses to examine them, but I am not a neutral investigator. My experiences at PLOS led me to view science as a domain in flux, in which scientists and open advocates alike perceive major problems with the integrity of knowledge production and scientific career structures. At PLOS, I embraced open science as a solution to many of these problems. Since then my view has been tempered by attention to many meanings of "open" in practice, and their varied effects in scientific lives. However, I carry with me a conviction that scientific systems require critique, care, even transformation, and that open science discourses and practices can be part of a helpful change in that direction. Moreover, I hold many open science advocates in great esteem; some are my personal acquaintances, and interviewees in this research. Although, as a researcher, I do analysis – a safe, equivocal position – my hope is to do so in a way that supports rather than undermines those who advocate for better science and scientific lives.

My position is no less shaped by my connection to biological science. Despite my move away from biology, I have found that my scientific identity is deeply embedded, and scientific ways thinking are blended into my social research lens. Perhaps the most important influence from this background is the empathy that I feel for varied experiences of scientific life: for over ten years, I have been close to practising biological scientists as bosses, supervisors, mentors, colleagues, friends, and family. Several have contributed to this research as interviewees. My sister is a biologist who is writing her PhD as I write mine: our conversations have allowed me to check my interpretations against an ongoing life in science. I have adopted a purposefully empathetic position as a counterpoint to tendencies in both STS and open advocacy to be cynical in interpreting (other) scientists' behaviour as, for example, primarily self-interested (e.g. Gieryn 1983). This sharpens my attention to the human element of this study; any temptation I might have to make one-dimensional interpretations of interview data is checked. This is not the same as being uncritical; empathy means acknowledging depth and complexity.

2.3.11 Limitations and potential to broaden scope

Although the interview findings have potential relevance beyond the immediate contexts studied, they are a partial perspective on scientists' constructions of openness. As I have noted, extending the study into commercial or government science contexts, beyond biological fields, and outside privileged institutional and national settings would yield valuable and complementary perspectives. In particular, this study lacks non-Anglophone and global south perspectives, which would challenge and/or align with the current findings in multiple and unexpected ways. Constructions of open access in particular (see Chapter 6) are tied to variable local conditions such as institutional access to subscription journals; and open access policy, administration, and fee support. Interpersonal openness in science (see Chapter 8), meanwhile, is likely related to broader cultures of interpersonal communication that vary between nations and regions.

Interviewing at prestige-oriented universities may emphasise particular dynamics of openness related, for example, to competition. Research institutions that cultivate a different type of image – or citizen science settings – would likely tell contrasting stories. There are also some professional perspectives that I would emphasise more in any future study: those of laboratory technicians, librarians, and administrators of open policy/practice. It would also be pertinent to do an equivalent study of a humanities or social science context. This was outside the present scope, which is contextualised by a long history of openness in the natural sciences: but "open" policy now applies across a spectrum of scholarship, and rich histories of openness could be told across a spectrum of knowledge-making. I would also consider, in future, recording intersectional identities (e.g. gender, ethnicity, race, class, disability, sexual orientation) if this could be done in a non-intrusive way around interviews. This would increase awareness and accuracy of the range of views included in the study.

Generational differences were important in this study and were well-represented, but merit ongoing, focused study.

Chapter 3 | "On the shoulders of giants"

Historical framings of openness in science

> *"Openness is a catch-all term that is deployed in many different ways, for different purposes, and with different reasons and motivations […] [It] is sometimes deployed as being very new: it's all about the web, and the internet, and what we can do today. [..] And at the same time, it's also frequently deployed […] as something that's very old, that sits in the roots. And that deliberate pointing out – that these two things [new and old] are the same – also has deep roots."*[6]
>
> *Cameron Neylon, Professor of Research Communication, Curtin University*

3.1 Introduction

In this chapter, I explore the historical construction of scientific openness. This would be a key dimension of any social-conceptual analysis; in this case, my motivation to study openness in science came in part from the historical resonance it seemed to carry (see Section 1.3). I commonly encountered the assertion that openness is an enduring, inevitable, even defining characteristic of science – one that goes back for centuries. This created a curious juxtaposition: scientific openness appeared to be both as old as science itself, and so new that social and policy movements embracing it have taken form in only a few recent decades. This juxtaposition led to the formulation of my first research sub-question: *how do contemporary "open science" discourses relate to older, traditional discourses about the essential value of openness in science?* This question guides the current chapter, which is in part a review of texts about scientific openness in past historical periods. It is also an analysis, in the sense that it examines both these and other relevant texts as artefacts of a certain time and place that reflect and construct historically situated ideas about scientific openness.

[6] Comments from a panel discussion in February 2019, part of an *Openness and Reproducibility in Science* workshop that I co-organised with staff at the Australian National Centre for the Public Awareness of Science and the ANU Research School of Biology.

The aim is to comment both upon the past of scientific openness, and upon the construction of that history.

The chapter begins with two contrasting accounts of scientific openness: firstly as an enduring, apparently traditional, and essential characteristic of scientific culture; and secondly as the contemporary, internet-associated phenomenon of "open science" – a diverse yet congruent set of movements and practices that seek to revolutionise science, redeeming its open character. I then analyse literature to contextualise these two accounts and the relationship between them. I interrogate the traditional ideal of scientific openness by considering its historically contingent origins; its intermingling with secrecy and closure across and within scientific cultures; and the social, relational roles that are played by openness and secrecy in context. Reflecting upon these analyses, I consider how accounts of science as inevitably or essentially open may be constructed through *boundary work* (Gieryn 1983) that delimits both "science" and "open". I present more tangible ways to characterise and situate historical traditions of "openness" in science – in relation to community, professional status, and publication practices. These analyses culminate with an argument that histories of scientific openness are difficult to disentangle from the contemporary "open science" lens that now shapes them.

3.2 Traditional openness: enduring, inevitable, and essential

Traditional stories of scientific openness go as follows: if scientists communicate – that is in some form share, make visible, or be "open" about – their results, scientific communities can collectively build on upon these findings and a progressive, self-organising, even self-correcting corpus of knowledge emerges. The scientist and philosopher Michael Polanyi (1962) tells a version of this story when he asserts that the work of individual scientists – "freely making their own choice of problems and pursuing them in the light of their own personal judgment" – is coordinated by their awareness of one another's results, without which "scientific progress would come to a standstill" (p. 54). He likens this system to the collaborative completion of a jigsaw puzzle, in which each worker can see and adjust to the contributions of the others (p. 55). In stories like these, openness is an organising principle, and a necessary condition for the production of science that is collective or cumulative – rather than individual or idiosyncratic – in character. Such stories and logics are not rare, esoteric or derived from a single source: Polanyi writes that much of what he has to say "will be common knowledge among scientists" (p. 54). The widespread appropriation of Newton's aphorism, "If I have seen further, it is by standing on the shoulders of giants" (1675) – including in the tagline of the Google Scholar search engine – indicates that a story of science as communal, cumulative, and predicated on the availability of past knowledge, is well acknowledged in scholarly and public discourse.

The sociologist of science Robert Merton tells his version of this story with an extra twist: he theorises that the open communication and communal sharing of scientific findings is not only a technically efficient way of building knowledge, but a norm institutionally reinforced in scientists. This "communism" (later "communalism", Vanderstraeten and Eykens 2018), is one of four norms defining the ethos of modern science according to Merton (1973 [1942], pp. 267–275), and he describes it as follows:

> *"The substantive findings of science are a product of social collaboration and are assigned to the community [...] The institutional conception of science as part of the public domain is linked with the imperative for communication of findings. Secrecy is the antithesis of this norm; full and open communication is its enactment." (pp. 273-274)*

Merton himself identifies Newton's "shoulders of giants" aphorism with his norm of communism, observing that it "expresses at once a sense of indebtedness to the common heritage and a recognition of the essentially cooperative and selectively cumulative quality of scientific achievement" (p. 275). Merton's assertion is not that open, communal knowledge sharing is universally practised by scientists, but that optimally functioning systems of science will encourage these norms, which will be "in varying degrees internalised by the scientist" (p. 269). Merton's perspective goes beyond scientific openness as an abstract logic governing the accumulation of knowledge: he situates it as a cultural ideal with which scientists are trained to have a meaningful relationship. Merton has many critics: his ethos of science is often seen as an outmoded and flawed (e.g. Barnes and Dolby 1970; Mulkay 1976; Rothman 1972).

Mulkay in particular, argues that Merton's norms can be better understood as "vocabularies of justification" that have an ideological function, promoting the status of science and scientists to government (pp. 653-654). However, Merton himself recognised that normative structures do not preclude ambivalence and the operation of "counter-norms" (as documented and developed by Mitroff 1974). Merton's articulation of science as ethos-driven, including by open, communal ideals, seems to have popular resonance beyond its academic life in social studies of science: "If the classic Mertonian norms of science [...] do not exist, then it seems today's scientists intend to invent them" (Kelty 2012:159).

Fifty years after Polanyi described communally built jigsaw puzzles, the UK's Royal Society – of which he had been a Fellow – produced a report reinforcing a commitment to openness in science, titled *Science as an open enterprise* (2012). In reflecting on over 300 years of scientific activity, the authors cite Polanyi (1962) and re-tell a familiar story:

> *"Much of the remarkable growth of scientific understanding in recent centuries is due to open practices; open communication and deliberation sit at the heart of scientific practice. Publishing scientific theories, including experimental and observational data, permits others to scrutinise them, to replicate experiments and to reuse data to create further understanding..." (p. 13)*

In the report, this statement is a prelude to a new and different narrative about openness in science: the story of "open science" that unfolds in the twenty-first century.

3.3 Twenty-first century "open science"

I have shown above that the idea of openness in science is not new in the twenty-first century, but rather deeply embedded in traditional stories about both the logic and culture of science. Since the turn of the twenty-first century a new set of stories concerning scientific openness have emerged. Where traditional stories are reassuring about the essential openness of science, new stories are less sanguine: after affirming that openness sits “at the heart of scientific practice”, the Royal Society report (2012) raises concern that “much of today’s scientific practice falls short of the ideals of intelligent openness” (p. 16). Simultaneously, the report presents a grand new opportunity for scientific openness: “The internet […] may pave the way for a second open science revolution, as great as that triggered by the creation of the first scientific journals” (p. 7). Similarly in the US, the website of the National Institutes of Health (NIH) features a video titled “open science” that identifies deep problems with current scientific communication systems, and a need to “ensure the central values of the scientific process, which are openness, integrity, and reproducibility” (Nosek 2014). The video associates a need to redeem the values of science with the 2013 founding of a Center for Open Science, based in Virginia. Stories of crisis, opportunity and revolution in the relationship between science and openness are characteristic of an “open science” discourse that has rapidly gained traction in international scholarly research contexts over the last twenty years.

Twenty-first century “open science” is associated with a diverse but interrelated set of practices, often explicitly labelled as “open”, each of which is supported by social movements, technological innovations, and in some cases, research policy directives (Fecher and Friesike 2014). Some of the most salient are:

- **open access**: free, public, online access to – and reusability of – published research, especially in the form of scholarly journal articles (Suber 2012);
- **open research data**: free, public, online access to primary research data including reusability (Murray-Rust 2008);
- **open preprints**: free, public, online access to scholarly journal articles before their formal publication (Desjardins-Proulx et al. 2013);
- **open peer review**: free, public, online transparency about aspects of editorial review that are traditionally hidden (e.g. reviewer identities or comments, Ford 2013); and
- **open notebook science**: in which detailed research notes – including ongoing primary data, ideas, plans, and actions – are documented in an online, cost-free, and publicly-available format (Bradley et al. 2008).

This is a partial selection, and the list continues to expand. For example, the involvement of lay publics in science, including an "opening up" of traditional boundaries separating lay from expert scientific communities, is sometimes considered to be a component of open science, under headings such as citizen science, science blogging and "wisdom of the crowds" (Fecher and Friesike 2014:19–25). Moreover, there are components of open science that focus on "opening" discrete aspects of research processes or products beyond those above: these include preregistered open research proposals (Nosek et al. 2018); open methodologies or protocols (e.g. Teytelman et al. 2016); and open software, code, or models (Easterbrook 2014). Open science is associated with claims that traditional structures of scientific communication, in particular the journal article, need to be supplemented or replaced by new formats (e.g. Nielsen 2009:32). Further elements of open science relate to the way that research is indexed, measured and assessed, including the "opening up" of citation data (Shotton 2018), and challenges to the journal impact factor as a proxy for of research impact (DORA 2012). Open science is often explained using visual metaphors, especially an umbrella, but also e.g. a rainbow or a wheel, indicating that multiple practices, movements, technologies or policies are being brought together as part of a congruent whole (see **Figure 4**).

Figure 4 | Diagrams that use visual metaphors to explain open science. (a) A commonly used drawing of open science as an umbrella (지우 황 2013) CC BY 2.0; (b) open science "wheel" created by the European Commission's Open Science Monitor (EC 2017:6); "rainbow of open science practices" by Kramer and Bosman (2018) CC BY 4.0.

Despite the breadth and diversity of open science concepts, nearly all are strongly associated with Internet technology and cultures, and have arisen in their present forms following the invention of the World Wide Web in the 1990s. The speed, breadth and accessibility of information-sharing enabled by the Internet is seen as a step-change opportunity: Nielsen (2009) describes it as the "first major opportunity to improve" the way science works after 300 years of "surprisingly little" change (p. 30). Although each "open" practice or movement is pursued in its own right, there is a tendency for them to be viewed not only as congruent elements of whole, but as extensions of one another; even a moving frontier that extends "openness" throughout scientific practice and systems. This tendency is evident in talk of an open science "revolution" (Bartling and Friesike 2014b; Nielsen 2011:184; The Royal Society 2012:7) that evokes all-encompassing transformation of science, as well as the implication that science should be on a path towards "complete" or "extreme" openness through the progressive opening up of different elements. Nielsen (2009) urges the creation of an open scientific culture in which "everything – data, scientific opinions, questions, ideas, folk knowledge, workflows and everything else" (p. 32) is present in an online network. He expresses his vision as a logical extension of the older, traditional story about communal knowledge-sharing when he writes that:

> *"…extreme openness is the ultimate expression of the idea that others may build upon and extend the work of individual scientists in ways that they themselves would never have conceived." (ibid.)*

A moving frontier is also in evidence when advocates of "open" shift or expand their attention progressively from one "open" practice and concept to another. This can be observed in a broad chronological shift in advocacy, technology and policy focus that begins in many contexts – certainly in the UK – with open access to the research literature, progresses through the opening up of primary data underlying research findings, and moves towards the incorporation practices that are currently on the fringe of mainstream research culture in many fields, such as open peer review and sharing of research articles – preprints – prior to publication (see Chapter 5).

Contemporary open science, then, can be understood as a set of diverse yet congruent components – sometimes framed as extensions of one another – directed towards the pursuit of a transformed open science future, and operating largely via the Internet. It can also be understood as a set of social movements and practices that have developed in the grassroots of research and entrepreneurial communities, but are increasingly endorsed, encouraged and even enforced top-down by institutions that govern research. Each component of open science is associated with its own material practice, cultural history, and relationship with research policy: collectively, these would be too extensive and intricate to include in this thesis. However, some of these details are particularly relevant for contextualising my empirical research findings.

Chapter 5 is devoted to a close examination of open access and open research data in particular, as part of a framing analysis of "open" in advocacy and policy contexts.

3.4 Tracing traditional narratives of scientific openness

I have now introduced two dominant accounts of scientific openness that appear in scholarly and policy literature, and have compelling narrative qualities. The traditional narrative conjures a seemingly timeless image of communal knowledge-sharing. It is followed in the twenty-first century by an explicitly "open" narrative that reaffirms and extends the traditional one. In doing so, it evokes a future in which today's flawed science has been restored or elevated to an ideal state of openness – using the Internet. I present these narratives as powerful sources of meaning for open science movement(s): stories that seem to have been told and re-told, shaping scientists', policymakers', and open science advocates' worldviews and constructions of scientific openness. In this sense, my initial answer to research sub-question (a) is that the relationship between traditional and contemporary "open science" discourses is a complementary and additive one. Accounts of scientific openness as very old strengthen very new "open science" by rooting it authentically in history, as the contemporary pursuit of age-old principles that define science. Contemporary "open science" narratives do not discard traditional narratives: they build upon them.

This analysis is not intended to imply falseness or strategy in the framing of scientific openness, or to suggest that that contemporary "open science" is without a rich history. Nor does my use of the words "narrative" and "story" imply that accounts of old or new scientific openness are untruthful. Rather, I use "narrative" to indicate that common, overarching threads of meaning run through accounts of scientific openness and hold the intuitive appeal of a story – one which has often been based in evidence, but can easily be distanced from it. Stories have their own effects: Haraway writes, "it matters what stories we tell to tell other stories with [...] It matters what stories make worlds, what worlds make stories" (1988:12)[7]. The relationship between these narratives – the framing of one by the other – shows how they interact at an intuitive level. I have distilled these narrative threads to make them visible; in the subsequent sections I will begin to unravel them, reintroducing contextual detail that can be elided by a narrative focus, in an attempt to tell a more complex and nuanced history of scientific openness. I begin by contextualising – through literature review and analysis – the older, traditional narrative.

3.4.1 Historical origins of openness in science

The idea that science is built through open, communal knowledge-sharing has a familiar, logical appeal – at least from a contemporary Western perspective – which belies its contextual features and its contingent place in history. Attempts to trace the origin of ideas, practice or cultures of openness in science quickly encounter the flexible boundaries of both "science" and "open". For instance, McMullin (1985) argues that the origins of science as an open, public enterprise can be found "in the Greek-speaking world more than two millennia ago" together with the shaping of "science" itself (p. 14). He considers an early, Platonic, Aristotelian ideal of science to be open because of its emphasis on knowledge creation via the public realm: "Any assertion which lays claim to being *epistēmē* (science) must be justified, and the process of justification is in principle a *public* one" (ibid., original emphasis). Eamon (1985), on the other hand, considers that scientific openness was *not* evident in classical Greece despite the "paradigm of competitive public debate", because it "tended to solidify knowledge into separate schools" (p. 321). Thus even given similar historical reference points, McMullin and Eamon reach different conclusions because of their varying specifications of scientific openness: for Eamon, "collaboration among scientists directed towards a single goal" (ibid.) is an implied requirement, whereas a public locus of knowledge-making is sufficient for McMullin.

[7] I am telling another story about these stories of openness, of course.

Eamon's view aligns with the "traditional" story of scientific openness that I told at the beginning of this chapter, with the accounts of Polanyi (1962) and Merton (1973 [1942]), and with some other histories (e.g. David 2008), in that openness is assumed to be communal, collaborative and directed toward the accumulation of knowledge. Eamon (1985) and David (2008) place the emergence of this well-recognised form of scientific openness in sixteenth and seventeenth century Europe, and each author reconstructs its socio-technical development. There is a common understanding that this form of openness arose in contrast with prevailing cultures of secrecy in medieval Europe – cultures that not only concealed knowledge, but rationalised secrecy as an epistemic priority. Literary traditions portrayed the goddess Natura as being "modest, covered with a veil, and hostile to an open disclosure of her secrets"; social hierarchies and surrounding political and religious attitudes encouraged the withholding of such secrets from the "vulgar"; and alchemical writers used an esoteric language – "obscure symbols, paradoxes, allegories and secret names" – to protect their knowledge, which was considered divine and personal, rather than public (Eamon 1985:321, 325). The appearance of an "open" knowledge culture in this context was not inevitable: David (2008) in particular emphasises the "extended, intricate and contingent historical process" by which it took hold (p. 6).

The invention of the printing press in the fifteenth century is identified as a background condition that later enabled more "open" dissemination of scientific knowledge, but not a significant turning point (David 2008:52; Eamon 1985:322). According to Eamon, gradual cultural-conceptual changes mark the rise of communal, cumulative openness. New metaphors for science saw it is a hunt, an "aggressive pursuit after the deepest secrets of nature" (p. 334), and the idea of scientific progress emerged. This latter idea, so familiar in scientific cultures today, was "almost totally absent from the great philosophical systems of classical antiquity" (p. 335) which emphasised refinement of classical truths rather than innovation[8]. McMullin (1985:17) also notes this as a shift in epistemic culture evident between medieval and renaissance cultures. In the seventeenth century Francis Bacon

[8] The idea of progress as a quality of modern science is also open to critique; I note that the late twentieth century, when these histories were written, featured much theorising on the topic of scientific progress (e.g. Kuhn 1962; McMullin 1979).

influentially critiqued scientific secrecy and promoted a new rhetoric of progress via "co-operation and communication within the scientific community" (Eamon 1985:334, 338–39). Bacon's ideas apparently inspired the formation of "informal study clubs" in England and France who communicated their research through personal correspondence – letters (ibid., p. 342). The establishment in 1660 Royal Society of London represented an institutionalisation of this informal arrangement in which the secretary, Henry Oldenburg,

> *"…vastly expanded the Royal Society's contacts with the European virtuosi, and within the space of only a few years made the Society an international centre for the exchange of scientific and technical information."* (Eamon 1985:343)

The contents of correspondence to the Royal Society were aired at its meetings, and detailed notes of these meetings were recorded and kept; thus a communication network based on private letters became a systematic and "public" scientific record (ibid., pp. 342, 344) – albeit one limited to a certain gentlemanly class, race and gender. Simultaneously it became a means for establishing priority – the status of having been first to make a discovery – and according to Eamon this "induced many natural philosophers to accept the new norm of free communication of scientific information" (ibid., p. 344). This association between openness and recognition of priority is core to norm of communism as conceptualised by Merton (1973 [1942], pp. 267–275). This norm was further institutionalised and materialised through scholarly journals: the Royal Society's *Philosophical Transactions* was one of the first of these, initially published and circulated in 1665.

In his economic-historical analysis of same period and concept of scientific openness, David (2008) argues that it was not the appeal of institutionalised knowledge-sharing and recognition alone that set in motion a broad cultural change away from secrecy (pp. 31-32). He concludes that a fragmentation of political power in post-feudal Western Europe was a pivotal factor promoting an institutionalisation of "open science". In this context, aristocrats sought competitive advantage by sponsoring natural philosophers of exceptional skill. Because these powerful patrons were unable to assess skills personally due to increasingly complex mathematics, a demand grew for public advertising and peer assessment of intellectual feats, conducted through networks of correspondence visible to the ruling classes: "challenges could be issued, contests and competitions could be staged, and collegiate reputations could be both secured and widely broadcast" (ibid., p.69). It was in this fragmented and competitive context, David argues, that scholarly societies and academies began to institutionalize knowledge-sharing in a familiar form. David introduces his argument with a "Presentist" prologue emphasising the fragility of the open culture that came about in this contingent manner, perceiving it as vulnerable to reductions in public funding or incursions from infrastructures of private property (ibid., pp. 4-5).

3.4.2 Openness, secrecy and closure in social and historical context

Some of the historical accounts above observe that practices or cultures of scientific openness are not, and did not become, ubiquitous – whether in classical Greece or in renaissance Europe. Cultures of secrecy flourished alongside them: McMullin (1985), for instance, contrasts "open" Platonic, Aristotelian knowledge cultures with Eastern-inspired "mystery religions" of the Greek world in which knowledge was apparently constituted by its "occult and hidden character" (pp. 14-15) – much like medieval cultures of epistemic secrecy. And Eamon stresses that in renaissance Europe, "traditional exhortations against 'forbidden knowledge' were still widespread, in particular against looking into the secrets of nature, the secrets of God, and the secrets of the state" (p. 333). Although Eamon describes a cultural shift towards openness at least in regard to "secrets of nature", craft knowledge was not governed by the same institutional norms as established by the Royal Society. The Society's attempts, from 1660, to document craft knowledge and offer recognition to inventors fell flat when craftsmen proved unwilling to reveal trade secrets (Eamon 1985:344–45; Iliffe 1992:31). Mathematical traditions at the end of the seventeenth century were also far from open and contrasted with some knowledge-witnessing conventions of the Royal Society. It was routine to announce a mathematical result in cypher – code – in order to claim priority while protecting the discovery from "appropriation by competitors" (Iliffe 1992:33).

The story of scientific openness is thus not as clear as it first appears. Different definitions and traditions of openness can be identified with different historical starting points, and despite their apparently compelling ideological qualities – at least in the case of the communal, progressive openness of the European Renaissance – they become dominant only in particular knowledge contexts. And indeed, it may only be through accidents of history – such as the fragmentation of scientific patronage regimes – as David (2008) claims, that such cultures gain momentum. However, if one examines the history of scientific openness still further, a more elementary observation becomes inescapable: cultures of secrecy not only run parallel to cultures of openness; they flourish within them. Moreover, through a fine-grained contextual lens, the flexible boundaries of scientific openness start to crumble: secrecy or closure are intermingled with openness at the level of a culture, an institution, an individual, or even a single action.

3.4.2.1 Historical example of openness and closure in practice

Iliffe (1992) writes a compelling account of a controversy between two Fellows of the Royal Society, Robert Hooke and Christiaan Huygens, at a time, 1675, and within an institutional setting credited with the establishment of openness cultures. Each man claimed invention of the balance-spring watch. In principle, the Royal Society had mechanisms in place to air and adjudicate these claims to priority. In practice, these natural philosophers each deployed tactics involving varied forms of openness and closure in an attempt to secure acknowledgement and patenting rights. Huygens initially circulated knowledge of his watch in closed form: a cypher was sent to Oldenburg, the Society secretary, who was entrusted to register receipt of such claims to priority. Huygens was pushed to reveal the meaning of the cypher early, when the watchmaker he trusted betrayed a vow of secrecy and attempted to claim the invention. When Hooke learned of Huygens' watch, he leaned on the records of the Royal Society and the testimony of trusted members to claim that he had invented a comparable watch years previously – whilst immediately and privately getting to work on a new version. Having been drawn into a defensive position, the burden of evidencing claims was placed on Huygens, who was asked to send increasingly detailed documentation – and the watch itself – for consideration by the Society. Hooke was able to build a competing watch while hiding its mechanism – some versions may even have been engineered to break if opened – by leveraging his connection with the King Charles II, who could vouch for the watch's functionality. This prevented an assessment of similarity between the two watches, which in all likelihood worked differently and thus need not have competed.

This is an example of Hooke successfully bypassing the authority and "open" conventions of the Royal Society to secure his invention according to secretive trade norms. "From Hooke's point of view, no craftsman would ever have made such a secret public, and once Huygens had done so, Hooke assumed he was fair game…" (p. 54). It also an example of how cultures of openness can not only contain secretive instances and individuals, but systematically engender particular kinds of secrecy. Announcement and full disclosure of discovery were not necessarily simultaneous. In fact they were routinely decoupled, and the conventions of the Royal Society enabled this. Members could submit cyphers – as Huygens did – or sealed letters in attempt to establish their priority without revealing the key details. Even hints dropped in personal correspondence or conversations with other members of the Society could be presented as evidence of prior discovery. A dichotomy between openness and closure is unravelled in such contexts. Early intimations of knowledge designed to secure priority represent partial openness and partial closure depending on factors such as the depth of communication, manner of encoding or concealment, and the size and character of the receiving audience. These degrees of openness and closure are exposed in context as performances that do socio-technical work, asserting or ceding power. The precise manner of simultaneously revealing and hiding a discovery can buy one time to "perfect" it and an assurance of priority; it can alarm, quell or motivate one's competitors; and can elicit counter-moves that also mix openness with closure. In showing his watch to the King, Hooke selectively revealed his knowledge in a way that wields power, simultaneously ensuring he does not need to reveal the secrets manifest within the physical object – but that Huygens must, in order to respond. These power dynamics of intermingled openness and closure around claims to priority are recognisable in the twenty-first century experiences of the scientists I interviewed, and are discussed particularly in Chapter 8.

3.4.2.2 Theorising openness and secrecy in the history of science

Such reflections on the contextual and intertwined nature of openness and secrecy have not been conventional in the historiography of science according to Vermeir (2012), who examines these subtleties in early modern European science. Balmer (2012) does similarly in relation to UK government science during the Cold War. Both authors show through case studies the ways in which "open" and "secret" do not function in a binary manner, as states with an inside and outside, or as polarised opposites. Balmer and Vermeir both refer to "degrees" of openness or secrecy: Balmer invites us to "…think of a concentric circle as a zoned metaphor for secrecy and openness. When a secret is revealed, there is always going to be something left concealed…" (p. 146). Vermeir considers "…a range of gradations between full openness and extreme secrecy" (p. 170). Neither imagines these to be fixed or inert states; both characterise openness and secrecy as constructed in relation to particular objects, temporalities, geographies, and people; and importantly, as conditions that contribute to *constructing* those realities. Balmer, for instance, writes that:

> *"Secrecy changes science. A close reading of events in the history of chemical and biological warfare research and policy demonstrates how it is difficult to maintain the idea of secrecy as simply a veil drawn around an immutable activity. Secrecy is enacted or performed; it produces, alters, re-configures whatever comes within its ambit. Different questions get asked, different lines of research are pursued, normative judgements are altered and actions are taken…" (p. 145)*

Vermeir similarly refers to the active, formative roles of openness and secrecy, for example: "*openly showing* that there is secrecy involved, is the best way to spur your readers on, to fascinate them" (p. 188, original emphasis). Balmer's and Vermeir's observations and arguments are consistent with the case above documented by Iliffe (1992), and as well as with the interview observations reported in this thesis.

Vermeir in particular argues that openness and secrecy should not be seen in opposition to one another, but rather as two positive categories that "are often interlocked, impossible to take apart [...] they might even reinforce each other" (p. 165). Newton, for instance, pursued strategic forms of openness and secrecy simultaneously by disseminating his ideas widely in several forms – conversations, letters and manuscripts – whilst intentionally restricting them from some audiences including close competitors (p. 175). Vermeir and Balmer both test and explore the idiosyncrasies and limits of openness and secrecy, exposed in examples and thought experiments: e.g. a "secret" kept by over a hundred people, or an "open" conversation between ten people in a classroom (Vermeir 2012:169–70); an author unable to access his own government work when its secrecy is reclassified, and a group of missile technicians finding confidential design details – denied to them – on models accessible at toy shops (Balmer 2012:1):

> *"Someone on the inside inadvertently finds himself on the outside; people excluded from a secret later find they had unknowingly accessed the inside by the most unlikely means." (ibid.)*

These authors make it clear that openness and secrecy elude absolute definition; their manifestations and effects can be counterintuitive; they can coexist; and they are given meaning in subjective context, by those who enact and encounter them. In this vein, Vermeir (2012) provisionally concludes that an *intentional* pursuit of openness or secrecy defines these categories (pp. 171-176); whereas Long (2001) defines openness as a lack of restrictiveness. In this thesis I more often use the term "closure", which implies intention to a lesser extent than secrecy, anticipating in contemporary scientists' accounts gradations not only of openness and closure – including overlaps – but gradations of intention in producing them.

3.4.3 Boundary work and the construction of openness in science

An awareness of the situated, equivocal role of openness in the history of science makes traditional stories of openness "at the heart" of science (The Royal Society 2012:13) all the more remarkable. There is an eagerness to elide messy relationships with closure and secrecy in order to find within science an essence of openness. This is despite a temporal and cultural closeness of science to secret knowledge-making; and despite the interweaving of openness and secrecy in the everyday actions and intentions of scientists pursuing a public or private accumulation of knowledge. This leads me to two observations: firstly, I suggest that *boundary work* (Gieryn 1983) is a useful concept for understanding what is taking place: social, rhetorical work is being done to delimit the boundaries of "science" despite its inherently flexible qualities. Secondly, although the relationship between science and openness is empirically messy, a widespread subjective valuing of openness by scientists – that involves boundary work – could be considered an important part of (some) scientific cultures.

The boundary work being done in histories of openness in science is of the traditional type described by Gieryn (ibid.): a particular quality – in this case openness – is established rhetorically as a distinguishing factor between science and non-science. That which is not open, is, in this view, not truly science[9]. Then in addition, an extended, circular form of boundary work can also occur, because of the inherent flexibility of not only science, but of openness: almost any practice or culture which is felt to be truly scientific can also be found, by some definition, to be open. Instances or patterns of closure and secrecy can be found in most activities and used to question their scientific status. Gieryn associated boundary work with a pursuit of partisan interests, for example scientists' interests in presenting their own field of work as more truly scientific than others in order to gain funding. Through this interests-based theoretical framing, Gieryn creates the impression that boundary work is performed instrumentally – a superficial rhetorical device that cannot be based in the true character of science, as no such essential character exists.

Such interests no doubt participate in the boundary work around openness: Hooke, for instance can be found espousing the value of openness in his publication *Micrographia*: "the Arts of life have been too long imprison'd in the dark shops of Mechanicks themselves, & there hindered from growth, either by ignorance, or self-interest…" (cited by Iliffe 1992:55). This may be strategic rhetoric, as Hooke acted

[9] For an interview-based example, see footnote 10, p. 99.

secretively on many occasions including in relation to the watch, and may have had some interest in contrasting an illuminated academic science with the "dark shops" of craftsmen. In the contemporary context of my interviews, scientists would often espouse the essential openness of their research or identity, while in the same conversation acknowledging patterns of closure in their day-to-day practice[10].

3.4.4 Beyond boundary work: scientific openness a discursive culture

However my interpretation of boundary work in relation to openness cannot, like Gieryn's, end with interests: rhetoric is situational and inconsistent for many reasons; it cannot be assumed that scientists are adopting tactical stances rather than expressing deeply held convictions about the value of openness when they do boundary work. My suggestion is that boundary work is in operation, but as more of a discursive tradition than a rhetorical device – one that has meaning for scientists, and is itself part of (some) scientific cultures. A culture of valuing or idealising openness could exist in tandem with patterns of practice that include or even prioritise situational closures. Such a culture would shape a sense of professional identity and ethos in scientists, and condition practice indirectly.

My contention is, then, that widespread subjective, discursive valuing of scientific openness could be considered part of a culture of openness in science. This may be of little interest to scholars who are most concerned with evidence of openness in practice: with what scientists do, rather than what they say. Barnes and Dolby (1970) for example distinguish between "professed" norms, which are "enjoined or celebrated in tract or speech", and "statistical" norms, "observable as a pattern of positively sanctioned activity" (p. 8). These authors described professed norms as of little value for understanding "ongoing scientific activity", except to the limited extent that they guide scientists' actions. Their article is a critique from Merton's norms (1973 [1942],

[10] For instance, Gavin asserted that "science is science if it's open. Whereas the things that corporations do is...I hate to go into the divide, it's almost like technology". At other times, he spoke about the necessity of certain forms of closure in (academic) science, including his own.

pp. 267–275), which are portrayed as unreliable extrapolations of professed norms. Their characterisation treats discursive acts as mere epiphenomena of real, "statistical" scientific activity, and turns away from ideas of ethos and identity in science because they are not considered necessary for theoretical explanations (ibid., pp. 23-24) when science is viewed as a sum, variety, or pattern of practices producing knowledge. However, I contend that the way scientists talk about their practice and profession is a component of scientific activity, even if it is not categorised as a "statistical" behaviour or a useful predictor thereof. Moreover, where openness in science is understood to be relational, subtly blended with closure, and conditioned by intention, scientists' subjective accounts offer insights that cannot be gained by observation from the outside: scientists' discursive constructions of openness, however complex their influences, must be taken seriously as instantiations of scientific culture. When the subject of the scientist is centred in this way, ethos and identity – and virtue – become relevant theoretical tools, albeit without Merton's essentialist framing (see Chapter 2).

3.4.5 Boundary work in scholarly histories of scientific openness

I maintain, however, that a cultural valuing of scientific openness results in – or is blended with – rhetorical boundary work that persists in simplifying and eliding a complex relationship between science, openness, and closure. This consideration is particularly acute in historical analysis, because of the rhetorical medium of historical narratives, and the subtle yet potent effect of present-day influences on views of the past. I suggest that key, influential scholarly histories of scientific openness reflect and participate in a cultural valuing of scientific openness, through a performance of boundary work. This is evident from the outset of my historical literature review (Section 3.4.1), where scholars co-locate the origins of "open" knowledge traditions with the origin of science – despite varying definitions of both openness and science.

McMullin (1985) does this by asserting that the idea of science "as an open enterprise […] goes back to the very beginnings, to the time with the notion of 'science' first took shape" in classical Greece (p. 14). Eamon (1985) and David (2008) perform a variation on McMullin's boundary work through a rhetorical co-construction of openness and *modern* science. Eamon's paper for example begins by describing various forms of openness as "essential features of modern science" (p. 321). David's account is particularly notable because he argues that "open science" ideas and practices were "a distinctive and vital organizational aspect of the Scientific Revolution" (p. ii) – the period synonymous with the birth of modern Western science. However, the existence of the scientific revolution as a "singular and discrete event, localized in time and space" is disputed by many historians (Shapin 1996:3). David's maintenance of this historical construct allows him to associate the origins of "open science" not only with a gradual emergence of modern science, but with the moment and cause of its conception – a powerful notion. While these histories depict complexity, scientific secrecy, and a contingent emergence of openness, they still in their narrative outlines use openness to reinforce the boundaries of science as we know it today. This scholarly boundary work is significant not only because it seems to reflect a cultural valuing of scientific openness – from the 1980s at least – but also because it influences today's constructions of scientific openness in open science movements.

3.4.6 Presentist applications of "openness" and "open science" to history

Boundary work in key histories of scientific openness has an additional feature that is worthy of note. These histories appear to introduce or add new emphasis to the terms "openness" or "open" as descriptors of public, communal, or cumulative knowledge-making, where such language was not necessarily central in past discussions of science – even those in English-language contexts of only a few decades previously. For instance, Merton (1973 [1942]) first wrote about the ethos of science including "communism" the 1940s and Polanyi (1962) evoked science as a communal jigsaw puzzle in the 1960s (see Section 3.2). Neither focused upon "open" terminology in these texts: Merton used "open" only once, in passing, in his characterisation of communism: "full and open communication" (p. 274); Polanyi did not use it at all in his article. When these texts are later interpreted, however, they are seen as foundations for "openness": Eamon (1985) and David (2008) both draw on Merton to anchor their histories of "openness in science" (Eamon) and "open science" (David); and Polanyi's article is cited by the Royal Society (2012) to support a claim that "open communication and deliberation sit at the heart of scientific practice" (p. 13).

These instances suggest that both the word "open" and its conceptual connotations are – subtly – presentist framings, reflecting more recent vocabularies and concerns associated with science. McMullin (1985), Bok (1982) and Long (2001), as well as Eamon (1985) and David (2008), all centre "openness" in their histories of public, communal, cumulative science. While this could be coincidental or part of a localised scholarly trend, it may reflect wider evolutions of language connected with societal developments from the 1980s onwards that make the "openness" a pressing concern[11] – for example, a rise in neoliberal politics, and private science (Agar 2012:440; Marginson 1996; Thackray 1998). Indeed, David (2008) explicitly frames his historical analysis with a "Presentist" prologue warning of incursions from the private realm (p. 4-5). And strikingly, David – whose history is the most recently written of the set, in 2008, centres the contemporary buzz phrase "open science" in his title, where the other authors use "openness". Thus, he introduces a phase connoting twenty-first century open science movements, and uses it to once more reframe the traditional, communal science that once may never have been characterised as "open" – and certainly not with the term's contemporary connotations.

As Vermeir (2012) observes, presentist or transhistorical analyses of openness and secrecy have an advantageous breadth of scope, but they divert emphasis from "the values as well as the practices of openness and secrecy [that] have varied strongly throughout history" (p. 167). Vermeir's recommendation is that:

> *"In order to historicize the discussion of openness and secrecy, we must pay close attention to [historical] actors' categories of analysis, the distinctions and oppositions they make, and we must differentiate between kinds of secrecy [or openness] … This should allow us to enrich the relevant historiographical vocabulary, to be more sensitive*

[11] In his history and analysis of Free Software, Kelty (2008) writes: "In the 1980s everyone seemed to want some kind of openness, whether among manufacturers or customers, from General Motors to the armed forces" (p. 148).

to lexical changes and relations between concepts of actors and historians, as well as to include a wider array of historical practices in our analysis." (ibid.)

Analyses like David's (2008) and those listed above are rich in detail and historiographic reflection. Their "open" framings tend to be just that – framings, present in terminology and narrative outlines. However, they remain powerful: particularly because they condition way these histories travel, no more so than in the present online era when new and influential "open science" movements seek to understand their histories, and literature searches for the history of "open" in science readily yield a backstory. David's work in particular is drawn upon in contemporary advocacy of "open science" and has contributed to the framing of science as open at heart for example by Nielsen (2011), who was subsequently cited by the Royal Society (2012).

It would be an exaggeration to say that past ideals of science as open are figments of our contemporary imagination; however, I have shown here that the interpretative flexibility of both "science" and "open" make it easier to conclude that the pursuit of scientific openness has been continuous historically or has always meant the same thing at heart. Moreover, broad outlines of historical analyses such as David's (2008) have been incorporated into contemporary discourses, where they position open science movements as reassertions of essential truths about science. On a reflexive note, this very movement of information, eased along tracks laid by "open" terminology, undoubtedly shaped the manner in which I came to construct its history: this chapter begins with the historical narrative that I first encountered as a researcher familiar with contemporary open science, searching for its history.

3.4.7 Situating and specifying historical "open science"

Following Vermeir's (2012) recommendation above, we can make sense of past modes of openness in science by seeking out and emphasising specific, historically situated qualities of "openness" – and also of "science". This builds a bridge between neat and messy accounts of the relationship between these two words and concepts. Some of these qualities I have articulated from the outset, but they warrant emphasis. Portrayals of past openness in science (e.g. David, 2008; Eamon, 1985; Merton, 1973 [1942]; Polanyi, 1962; The Royal Society, 2012, p. 13) – whether labelled as "open" or later interpreted as such – depict a public, communal, centralised accumulation of knowledge rather than individual, idiosyncratic, local, or unsystematised knowledge work. These qualities are not abstract, but conditioned by historically located socio-technical infrastructures: scholarly societies; their coordinated systems of written correspondence; and ultimately the systematic verification, compilation, publication, and circulation of physical scientific papers in scholarly journals. Communal and public features of past science can be historicized by examining notions of "scientific community" and "the public" – their origins, and their limits. These specifics allow a clearer view of how past science could be considered "open", and how this may differ from, or align with, contemporary "open science".

There is a well-acknowledged link, including in contemporary advocacy discourses (Bartling and Friesike 2014b; Nielsen 2011) between past scientific openness and the history of journal publication. This aligns with a historicized reading of Merton (1973 [1942]) and Polanyi (1962). Both, in their oblique references to publication, portray it as the taken-for-granted medium through which knowledge becomes communal. An article by Zuckerman and Merton (1971) confirms that the "technological basis" for the Mertonian norm of communism is the printed word (p. 69)[12]. Thus it appears that by the mid-twentieth century, journal publication had become a practice that defined public, communal, cumulative science. This does not mean that journal systems then held the same status as early as the seventeenth century, when the earliest of them began; but it does anchor *an* historical vision of scientific openness that predates today's phenomenon.

The communal setting of early "open" science must be acknowledged alongside its infrastructural basis in publication systems, because the two are intimately

[12] However, Merton's famous evocation of "full and open communication" (1973 [1942]:274) makes his norm of openness highly appropriable outside the print/publication context.

connected: scholarly publishing was established by academies and learned societies which were operating as communities, establishing group identities; and in turn, scholarly publishing practices have shaped notions of community. This means that the history of scientific openness is the history of scientific community as much as it is the history of scholarly journals, or the history of written communication or publication of science. The development of collective social structures and relationships in science has arguably been as important for the emergence of historical "openness" as technical, conceptual or political precursors such as the invention of the printing press, ideas of scientific progress, or the fragmentation of power in post-feudal Western Europe (see Section 3.4.1). A focus on scientific community also brings to the fore a situated characteristic of historical openness that is often left implicit: communities are bounded. "Openness" of this kind is practised by, and for, certain people: an in-group. It is illuminating that framings of community and communality are far more central to Merton (1973 [1942]) and Polanyi's (1962) accounts of science than any mention of "openness". Moreover, the openness, communication, visibility and collaboration they describe are not universal but located between community members.

A contemporary interpretation might be that in the past, scientific "openness" occurred between scientists – members of "the scientific community"– and not with "the public"[13], distinguishing it from publicly-oriented values and goals of today's open science movements. However, early scholarly societies and academies pre-date the professionalization and systematic state funding of science, meaning that boundaries between scientists and the public were drawn differently. The early members of the Royal Society tended to be gentlemen, distinguished not primarily by their expertise but by their social class: Skinner (1969) described it as "like a gentlemen's club" (p. 238). Shapin (1988, 1994) shows that during this time and place, notions of truth and trustworthiness were bound up with codes of English gentility such that class was treated as a qualification for knowledge-making: "roughly speaking, the distribution of credibility followed the contours of English society…" (1988:376). Similarly, the witnessing of experiments in the Royal Society's "public"

[13] Both phrases used by Polanyi (1962).

rooms was open – to a certain echelon of society – and the houses of gentlemen scientists were “open to the legitimate visits of other gentlemen” as a “point of honour” in codes of gentility (Shapin 1984:488, 1988:387). It was not until late nineteenth or early twentieth century that comparable boundaries were drawn according to professional qualification and expertise, and “the public” was configured as a distinct outside group: “the professional world of science and the popularization of science to a variety of publics emerged together [...] before 1900 there was no professional science from which we would distinguish popular science” (Pandora and Rader 2008:355). Therefore early scientific communities, to the extent that they resembled the Royal Society, had a public, vocational character that was closer to the “lay” realm than today’s professional science – chiming with notions of past scientific openness. However, this past openness was exclusive according to class, gender, and race at the very least – forms of closure with a degree of continuity to the present day. By the mid-twentieth century, Merton and Polanyi were characterising a science significantly altered not only by professionalization, but by the introduction of systematic state funding, endowing some settings of science with a new “open” orientation associated with the taxpaying public. This public orientation is again, historically located rather than an essential scientific quality.

These points are not exhaustive, but demonstrate how past scientific “openness” is configured in relation to particular social and historical circumstances. Narratives that frame science as traditionally or essentially open – as distilled at the opening of this chapter – can in this way be taken to describe specific, situated kinds of “openness”. Descriptors like public, communal, and cumulative – and socio-technical settings, like scholarly journal publishing and scientific community – can be used to characterise some scientific traditions of decades or centuries ago without the implication that they are essentially open in a way that we would recognise today. This distinction is important because of the particular weight that “open” holds in contemporary knowledge movements.

3.5 Conclusion: relating older, traditional "openness" to contemporary "open science"

Having evoked, unpacked, and situated historical openness in science, I address my first research sub-question in several ways. *How do contemporary "open science" discourses relate to older, traditional discourses about the essential value of openness in science?* My first observation has been that these two discourses are not independent, and cannot be clearly separated. They both tell a story of openness as a core, defining feature of science – in one case defining science in its earliest form, whether in ancient Greece or in renaissance Europe; and in another defining what has been lost, and must be regained, in today's science. In a narrative sense, traditional discourses about a long history of open science seem to have a function: strengthening contemporary narratives by supplying depth and authenticity. Flexible qualities of both "open" and "science" ease this alignment. This interpretative flexibility also makes it difficult to trace the history of scientific openness. Many accounts trace it to seventeenth century Western Europe, co-locating it with the first scholarly journals and the birth of "modern science" – and with a repudiation of secret, obscure knowledge. Closer examination of openness and secrecy as social practices, however, shows that the two are often intermingled at the level of a culture or an individual, or even a single action. It is important, then, to understand seventeenth-century Western science as open in specific and situated ways. I have suggested that it is the public, communal, cumulative qualities of some science in this period – enabled in particular by journal publishing and community structures – allow us to identify it as "open". This traditional kind of scientific openness, however, is distinguishable from today's open science – it has shifted as the contexts that define it have changed.

Interpretative flexibility allows a circular kind of boundary work: openness distinguishes science from non-science or bad science, where both openness and science are in the eye of the beholder. I argue, however, that this is not necessarily an instrumental or purely rhetorical ploy to co-define science and openness. A culture of valuing or idealising openness despite the messy reality of scientific practice seems to be an important subjective dimension of professional identity and ethos for scientists, albeit one that it difficult to trace historically. This difficultly arises partly because some historians participate in a cultural valuing of scientific openness, engaging subtly in boundary work. Through this route, "open" terminology itself seems to emerge from present-day concerns, and be carried into retellings of the past. As late as the mid-twentieth century, Merton's choice of words for his norm was "communism" rather than "openness". But by the 1980s, scholarly histories of "openness" in science were being written; by the 2000s, David's (2008) was a history of "open science" – today's buzz phrase. My suggestion therefore is that discourses about an older, traditional openness in science do not only – in a narrative sense – add depth to contemporary "open science" discourses. I propose that these "older" discourses are part of a recent phenomenon that re-imagines past science in relation to "openness" and subsequently "open science" – in response initially to neoliberal societal trends from the 1980s, and then in response to an explicit and organised pursuit of "open" in science from the 2000s.

I conclude therefore that a long history of scientific openness is undoubtedly important: traditions of public, published, communal, cumulative science can be traced back hundreds of years. This history, however, is deeply entangled with our present perspective that places great and specific contemporary value on "openness", and is eager to see this unusually flexible quality reflected in the past. In their outlines and emphases therefore, historical accounts of openness accounts illuminate the present. In their specific details, they situate past science, allow continuities and contrasts to be drawn. I turn in the next chapter towards a more recent history, leading up to the emergence of contemporary "open science" in the twenty-first century.

Chapter 4 | “Open science is the box…”

Twentieth century preludes to “open science”

4.1 Introduction

Contemporary open science is animated by powerful narratives, reviewed at the opening of the previous chapter. These narratives tend to evoke a long history of openness, affirming its essential place in science, and to communicate two types of subtext: firstly, that openness of science is now lacking, and must be restored; and secondly, that the Internet provides an opportunity to reach an unprecedented, even ideal, state of openness in science. Unlike their historical counterparts, these contemporary narratives and discourses make explicit and frequent use of the term “open”: e.g. open access, open research data, open peer review, open notebook science. These narratives, movements and practices seem to appear suddenly in historical terms, rising to salience in international scientific discourses around the turn of the twenty-first century, seemingly prompted by the technological capacity of Internet technologies – in particular the World Wide Web from the 1990s (Berners-Lee et al. 1994) and subsequently the more interactive qualities of Web 2.0 (Procter et al. 2010:4039).

While Internet-based technologies have undoubtedly been a key and proximal influence enabling the emergence of contemporary open science, my purpose in the coming sections is to document a much richer social and technological twentieth century history that set the stage for this phenomenon. Each element of this history is well known in certain fields, but these elements have not typically been gathered together as a holistic backstory to the emergence of “open science” – perhaps because they do not all have clear links to openness, and some are more linked with closure or secrecy. Aspects of this history have been traced in a holistic way, however (e.g. Delfanti and Pitrelli 2015; Prainsack and Leonelli 2018; Strasser 2019). I aim to add capture extra breadth in order to orient towards research sub-question (b): *How has "open" in science recently gained such salience and discursive power, despite its multiple meanings and lack of specificity?* As well as addressing this question, I show in this chapter that the significance of both "open" and "science" shifts during the twentieth century. This means that the appearance of "open science" in the 2000s is not straightforwardly an "opening" of science. Despite my aim to capture breadth, it is important to note that this history is not exhaustive, and is centred in the global north.

4.2 Changes in scale: big science

During the twentieth century, the growth of science – "perhaps its most notable historical characteristic" (Capshew and Rader 1992:3) – became a topic of study under the heading "big science". This concept was established by physicist Alvin Weinberg (1961, 1967) and extended by historian of science Derek de Solla Price (1963). The term refers to growth across multiple dimensions, including the size of scientific teams; the ambition and cost of projects; and the physical scale, complexity, and expense of scientific instruments: "money, manpower, and machines" (Capshew and Rader 1992:4). It can also refer to generalised statistical growth, for example in numbers of professionally trained scientists and scholarly journals (Price 1963:6–7); as well as intensified qualitative features such as centralisation, industrialisation, internationalisation, and multi-disciplinary collaboration (Vermeulen 2016:199–200). Although science has grown throughout its history and there are earlier examples of large-scale scientific ventures (Capshew and Rader 1992:19–22; Price 1963:1–32), "big science" was conceptualised in relation to types of twentieth century physics that were centrally organised, concentrated around large and expensive machinery, and supported by government as a source of power in relation to war (Vermeulen 2016:207). Vermeulen shows that "big biology" emerged towards the end of the twentieth century – the Human Genome Project is an iconic case – but tends to be decentralised across virtual collaborative networks, and has been shaped by a more conditional, neoliberal funding environment. Big science is not self-evidently skewed towards openness or closure, but is an altered form and context of science, with different relationship to society.

4.3 Organisational "openness" and preprints in big physics

Archetypal big physics created social structures far removed from the self-coordinating community of independent scientists idealised in Polanyi's (1962) account. Agar (2012) refers to the "large-scale hierarchical teamwork" of wartime and post-Cold war physics in which researchers work as "managed populations" (pp. 309-310). Agar quotes an academic administrator, who stated in 1953:

> *"There was a time when scientific investigation was largely a matter of individual enterprise but the war taught scientists to work together in groups; they learned to think of a common project, they were impressed by the progress to be made through unified action. A notable degree of this spirit has been transfused into the life of our larger universities…"* (quoted in Agar 2012:309–10)

These organisational features have the potential to strengthen community ties and decrease competitive pressures, encouraging and enabling earlier, faster, less formalised communication: a form of openness in today's discourse. Kling and McKim (2000) described such patterns in the iconic big science of high energy physics: more willingness by physicists to trust and circulate unpublished working papers – partly due to internal institutional review processes – and more "mutual visibility of ongoing work" than in other fields (p. 1313).

These are the kinds of conditions under which preprints became a traditional communication format for high energy physics communities from the 1960s. Originally preprints were paper manuscripts, photocopied and bulk-mailed to other physicists and their institutions, simultaneously with their submission to peer-reviewed journals (Aymar 2009:36; Taubes 1993). Librarians at the major facilities created a "bibliographic infrastructure" of preprints: a more extensive and more immediately available resource for scientific knowledge than could be gained via the selective and formal processes of scholarly journals (Kling 2001:597–98). This preprint system preceded the rise of online open science by many decades, and provides an example of a cultural transformation towards openness that did not require the Web; and in which cultural, organisational, and economic aspects of big science had particular influence. This culture did, however, translate well into an online context in 1990s in the form of arXiv, a ubiquitous reference point in mathematical, physical, and other quantitative fields (see Section 4.8).

4.4 Secrecy and openness in wartime

While big science had the capacity to engender organisational openness, it is also associated with conspicuous instances of scientific secrecy during wartime. The Manhattan Project, an emblem of big science, was not only secret to outsiders; groups within the project were forbidden from communicating openly with one another (Vermeir and Margócsy 2012:159). However, this secrecy was not universal, nor was it mutually exclusive with certain forms of openness (as discussed, p. 72). Some academic scientists experienced a kind of opening up through their involvement in wartime efforts: enriched connections with fellow scientists and a wider community; closer familiarity with contexts of application; and a sense of efficacy and freedom associated with increased funding. Irish crystallographer and molecular biologist J. D. Bernal reflected in 1975:

> *"As the war proceeded the scientific workers found themselves less and less isolated and more and more part of a general directed effort. By coming closer to actual experience, they learned the existence of problems at first hand [...] It was most astonishing and invigorating to the academic scientist, cramped for years in a small laboratory and perpetually begging from this or that charitable fund for minute grants, to be allowed to spend thousands of pounds where he had previously spent ten pounds..." (pp. 557-558)*

Bernal also described how networks of scientific communication in WWII played a role previously fulfilled by scientific societies, but with a more strategic focus (ibid., p. 558). This illustrates a profound upheaval of factors that situated "traditional" scientific openness – community, communication, and relationship with the public – during world wars of the twentieth century.

4.5 A post-war "scientific information crisis"

Emerging out of WWII there was "the perception of a scientific information crisis", according to Wykle (2014:1), which challenged the primacy of journals as the arbiters of scientific communication. Bernal (1975) wrote that publication systems had already been "completely overloaded" before WWII and were ill-equipped to tackle a rising tide of knowledge, not to mention having been damaged by wartime shortages of paper and labour, and a loss of indexing and summarising functions from Germany (p. 571). However, Bernal also wrote that by the end of the war, "a far more effective system of scientific communication [...] than most scientists had enjoyed before the war" was in place (p. 560). This was an intersecting background context for the rise of preprint culture in high energy physics in the 1960s, and one which demonstrates that there was a widespread desire for alternatives to traditional publication. By 1959, UNESCO was studying "the increasing inadequacy of the scientific research periodical as a method of communication" (Wykle 2014:3).

In fact, in the same decade as preprints were established in big physics – the 1960s – they were trialled in biology as part of an experiment by the US National Institutes of Health (NIH). Seven "Information Exchange Groups" (IEGs), each based in a different specialist community, were funded to exchange preprints (ibid., p. 4). Although most of the IEGs were "successfully accomplishing rapid informal selective communications", the NIH ended the experiment in 1967 citing an inability to fund its growth (Confrey 1966). One IEG chairman reported that "opposition from scientific journals had a crucial influence on the decision to terminate the IEGs": this included a refusal by the editors of several major biochemical journals to publish articles that had circulated through IEGs, or to allow published articles to circulate subsequently (Till 2001:10).

Thus, in the mid-twentieth century many scientific fields were caught in tension between the infrastructures of traditional journal-based communication, and more immediate, informal practices that could be described as open. It was not only the organization of a field – "big" or little, close-knit or diffuse – that weighted this tension, but also, perhaps decisively, how power was wielded by actors controlling the publication system.

4.6 Big, commercial publishing and the "serials crisis"

When biochemical journal editors opposed IEG preprints, they were not merely preserving the established role of journal publication in legitimising knowledge: they were protecting a type of commercial interest that gained particular value during the latter half of the twentieth century. After WWII, a number of factors combined to make scholarly publishing into big business. Previously the publication of journals by societies and university presses tended not to yield profit: costs of production were significant, the market was small and uncertain, and subsidies were required to continue a mission of scholarly dissemination (Fyfe et al. 2017; Guédon 2001). After the war, academic research expanded, associated with growth in both publications and the budgets of university libraries, and creating a ready market (Fyfe et al. 2017; Guédon 2001) . Commercial publishers were adapting by taking on the publication of primary research rather than news and short reports; by selling directly to institutions who could afford higher costs per subscription; and by expanding into an international market (Fyfe et al. 2017:9). Economies of scale incentivised corporate growth, and from the 1960s larger commercial publishers began to gain control of publication previously in the hands of societies and other smaller publishers. By 2017, via acquisitions and mergers, an "oligopoly" of four corporations – Springer Nature, Elsevier, Wiley-Blackwell, and Taylor & Francis – had emerged, each of which now publishes thousands of journals (ibid., pp. 9-10).

The transformation of scholarly publishing into big business was also influenced by metrics that reconceptualised and filtered a growing literature. In the 1960s, Eugene Garfield created the Science Citation Index (SCI): a list that drew attention to a "hard core" (Garfield 1970:670) of journals capturing the "essential scientific literature" (Gaudet 2014:6–7). In association with the SCI, Garfield also created the journal impact factor (Garfield 2006) – a metric which has gained particular power as a marker of prestige, as is evident in my interviews with scientists (see Chapter 6). The construction of "core" journals guided library subscription policies, resulting in an "inelastic market" in which publishers could increase prices without reducing demand (Guédon 2001:28); as a result, large publishers established extremely high profit margins (Gaudet 2014:7). Also in the 1960s, journals were positioned as exclusive news sources by the "Ingelfinger rule": a policy of rejecting papers that been published in some form previously (named after the then-editor of the New England Journal of Medicine; Altman 1996; Gaudet 2014:6). This was the kind of editorial stance that suppressed experimentation with preprints in the biological sciences (Till 2001:10). The Ingelfinger rule also reinforced peer review as a dominant marker of credibility, and decreased the authority of knowledge exchanges outside the "merchant relation" of journal subscription (Gaudet 2014:6), tightening the grip of commercial publishers.

These developments set the scene for the "serials crisis" beginning in the early 1970s, in which the cost of journal subscriptions rose at a rate above inflation that made them less affordable to academic libraries (Eve 2014:13; Fyfe et al. 2017:14; Suber 2012:29). As well as price increases, this crisis was fuelled – again – by a growth in the production of academic research (Eve 2014) and a reduction in university funding in the 1980s (Fyfe et al. 2017). This escalating struggle to afford subscriptions, resulting in restricted and unequal access to research on an institution-by-institution basis, is a highly recognisable "closure" of science that has directly motivated open access movements (Suber 2012). Eve (2014) identifies this economic history of journal publishing as one of two converging antecedents to open access, the other originating in computer software cultures (see p. 103).

However once again it is clear that during this period the meanings of both science and openness changed. Scientific communities had grown and been enabled by commercialisation to disseminate more findings – greater "openness", in one sense. But the scholarly societies around which they had been organised were reconfigured by a profit motive. Corporations had the power to reshape communities – now, as markets for exclusive content – whilst retaining the allegiance of academics as authors, editors, peer reviewers and readers who seek association with "core" prestigious journals. If publishing was considered a traditional form of community openness, it had now become metrics-driven, internationally competitive, and exclusive. It was not merely act of dissemination, but a symbol of status and success beyond the published finding – driven also by output- and metrics-based evaluations of academics' professional reputations (Fyfe et al. 2017:3–4, 7, 13). Under these conditions, there was not only an overt curtailing of scientific openness according to institutional affiliation and wealth. There was also a narrowing of avenues for legitimate and well-regarded scholarly communication, and a circumscribing of good quality or successful science – and scientists – according to prestige-driven branding[14]. Alongside the serials crisis, these broader historical features shape the science that was ripe to be "opened" in the twenty-first century. When Nielsen (2009:30) asserted that publication systems have "changed surprisingly little in the last 300 years", he was referring to the technology of publication, rather than these social and economic transformations that profoundly altered the significance of publishing as "openness" before its move online.

4.7 Shifting modes of knowledge production, and commercial biology

Multiple studies have theorised a late twentieth century shift in modes of knowledge (reviewed in Hessels and van Lente 2008). One of the most influential accounts argued that there has been a rise in "Mode 2" science, which – instead of being primarily academic, disciplinary, and autonomous (Mode 1) – is heterogeneous, socially accountable, transdisciplinary, and "carried out in a context of application" (Gibbons et al. 1994:3). Other analysts have proposed a new "post-academic science" (Ziman 2000:67–82); a rise in "academic capitalism" in which university science is increasingly market-driven (Slaughter and Leslie 1997); and the "Triple Helix" model in which industry, university and government sectors are increasingly interlinked (Etzkowitz and Leydesdorff 2000). The concept of Mode 2 has been critiqued,

[14] Parts of open science movements that aim to move away from narrow, metrics-based prestige have their roots here (e.g. DORA 2012)

including by the latter authors, who point out that science has been situated in contexts of application – including industry – for centuries, and argue that Mode 1 arose later, with a nineteenth century institutionalisation of academic science (ibid., pp. 115-116). Similarly, historian Jon Agar argues that "science, located in working worlds, has always, typically, been more mode 2 than mode 1" (2012:434). However, even though Mode 2 has a long history, the latter decades of the twentieth century have featured changes – including an intensification of market-driven activity in universities - that are significant on a shorter timescale and in particular disciplines. Certainly, such changes shape the collective memories and identities of today's academic scientists, as documented by Holden (2014) in a UK biomedical context.

A shift towards entrepreneurial, application-focused research may seem particularly significant to scientists in biological fields due to the rapid commercialization of molecular biology in the 1970s and 1980s (Wright 1986). Between the early twentieth century and the invention of recombinant DNA techniques in the 1970s, molecular biology was "almost exclusively an academic field" (Wright 1986:303); it thus witnessed a form of the originally proposed transition from Mode 1 to Mode 2. In these latter decades its promise of practical applicability led to a boom in private investment, and there was a rapid cultural change in which molecular biologists became "equity owners, corporate executives, and consultants" (ibid.). This transformation was centred in the US, encouraged by a shifting relationship between science and government: market deregulation policies encouraged private investment in research with little oversight (ibid., p. 345). Commodification was stimulated by the 1980 Bayh-Dole Act, which allowed universities to patent discoveries or inventions arising from government funding (Agar 2012:440). Deregulation agendas affecting university research became widespread around this time, including – pertinent to my study – in the UK and Australia (Abir-Am 1997:514; Marginson 1996:78–81). By 1990 the Human Genome Project had begun, establishing an era of "new biology" in which molecular life sciences not only scaled up, but gave rise to an interdisciplinary, cross-sectoral, capitalist "bioeconomy" focused on health, and featuring upheavals to the conventional pharmaceutical industry (Mittra 2016).

These shifting modes of knowledge production, and intensified networks of academic–industry relationships, may set the stage for "open science" in multiple, contrasting ways. On one hand, traditional commercial contexts bring in patenting, trade secrets and other competitive practices that conflict with the notional Mertonian norm of communism (1973 [1942], p. 275). Where such practices inhibit publication or other expected ideals and practices of "openness", especially in traditional academic settings, they may add to a sense of breakdown in science – similar to the loss of a "golden age" lamented by biomedical scientists in Holden's (2014) analysis. On the other hand, neoliberal economics strike a chord with other narratives around openness in science: for instance, the idea that circulation of knowledge should be as regulation-free as possible in order to maximise scientific, economic, and societal benefits. Polanyi (1962) made explicit his view that the "invisible hand" of the market produces "maximum advancement of science" when scientists are acting freely and independently, and sharing their findings (p. 56). The neoliberal appeal of openness, which has been noted on multiple occasions (Lawson, Sanders, and Smith 2015; Mirowski 2018; Moore 2017; Tkacz 2012) no doubt influenced the form and political acceptability of "open science" as it emerged into a neoliberal political environment in the twenty-first century, even though some forms of openness intend to counteract commodification (e.g. Moore 2019b).

4.8 Networked computing and the early Internet

Alongside these late twentieth century dynamics, the Internet was in development. Like early big science projects, it began with generous funding from military sources: ARPANET, the foundational architecture of the Internet, was founded by the US Defense Department in the 1950s (Castells 2002:2–3). From the beginning, the project was close to science and academia: it drew on the expertise of an "informal, yet exclusive" network of university researchers, engineers and computer scientists who were given great autonomy over their work (Castells 2002:3,11).

The Internet, email, and the World-Wide Web – all developed in research contexts – were obvious mechanisms for scientific communication, and especially for responding to existing calls for faster, more systematic sharing of ongoing work. Particle physics communities were in an excellent position to respond to this opportunity; and indeed, the Web was designed with them in mind (Berners-Lee et al., 1994:76). In 1991, the theoretical physicist Paul Ginsparg brought his field's preprint culture online by building an email-based system to automatically receive, distribute and store full-text articles (Taubes 1993:1247). This system, called arXiv, is now the central tool of scientific communication for particle physics, as well as other physical and mathematical fields (Delfanti 2016). As well as fostering and speeding up previously established networks, arXiv "democratized" networks which previously favoured a more exclusive set of individuals and institutions (Taubes 1993:1247). However, these early online spaces preceded the rise of explicitly "open science", and were spaces of experimentation. Moore (2017:9) suggests that the free accessibility of arXiv was a by-product rather than an intention; contingent upon an early Web which was not a global resource but "something of a private playground for academics" (Ginsparg 2011:5). An analogy could perhaps be drawn with the "openness" of the early Royal Society; gentlemen conducting science in the public realm, where the public was differently and much more exclusively configured (see Section 3.4.7).

As physics communities evolved in close correspondence with the early Internet – a mutual shaping of technology and culture – other areas of science including biology seem to have been more distanced from such an interaction (Walsh and Bayma 1996:682). November (2012) described life sciences as "the least computerized area of science" in the 1960s, the technological and large-scale collaborative opportunities of WWII having bypassed biomedical fields (p. 6). This changed over subsequent decades, through a series of contingent developments (ibid.), and some biological fields – in particular those working with DNA sequences – developed close and mutually constitutive relationships with computing (Chow-White and Garcia-Sancho 2011). However, Strasser (2019) is careful to note that these formative relationships – with fields such as genomics and bioinformatics – cannot be generalised to all of biology (p. 7). Moreover, biology as it developed in the twentieth century has a more pluralistic, disunified character than physics or chemistry (Mittra 2016:6). Together these observations suggest that many scientists in biological fields, or other fields outside particle physics, would have been neither technologically nor culturally positioned to embrace the widely accessible, cheap, rapid scientific communication opportunities of the early Internet and Web: without either a cultural blueprint for systematic preprint exchange, nor socio-technological immersion in digital, networked spaces.

Following non-profit community ventures such as arXiv, big commercial publishers quickly adapted to the Web, continued to charge subscriptions, and adopted new revenue-generating strategies such as bundled subscriptions – “Big Deals” (Fyfe et al. 2017:15). This meant that scientists without established online communities and communications in the late twentieth century were likely positioned more as consumers than designers of Web-based scientific communication – a tendency reinforced by editorial leadership, commercial forces and metrics that maintained position of traditional journals as exclusive content vendors. Thus in many fields of biology, a desire for new sharing mechanisms that seems to have existed since at least WWII (see Section 4.5) may have remained in tension with inapposite cultural and technological circumstances even as the Web unfolded. Meanwhile, not-for-profit “DIY publishers” – working academics – established some of the first online open access journals in the late 1980s and 1990s: in contrast to arXiv, many of these were in the humanities and social sciences, consciously critiqued commercial publishing, and were interested in reaching new audiences (Moore 2019b). Moore argues that these journals form a “neglected prehistory of OA [open access]” – they came to be overshadowed by the rise to salience of open access movements in the sciences, particularly the biomedical sciences (see Section 5.2). In this way, the rise of “open” journal articles has come to be associated with biomedical fields despite – or in reaction to – a contingent historical path that steered many of these fields away from rapid, scholar-led, public, online publishing.

4.9 Model organism communities, and resource sharing in biology

Positive antecedents to online openness in the life sciences can instead be found in histories of epistemic communities formed by biologists, and their resource sharing cultures. Communities formed around a "model organism" approach have been particularly well-studied (Ankeny 2010). Model organisms are species studied, often in standardised form, both as "samples of nature" and as representatives of a broader range of organisms and biological principles (Leonelli 2007), and they are associated with distinctive epistemic goals and social conventions (Ankeny and Leonelli 2011). As a selective use of species, the approach is centuries old (ibid.), but particularly characterised experimental life sciences from the late nineteenth century (Strasser 2019:29) and gained popularity in the 1960s and 1970s with the rise of molecular biology (Ankeny 2010:93). Robert E. Kohler (1999) famously studied the community of fruit fly (*Drosophila*) geneticists established by Thomas Hunt Morgan in the US in the early twentieth century. The "fly group" had a distinctive, self-aware communality and egalitarianism: shared access to physical resources including stocks of mutant flies; collective ownership of ideas with credit to those who accomplished them; and distributed authority. They went on to shape a wider fly community through a "moral economy" in which mutant stocks were generously shared with outsiders, with an expectation of reciprocal sharing, disclosure of research intentions, and limited ownership of research problems: "no trade secrets, no monopolies, no poaching, no ambushes" (ibid., p. 254).

Kohler linked this culture and moral economy with the co-shaped features of the fly and the community: *Drosophila* can be produced rapidly and in abundance, and its use for genetic mapping threw up a "cornucopia" of research problems – enough for everyone. Additionally, Morgan came from a tradition of US experimental biology in which mutual aid was valued. Benefits came from apparent altruism: the fly group spread its reputation and way of doing science by making their fly the "standard fly". They also had an ongoing knowledge of others' research through moral economy exchanges. This communal culture is described as persistent through generations of fly researchers, in the way that a culture of preprint exchange has persisted in particle physics. It is an offline tradition of "openness" in science that is bounded and historically situated – traceable within the twentieth century – enabled by key social and material conditions.

The fly community was one of several prominent model organism communities that grew in the twentieth century: others assembled, for example, around the mouse (Rader 1998), the nematode worm *Caenorhabditis elegans* (Ankeny 2001; de Chadarevian 1998), and the flowering plant *Arabidopsis thaliana* (Leonelli 2007). Like the fly group, these communities have each tended to foster a group identity and an ethos of co-operation and resource sharing (Ankeny and Leonelli 2011:317–18), although each in a distinctive and evolving way, with features that mix openness and closure. For example, Rader (1998) described the founding "mouse people" as bound together through shared experiences of scarcity, in contrast to the abundance that enabled communalism for the "fly people". W. E. Castle, the mouse group leader of the early 1900s, managed scarcity through sequential and centralised graduate training – an "extended one-on-one apprenticeship period with the master professor himself"– allowing each student equal, uncompetitive access to resources (p. 340). This built a "group-oriented but individualistically organized" culture in which graduates were taught to be self-sufficient and flexible (pp. 342, 345). In contrast, the *Arabidopsis* community – which blossomed later, in the 1980s – began with a group of young researchers who were able to create a collaborative, co-ordinated community through centralised committee leadership that attracted generous funding (Leonelli 2007). This mirrors qualities of the fly community – but ironically, the *Arabidopsis* community emerged after success and growth had made fly research competitive and less hospitable – less open, perhaps – for newcomers (ibid., p. 201).

Openness as established in model organism communities is therefore conditional – and mixed with closure – in interesting ways. One of the typical pre-digital means of communication in these communities, beyond individual laboratories and institutes, was the newsletter: the Drosophila Information Service (DIS); the Mouse Newsletter; the Worm Breeder's Gazette; the Arabidopsis Newsletter; and many others (Kelty

2012:144). Through newsletters, organism-specific information was shared and standardised, manifesting and reinforcing communities. In a thoughtful analysis of the DIS in particular, Kelty (ibid.) identifies newsletters as a "meso-scale" of scientific practice between the lab notebook and "purified" scientific publications, where ongoing work could be communicated on the understanding that it was provisional, unpublished, and to be treated with sensitivity – something akin to a personal conversation or letter, where one might seek permission or acknowledgement to share further (pp. 141, 152). In this way the information shared was not universally "open", and nor was the community. Kelty describes it as "partially porous", as the newsletter was for "those actively engaged in *Drosophila* research" (pp. 152, 156-157). Outsiders could theoretically join – but were expected to contribute. This echoes the establishing context of Morgan's group (Kohler 1999), in which an apparently open exchange of materials and knowledge was predicated on reciprocity, trust, and disclosure of ongoing work. It also rings true with the observation that fly communities – at least in their later twentieth century form – could be unwelcoming to outsiders (Leonelli 2007:201). These qualities of community openness have particular resonance with my empirical results, notably those reported in Chapter 8.

The establishment of model organism communities in the twentieth century is an important prehistory to the emergence of twenty-first century "open science" on three grounds. Firstly it demonstrates that a cultural valuing of openness has been embedded in many biological research communities that have continuity to the present day – even if those cultures are "open" in qualified and contingent ways, and do not date back to the seventeenth century. Secondly, these communities have been important sites for digital and online transformation, especially in genetic fields: as data-gathering began to outgrow the newsletter format in the 1980s, and genetic sequencing projects of the 1990s and 2000s generated further data, model organism communities established online databases – including FlyBase, Mouse Genome Informatics, WormBase, and The Arabidopsis Information Resource (Leonelli and Ankeny 2012). Leonelli and Ankeny (ibid.) show that these databases have – through curation of standards of terminology and evidence, and co-ordination the sharing of physical materials – reinforced and continued to shape a collaborative ethos in these communities[15]. It is particularly notable that the public, online accessibility of these databases alters dynamics of community and identity: membership of the in-group is no longer required (ibid., p. 34). This changes what is meant by openness, as traditional foundations for sharing in these communities – personal interaction, trust, reciprocity, and conditional disclosure of ongoing work – are bypassed by the more unconditional qualities of online access. This is an important consideration in relation to biological scientists' encounters with online "open science" in my interview study.

Thirdly, model organism communities shape the well-known collision of public and private science that characterised the Human Genome Project (HGP) in the late twentieth and early twenty-first centuries. According to Ankeny and Leonelli (2011:317), many key figures in the public HGP effort came from model organism communities, likely influencing its ethos. In turn, the public HGP established immediate, online data sharing as an internationally agreed norm for large-scale sequencing projects – an important origin for more widespread open data movements in the 2000s (Maxson Jones et al. 2018; Strasser 2019; Wellcome Trust 2003), that was not yet framed as "open data".

4.10 Free software, open source software, and free culture movements

The final thread I will add to the backstory of twenty-first century "open science" emerges from software development cultures and communities in the 1980s. Prompted by analogous commercial, proprietary trends in computing to those in publishing (p. 93) and molecular biology (p. 95), programmer Richard Stallman founded the Free Software Foundation in 1983, which promoted the freedom to use, copy, modify and redistribute software (Free Software Foundation n.d.). "Free" referred to liberty, not price – as is often quoted, "'free' as in 'free speech' not as in 'free beer'" (Free Software Foundation n.d.). This became an important social, ethical

[15] With the possible exception of mouse communities, which have been more entangled with commercial and competitive pressures (Leonelli and Ankeny 2012:35).

and practical basis for grassroots software practice. In 1998, an offshoot called "open source software" emerged, which emphasised the pragmatism and cost-effectiveness of free software methodology for commercial purposes (Kelty 2008:99). Open source, which was "associated with the dotcom boom and the evangelism of the libertarian pro-business hacker Eric Raymond" (ibid.) was framed by Stallman (2016) as a corruption of "free" values. To Raymond, however, Stallman and the Free Software Foundation represented "a kind of dogmatic, impossible communism" (Kelty 2008:109). Kelty argues that this ideological division – which includes a determined denial of ideology – fuelled the birth in 1998 of free and/or open source software as a movement. But, he argues, it was not a typical social movement: "Free Software and Open Source share practices first, and ideologies second" (ibid., p. 113). This movement emerged out of geek or hacker culture, which was already ideologically diverse then (p. 93); it was a movement defined by "the practice, among geeks, of arguing about and discussing the structure and meaning of Free Software: what it consists of, what it is for, and whether or not it is a movement" (p. 98).

Associated with free software was Stallman's repurposing of copyright: he wrote the GNU Public Licence (GPL) which – instead of restricting re-use – requires that a program's source code be publicly accessible, redistributable, and modifiable (Eve 2014:17). Moreover, the GPL is "viral", requiring that modified versions are redistributed on the same "free" terms, an arrangement known as CopyLeft (ibid.). The licensing dimension of free and open source software then became associated with a broader "free culture" movement, in which the Harvard-based lawyer Lawrence Lessig was an important figure. He developed licences that could be applied to any type of content, and again, built upon copyright law to enshrine re-use rights (ibid., p. 20). This approach resulted in the now widely used Creative Commons (CC) licences, established in the early 2000s (Creative Commons Wiki 2011), of which there are several variants specifying different re-use conditions. Free and open source software – and subsequently free culture – were also entangled with another dimension of 1980s and 1990s computing, which was the idea of "open systems" (Kelty 2008:143–78). Open systems were seen as the solution to incompatibility between different computer components, including software: it meant achieving wide interoperability through standardisation, to provide "seamless integration" in marketing terms (ibid., p. 147). Kelty observes that in the 1980s this openness was a "kind of cultural imperative" tied to free market ideals, and upheld – at least rhetorically – in many industries (ibid., p. 148). However, it was difficult to either define or achieve; amongst competing companies it was not clear whose standards should be upheld. Kelty describes this openness as "unruly" and "hopelessly plural" (ibid., pp. 143, 145). The explicit licensing arrangements of free and open source software, and the free culture movement, emerged from the intellectual property "blind spot" of open systems (ibid., p. 178).

The connections between free, open source cultures and twenty-first century open science are overt, and multiple. The use of the word "open" may be, in part, a direct inheritance: David (2014) notes that in October 1999, a conference at the Brookhaven National Laboratory was titled "Open Source/Open Science" – the first use of the term "Open Science" at a prominent, public scientific occasion, according to David's knowledge and that of conference organisers (p. 34). Wikipedia places the origin of the term as at a similar time, and in a computing context. Its sources link to claims by the inventor and professor of computer engineering Steve Mann that he coined "open science", and registered the domain openscience.com, in 1998 or 1999 (Mann 2016; Mann et al. 2016; Wikipedia contributors n.d.). Although these instances suggest that analogies were being drawn between the "open" of software and science from the late 1990s, open science had not yet gained its current status as a broad "umbrella term" (see Section 5.4)[16].

Regardless of whether "open" was inherited directly from open source to open science – or whether a broader cultural emphasis on openness from the 1980s (Kelty 2008:148) shaped today's terminology – several authors have argued that that cultures and ideologies from open source run through today's "open" movements in science and research. In the relatively early days of global open access movements, Willinsky

[16] Maxson Jones et al. (2018) assert that "open science" has been a buzzword for scientists and policy analysists "since at least the start of the HGP" (pp. 772-773), i.e. since the 1990s, but I could not trace this claim in their references.

(2005) drew attention to a "convergence" between open source, open access, and open science, identifying this as a basis for a popular movement with a common cause. In doing so, he linked the meaning of "open" in software movements not only with open access to research and scholarship, but with visions of science as historically "open"[17]. Eve (2014), following Willinsky's observations, proposed that open access emerged from the convergence of two histories: those of software-associated free culture, and the economics of journal publishing.

Tkacz (2012) takes a broader view of contemporary openness, again identifying its roots in software culture. More specifically, he argues that a neoliberal of version openness – representing open source over free software values, and traceable back to Karl Popper's *The Open Society and Its Enemies* (1945) – prevails in contemporary movements. Examining open access, Moore (2017) concurs with Tkacz that open source software and free culture – with neoliberal overtones – have shaped a distinct lineage of open access which, particularly in the sciences, has adopted the Creative Commons Attribution Licence (CC BY) as a minimum acceptable standard of openness. However, Moore also argues that another distinct lineage of open access is concerned more with public access to scholarship than with licensing for re-use: a basis outside neoliberalism. These two lineages are similar to those identified by Eve (2014), but Moore emphasises an ongoing ideological struggle between different lineages of contemporary openness, rather than a convergence. Delfanti (2013) also locates a strong influences from open source software cultures in open science, specifically in contemporary biology. He argues that there has been a "remix" between a hacker ethic and traditional Mertonian values, suggesting that contemporary open science inherits a highly diverse[18] set of ideologies from hacker culture, as evoked in his opening sentence: "crack the code, share your data, have fun, save the world, be independent, become famous and make a lot of money" (p. 1).

[17] Willinsky is among those who draw on the work of David (2008) in constructing the past of contemporary open movements, as discussed in Chapter 3.
[18] A diversity also noted by Kelty (2008).

It is clear therefore that software cultures, especially those taking shape shortly before the turn of the twenty-first century, have had a strong proximal – but not total or determining – influence on the shape of today's open science. That influence is traceable through "open" language, ideologies, and specific practices such as the application of Creative Commons licences as a standard of openness. It may even have led to the conception of contemporary open science as a "movement" (or set of movements), analogous to the open source movement. These influences continue to become clear in Chapter 5 which unfolds twenty-first century open science itself, as well as in my interview findings. But as I have shown in this chapter, software culture is one of many antecedents that built up a context for open science: I will now review them and bring them together.

4.11 Conclusion: converging on twenty-first century "open science"

These multiple, accumulating histories present a nuanced backstory to the emergence of "open science" in the twenty-first century. Some depict encroaching closure or secrecy, suggesting a relationship between openness and science that is under threat or being corrupted. Others depict new or reinforced cultures of scientific openness. Since neither "science" or "open" are fixed concepts, it might be more accurate to say that the twentieth century profoundly altered – on multiple dimensions – what it means to do science, and consequently what it means, and what it looks like, for science to be opened or closed. At the start of the century, science was already altered from its gentlemanly, seventeenth century form. Significantly, a new boundary was being drawn between science as a profession and the public realm (Pandora and Rader 2008:355). Then, "big science" changed the organisational structures of some fields, notably in physics. There was more reason to communicate science in a rapid, ongoing way some of these larger community structures. WWII was central to these changes, funding big science (especially physics) and giving cause for centralised, rapid communication – but also certain forms of secrecy. All these trends altered the relationship of science to its traditional process for openness – journal publication. In high energy physics a paper-based preprint culture was born. In other fields, including biology, preprint cultures were suppressed by editors who began to see scientific papers as unique products.

The mass commodification of journal publishing after WWII opened it up to a larger market, but reduced its accessibility in other ways as prices increased, and scientific communities lost financial control over their means of "openness". Commercial publishers co-opted journal metrics – developed to grapple with literature growth – and the community tradition of peer review, entrenching and narrowing what it meant to publish well. Similar measures were adopted by universities for assessing academic merit. By this point, seeds had been sown for both the "serials crisis" – and thus eventually for open access movements – and academic cultures that prioritise a single means to "openness": publishing, exclusively and prestigiously. In biology journal publishing largely prevailed over a decades-long desire for alternative means to "openness" – like preprints – due to these social and economic forces, rather than a deficit of technological options. By the 1970s and 1980s, deregulated markets and private investment were encouraging a faster, but less fully shared science, prompting some to celebrate the dividends of an open market while others saw greater openness in publicly owned knowledge, and mourned local losses of Mertonian communalism. These changes may have primed both the acceptability of open science in a neoliberal political environment, and movements of resistance against commercial closures.

Through all this, the Internet grew, infused with a traditional narrative and a technical architecture of openness, and holding a potential for faster, deeper, wider sharing. But it could only change existing cultures so far: it helped physicists modernise and make public their preprints system – before it was "open science" – and helped publishers grow their profits, and further entrench prestige publishing. Meanwhile in some enclaves of biology, model organism communities had developed their own cultures of openness, based not on universal public access, but community-based reciprocity and trust. Like physics preprints, these cultures served as blueprints for online resource sharing within particular scientific communities, and also influenced the adoption of online data sharing as a norm in genomics via the Human Genome Project.

A final, vital influence at the end of the twentieth century was from resistance to commodification in a different domain – free software. Prioritising liberty to access, modify, and redistribute software – and creating the legal means to enact this through licensing – free software gave rise to open source, which valued the pragmatism of these free practices rather than their social, ethical foundations. The resulting movement in hacker and geek culture was fuelled by diverse, conflicting ideologies, and brought new and specific meaning to "open", which was already culturally valued in the neoliberal 1980s (Kelty 2008:148). There are indications that twenty-first century "open science" directly inherited its buzzword from this software movement. It certainly inherited other features, including a prioritisation of open licensing for reuse, and a conflicted ideological breadth that places free market efficiency under the same banner as a pursuit of social equality. This connection also potentially establishes software as a metaphor for science, suggesting that science might be improved if its source code – all its unpublished inner workings – could be exactly recorded and made visible, repeatable, and shareable to others. This is the "extreme" vision for open science (Nielsen 2009:32), similar to open notebook science (Bradley et al. 2008) that goes beyond openness of papers or data. It chimes with the value placed by scientists on objectivity and reproducibility, and is consistent with the suggestion by Delfanti (2013) and David (2014) that "open science" results from a remix between hacker culture and traditional scientific ideals.

Together, these histories challenge tendencies towards technological determinism, in which the Internet is framed as the driving force behind an inevitable emergence of open science. Related tendencies frame openness in science primarily as a technological problem or accomplishment that can be solved or achieved with the correct technological state, practice, or infrastructure. This is perhaps not surprising considering that software and Internet cultures were vital influences precipitating the emergence of open science. Admittedly, technological framings are often implied, or used for rhetorical effect: few if any commentators claim that open science lacks a complex, socially entangled past – or present. Framings matter though (see Haraway quote, Section 3.4): for instance, Nielsen (2009, 2011) and Bartling and Friesike (2014b) anchor their expectations of an open science "revolution" in the technological shift from paper-based journals to online, networked communication (see **Figure 5**). At the same time, a common emphasis – including by these authors – is the difficulty of bringing about cultural change; an acknowledgement of complexity. By presenting multiple histories of open science I bring together and thereby broaden the resources and vocabularies available for explaining its emergence – and anticipating its future – beyond a technological framing, or indeed any single explanatory framing.

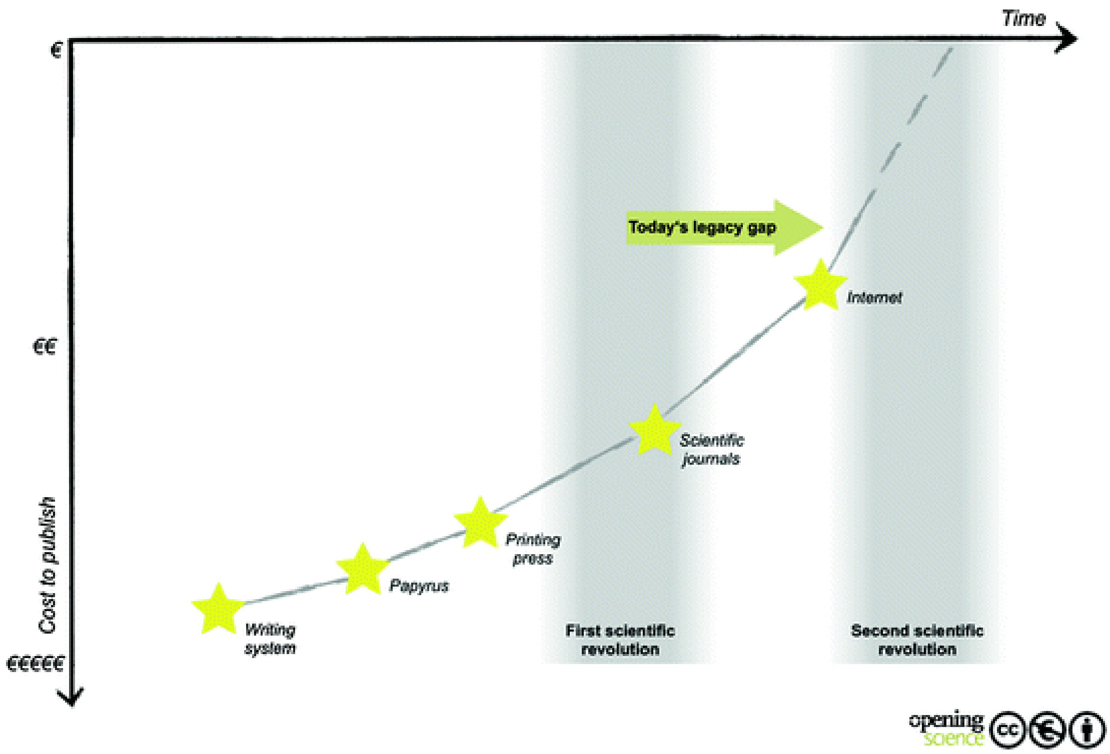


Figure 5 | An illustration reproduced from Bartling and Friesike (2014b:8), in which they depict an historical succession of scientific communication technologies, highlighting the expected "second scientific revolution" in association with the Internet. Technological determinism is evident in this framing: it suggests that technologies direct subsequent cultural changes – and that at present, we are in a "legacy gap" between the availability of a new technology and its cultural integration. Figure licenced under a Creative Commons Attribution Noncommercial Licence.

Together, these histories also show why the contemporary "open science" phenomenon is so diverse, connoting multiple meanings and values: it clearly has numerous historical roots. Today's "open" movements – open access, open research data, open preprints, open notebook science, and so on – each originate in one or several of these individual histories, as well as in overarching trends of growth, commodification, resistance to commodification, and socio-technological change.

What is remarkable is that in the 2000s, these lineages seem to have converged under the heading "open". It is in this context that I ask my second research sub-question: *How has "open" in science recently gained such salience and discursive power, despite its multiple meanings and lack of specificity?* The histories that I have collected begin to provide an answer by illuminating a slow and layered build up to today's movements that otherwise seem to appear suddenly. This means a slow ferment of ideas, causes, and cultures which have traction prior to their incorporation in an "open" movement – including a long-held desire for preprint systems or other alternatives to journal publication; dissatisfaction with rising subscription prices; pressure to publish with prestige; disintegrating community and ethos in some settings, while elsewhere communality is valorised; and excitement about the potential of the Internet, and open software, as a way forward. Each one of these ideas and causes brings its own constituency of actors to "open", multiplying salience and social power. The attraction of technological openness may be particularly high where tension has been building for decades: for biologists embedded in cultures of prestige publishing for example, without recourse to preprints, and outside of communally-minded enclaves like model organism communities.

How it is that "open" brought these constituencies together remains curious. I propose that the flexibility and ambiguity of "open" is its strength in this regard. Many have commented on this quality, and in Chapter 3 I showed how interpretative flexibility eases the reframing of past science as "open". Fecher and Friesike (2014), for example, write that "openness could refer to pretty much anything" (p. 18). In the context of computing, Kelty (2008) comments: "Openness is an unruly concept [...] Everyone claims to be open, everyone has something to share, everyone agrees that being open is the obvious thing to do..." (p. 143). My suggestion is that the broad appeal of "open" has allowed it to become a rallying point for meaning. It can refer to the openness of a free market, which upholds efficiency and progress as markers of scientific success; the openness of democracy and equality that shares across social boundaries; the openness of information unencumbered by copyright and conditions; the openness of transparency and accountability; the openness of being free to act; the openness of publishing one's work; and the open exchanges of a trusting community. Of course, there is nothing inevitable about this acquisition of meaning: the potential of “open” seems to have been seeded in the 1980s (see Sections 3.4.6 and 4.10) when it became a popular term in the context of neoliberalism; its increase in cachet may have been gradual with its influential deployment by open source, and then open access movements, prior to the “open science” umbrella term. Moore (2017) conceives of open access in particular as a “boundary object” – “a concept that has a specific understanding in a local community of practice but is rigid enough to maintain its definition across communities” (ibid., p. 2; concept proposed by Star and Griesemer 1989), a theorisation that could be useful for open movements broadly. Crucially, such objects are maintained by their plasticity, and cease to be unifying when made specific – for example, in policy (ibid., p. 12). This suggests that ideological conflicts within “open” do not diminish its salience and power, and in fact do the opposite. Kelty (2008) suggests that open source became a “movement” *through* internal disputes: everyone had something to believe in (p. 113). I propose that precisely because it does not commit to any single agenda, but holds positive meaning for many, “open” continues to collect causes, holding them together loosely. As a result, the term now captures a great deal of all that is considered good – and wanting – in contemporary science.

Chapter 5 | “An old tradition and a new technology”
Advocacy and policy framings of open science

5.1 Introduction

Chapters 3 and 4 showed that openness in science is an idea with great depth and breadth: it has resonance through hundreds of years of scientific history, and it has also absorbed a wide array of meanings over the twentieth century that continue to multiply. A more recent period – coinciding approximately with the twenty-first century – is characterised by the rapid rise in salience of explicitly “open” science instantiated by growing advocacy and policy movements. These have tended to emerge in grassroots communities of researchers and entrepreneurs before uptake and wider establishment by institutions that govern research. The aim of this chapter is to move from a wide-angle perspective on the historical place and meaning of openness in science towards a narrower focus on the contemporary period and the way that pertinent aspects of today’s open science movements have unfolded.

In particular, this chapter will unpack the recent emergence of *open access* and *open data* as two of the most significant advocacy and policy movements defining contemporary open science, before turning to *open science* itself as an overarching advocacy and policy movement. This will involve a more detailed characterisation of the conceptual and practical meaning of open access, open data and open science as constructed in advocacy and policy contexts. This is an empirical engagement with the contemporary dimension of my overarching research question: *How is the meaning of “open” (or “openness”) being constructed in the context of science?* The advocacy and policy context is also an immediate influence upon the construction of openness by scientists. The current chapter thus contextualises the interview-based empirical engagements presented in subsequent chapters. I have chosen to focus upon open access and open data not only because of their formative influence on open science more broadly, but because they emerge as two of the three major categories of meaning discussed by interviewees, and analysed in Chapter 6 (open access) and Chapter 7 (data openness). The third category – interpersonal openness (Chapter 8) – is significant in its relative absence from the advocacy and policy framings of open science presented in this chapter.

This topic presents the challenge of telling recent histories that are still in the making. I do not have the benefit of significant hindsight or an established scholarly literature with which to make sense of a multiplicity of events, actors and connections with varying levels of influence. To be comprehensive would also be beyond the scope of this chapter. I aim instead to provide enough breadth, and selective depth, for the reader to appreciate key events, trends and themes in the unfolding of open science in the UK and Australia over the period 2000-2019.

Document case studies

To allow for deeper engagement with my research questions alongside broader commentary, I include nine short **document case studies** in which I analyse the construction of "open" in key contexts. In each example I consider the context, form, and role of the text, and how it relates to the themes of my four research sub-questions: (a) historical framings; (b) diversity of "open" meanings; (c) epistemic virtue, and (d) portrayal of researchers and research culture (full questions in Section 1.7). The features of these documents are summarised in **Table 4**, and aspects of my analysis are summarised in **Table 5**. The documents were primarily produced in the UK, in or in an international setting that has affected global open science discourse.

Table 4 | Key details summary of the nine document case studies: authorship/source, type of openness, citation, and length.

	Document	Authorship/Source	Authorship category	Open category	Citation (Year)	Length
1	PLoS open letter	Biomedical scientists Harold Varmus, Patrick Brown, and Michael Eisen (before they founded PLoS as a publisher)	Grassroots (influential scientists)	open access	PLOS (2000)	<1 page
2	Budapest Open Access Initiative (BOAI)	16 "leading proponents of open access" at a meeting called by the Open Society Institution (OSI)	Grassroots (international meeting)	open access	BOAI (2002)	~2
3	Finch report	Working group reporting to UK government, chaired by Dame Janet Finch	Policy (for UK government)	open access	Finch (2012)	11
4	Open definition 1.0, 2.1 & the Panton Principles	Open Knowledge Foundation (OKF) & associates: Peter Murray-Rust, Cameron Neylon, Rufus Pollock, John Wilbanks	Grassroots (advocacy organisation)	open knowledge & data	OKF (2006, 2015) & Murray-Rust et al. (2010)	~2-3 each
5	*Science as an open enterprise*	The Royal Society, working group chaired by Prof Geoffrey Boulton	Policy (learned society)	open data	The Royal Society (2012)	105
6	*Concordat on Open Research Data*	HEFCE, RCUK, UUK, Wellcome Trust	Policy (UK funders, institutions)	open data	HEFCE et al. (2016)	24
7	*Towards Another Scientific Revolution*	Sönke Bartling and Sascha Friesike: medical imaging scientist and digital innovation researcher	Grassroots (academic literature)	open science	Bartling and Friesike (2014b)	13
8	*Open Science and its role in universities: a roadmap for cultural change*	League of European Research Universities (LERU), lead author Dr Paul Ayris	Policy (universities, UK & Europe)	open science	LERU (2018)	32
9	*Foundations for Open Scholarship Strategy Development*	Collaboratively drafted by 32 authors	Grassroots (community organising)	open scholarship	Tennant et al. (2019)	39

Table 5 | Summary of analytical points for the nine documents, including engagement with the themes of my research sub-questions. Column (b) summarises values of "open" that are salient in the document: Ep=Epistemic benefits, Pr=Productivity/efficiency, Co=Commercial/economic interests, TL=Tech libre values, In= Inclusivity/community/interdisciplinarity, Ac=Accountability/transparency/integrity, Pu=Public access & trust.

	Document	Role/Purpose	(a) Historical	(b) "Open" values	(c) Epistemic virtue?	(d) Scientist role
1	PLoS open letter	Asks publishers to release biomedical articles; threatens boycott; signed by >34000 scientists; led to founding of PLOS		Ep, Pr	Epistemic & moral (separate)	Active agents
2	Budapest Open Access Initiative (BOAI)	Established open access definition together with Bethesda Statement and Berlin Declaration (BBB definition)	Tradition of non-profit publishing renders OA possible	Ep, Pr, Co, TL, In	Epistemic & moral (blended)	Active agents
3	Finch report	Successfully recommended open access policy to UK government with a "gold" (publisher) emphasis	Publication at heart of science & research for >350 years	Ep, Co, Ac, Pu	Epistemic & moral (blended)	Low-priority stakeholders, need incentives
4	Open definition 1.0, 2.1 & the Panton Principles	Established widely used open definition especially for data (Panton Principles), strong links to open software movements		Co, TL, Pu	Essential for good science	N/A - "open" centred on knowledge not people; no ownership of data
5	*Science as an open enterprise*	Endorsement of open data policies by establishment, recommendations to many groups, precursor to FAIR data concept	Openness at the heart of scientific practice for centuries	Ep, Co, Ac, Pu	Essential for good science	Medium-priority stakeholders, need incentives
6	*Concordat on Open Research Data*	Agreement on open research data principles by main funder UK research funders		Ep, Co, Pu	Leads to higher quality research	High-priority stakeholders, diverse
7	*Towards Another Scientific Revolution*	Academic advocacy of open science, frames it as tech-led "revolution", continuation of historical legacy	Science open since the 17th century	Ep, Pr, Pu	Leads to higher quality research, and better knowledge	Key agents, need for engagement
8	*Open Science and its role in universities: a roadmap for cultural change*	Conceptualises open science as a whole, which requires cultural change to adopt, presents a plan for universities to do this	Cultural change analogy with 16th century moveable type	Pr, In, Ac, Pu	Leads to higher quality research	High-priority stakeholders, need for incentives
9	*Foundations for Open Scholarship Strategy Development*	Presents early-, mid-, long-term strategy for a unified open scholarship mission, at individual to national levels	"Proud history and tradition" of openness in science	Ep, Pr, In, Ac	"Open" is a mindset & identity anchoring a knowledge movement	Reluctant agents, need for incentives

5.2 Open access to research articles

Open access is possibly the earliest and most prominent element of open science discourse and practice, made visible by more than 20 years of grassroots and entrepreneurial activity followed by increasing top-down endorsement by research funders, institutions and publishers. Within these discourses, open access most often refers to free, public, online access to scholarly literature – especially in the form of peer-reviewed journal articles, but also applying to online books and other content – that would otherwise be accessible only via costly institutional subscriptions or individual payments (Eve 2014; Suber 2012). Many advocates of open access including Eve (ibid.) and Suber (ibid.) emphasise reusability as well as access, so are interested in removing legal and technical as well as financial barriers to reuse. For example, the Creative Commons Attribution licence (CC BY) – often upheld as a copyright standard for open access – requires that original authors are cited, but allows anyone to “distribute, remix, tweak, and build upon [the] work” without seeking permission. This approach is an influence of open source software and free culture movements upon open access, as discussed in Chapter 4 (Section 4.10).

In the early 2000s, consensus and salience were built around this latter, technically and legally specified use of “open access” through a series of international discussions and declarations: the Budapest Open Access Initiative (BOAI 2002), the Bethesda Statement on Open Access Publishing (2003), and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities (2003). **The BOAI is document #2, analysed in Section 5.2.4**. As the last title indicates, there was participation in these discussions from a range of disciplines, not all scientific. However, Moore (2019b) argues that this was a key moment at which a scientific framing of open access – represented both in the wording of the BOAI and its signatories – became dominant, establishing liberal licensing as a tenet of open access (pp. 2, 8). Prior to this point, at which the term “open access” was crystallised, another lineage of grassroots movements – primarily in the social science and humanities – experimented with free online access to scholarship without the *libre* (licencing for reuse) imperative (ibid.). Like open science as a whole therefore, open access is characterised by plurality, but a variety that focuses on “open” reuse is more influential, at least in the global north.

5.2.1 Green and gold open access

Practices and infrastructures of open access are usually divided into two main categories according to how the material is made accessible: repository-mediated open access, sometimes referred to as “green” open access; and journal-mediated, or “gold”, open access (Eve 2014:8–12; Suber 2012:6)[19]. Both of these routes to access are proposed by the BOAI, but each has developed a complex history in the politics of open access movements, representing a different set of values and compromises. Each also involves a different material practice and set of actors.

Green open access involves the deposition of an article into a publicly accessible, online repository that is hosted by a research institution, or may be curated in a particular subject or disciplinary area. If the article has also been accepted or published by a journal, the legal capacity of researchers to distribute it via the green route depends on the rules of the publisher in question: traditional publishers tend to assert their interests by restricting distribution of the final, typeset “publisher version” of the article, which they sell via subscriptions or individual purchases. However, many allow an “author version” of the article, which is peer-reviewed and finalised in its content, but not typeset and branded, to be distributed via repositories, sometimes after an embargo period. As a result, this mode of open access is free of financial cost to practise as well as receive, but requires knowledge of repositories and publisher rules, a willingness to distribute a pre-typeset version of their article, and a possible embargo delay. Online tools exist that allow researchers to navigate individual journals’ policies, in particular SHERPA/RoMEO hosted by the University of Nottingham (Eve 2014:10). The experience of accessing green open access articles is

[19] Other colour-coded categories of open access have also been suggested including bronze (Piwowar et al. 2018) and black (Björk 2017a), but these have not gained the salience of the green and gold categories.

shaped by this infrastructure as they do not appear on official publisher websites unless they are also open access via another route. Their discoverability depends on the repository's indexing practices, and potentially on efforts by the researcher and their institution to promote and maintain publication records. Many repositories also provide the function of digital preservation, which protects access to the article in the long term, independent of publisher websites (ibid., p. 11).

Gold open access, in contrast, is a model of journal publishing in which the final, typeset "publisher version" of the article is free and accessible online. Contention surrounds the financing of this model. Some open access journals are supported by the publisher's internal financial reserves, a sponsor, or another revenue source. Many other journals, however, support open access by requiring article processing charges (APCs) from authors, who are expected to pay from their own resources, grants, or institutional funds (Kozak and Hartley 2013). Some publishers waive these charges in cases of economic disadvantage (e.g. PLOS n.d.). Many open access journals – the majority, as measured by the Directory of Open Access Journals (DOAJ n.d.)[20] – do not charge APCs. Fees for open access are, however, particularly prevalent in medicine and the sciences. They are favoured by "private and large publishing houses" (Kozak and Hartley 2013:2591), and are more common for journals certified by DOAJ for meeting best practice openness standards. Additionally, in the global north and the life sciences particularly, the founding of open access publishers as start-up companies[21], funded primarily or entirely by APCs, has been an influential part of the open access landscape. Gold open access articles are available free in their final typeset and branded form from publishers' websites as well as any indexing services linked with the publisher and the author's institution. They tend to be "open" in a reuse sense, as the publisher controls licensing – although the permissiveness of the licence offered may vary – in contrast with green open access articles, which tend to be "open" in cost terms only (Björk et al. 2013).

[20] On June 1 2020, the DOAJ recorded 4,014 journals that charge APCs and 10,624 journals that do not. However, of those certified by the DOAJ as following best practice and a high level of openness, most (1,052) charge APCs and a minority (331) do not.
[21] Examples of open access publishers as start-up companies include: BioMed Central (BMC), PLOS, PeerJ, F1000 and eLife.

A distinction is made between fully open access journals in which every article is openly available, and "hybrid" journals that operate via a traditional subscription model, with the option of paid gold open access on an individual article basis. The DOAJ as an advocate voice does not include hybrid journals in their certification of open access journals, but at the time of my research hybrid was widely accepted as a form of open access in funder policies. The number of journals offering a hybrid option and its uptake by researchers have grown fast: Björk (2017b) records a five-fold increase in hybrid journals and articles from 2009 to 2016. The Wellcome Trust (2020) also records higher spending on hybrid than fully open access journals from 2016/17 to 2018/19 amongst the researchers they fund – both because of a higher number of hybrid articles, and a higher average APC cost at hybrid journals. It may be that some hybrid journals are able to charge higher APCs in correspondence with their higher prestige; there has also been controversy over apparent "double-dipping" as hybrid incomes combine subscriptions and APCs, the economics of which are opaque (Mittermaier 2015; Pinfield, Salter, and Bath 2015).

APC-based gold open access publishing has also become associated with the phenomenon of "predatory publishing" (Beall 2012), which co-opts the norm of charging fees for online publication in order to yield profits with little or no attention to peer review and other functions traditionally administered by journals, such as editorial checking and indexing. It has subsequently been observed that the boundary between predatory journals and those viewed as legitimate is not clear-cut, and that misleading, indiscriminate, and low quality journal practices abound without necessarily being linked to open access (Cobey et al. 2018). However, the association arises persistently and was raised once again in my interviews.

5.2.2 The rising salience of open access

The recent history of open access movements in the global north – particularly in relation to the life sciences – is marked by several developments in addition to formal definition-making. In 2000, PubMed Central (PMC) – “a free archive of biomedical and life sciences journal literature” (NCBI n.d.) – was established by the US National Institutes of Health. In the same year, more than 30,000 scientists signed an open letter (**document #1**, **analysed in Section 5.2.3**) asking that publishers make research articles freely available through PMC or a similar resource, or face withdrawal of research community support (PLOS 2000). The founding of two internationally influential open access publishers, BioMed Central (BMC) and PLOS, was closely related to these developments: entrepreneur Vitek Tracz founded BMC in 1998 and had also been in contact with then NIH Director Harold Varmus about the creation of PMC (Poynder 2006); and Varmus together with fellow life scientists Patrick Brown and Michael Eisen were the instigators of the open letter. The latter group founded PLOS in 2001 as a direct follow up to their letter, having encountered resistance from traditional publishers and reasoning that direct entry into the publishing industry may be a more effective way to bring about change (PLOS n.d.). The creation of these new scholarly publishing avenues, resources and brands among others, together with a political consciousness fostered by acts such as the open letter, widened visibility for open access initiatives beyond grassroots and entrepreneurial advocacy communities into more mainstream research communities, especially in the life sciences.

5.2.3 Document #1: the PLoS open letter (2000)

- The Public Library of Science (now PLOS) began in 2000 as an advocacy initiative, with the circulation of a short open letter.
- Aimed at traditional publishers of medical and life sciences, instigated by three influential US life scientists.
- High visibility and traction, including in mass media and prestigious journals (Hagerlid 2012:1); "signed by nearly 34,000 scientists from 180 nations".
- Shows discourse catalysing "open" movements before they took form.
- The three-paragraph text:
    - Envisions "online public library" for published biomedical literature;
    - Asserts a right to public ownership and accessibility to this library;
    - Threatens to withdraw community support from publishers who do not allow redistribution of work for this purpose.

  > *"Establishment of this public library would vastly increase the accessibility and utility of the scientific literature, enhance scientific productivity, and catalyze integration of the disparate communities of knowledge and ideas in biomedical sciences."*

- Epistemic and instrumental goals are prioritised: making science more productive and integrated; chiming with both communitarian ideals and neoliberal economics. These meanings appear in subsequent "open" discourse.
- *Libre* licensing and transparency do not feature.
- "Public" frequently invoked but value of access beyond scientists not articulated: scientists positioned as members of the public with access rights.
- Establishes negotiating position, accepting compromise with commercial publishers: six months of exclusivity, before public access.
- No historical framing, which appears more in later open discourse.
- Scientists as instigators of action, not barriers to progress, characterising this grassroots phase. Instigators are high-profile, not disempowered, however.
- Access is framed as epistemically useful and morally correct, but for different reasons: not an epistemic virtue framing, but a foundation for it.

5.2.4 Document #2: the Budapest Open Access Initiative (2002)

- First of three international statements in 2002-2003 that catalysed open access movements and provided reference definitions.
- BOAI meeting was called by George Soros' Open Society Institute (OSI).
- Brought together "leading proponents" of existing open access initiatives.
- The 16 original authors were primarily early innovators in online scholarly communication – among them representatives of the recently founded open access publishers PLoS and BioMed Central (BMC).
- Concise document, just over 1000 words.

The text:

- Begins with a striking line: "An old tradition and a new technology have converged to make possible an unprecedented public good".
- Establishes historical resonance, but only with the specific tradition of non-profit publishing before mass commercialisation (Fyfe et al. 2017).
- "Open access" used frequently, but "open" is not detached as a signifier.
- Diverse values and goals brought together: efficiency and productivity in research – the faster the better; social value of education; equality and the upholding of different kinds of knowledge as valuable; epistemic values as uniting, human values; cost-efficiency; individual visibility and impact.

> *"Removing access barriers to this literature will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge."*

- Reflects meanings of "open" from software cultures and the early Internet: *libre* licensing is not specified, but liberal digital re-use rights are mentioned.
- Technology-centric: language of "users", emphasis on digital processing.
- Publisher control is not mentioned: open access as a technological opportunity, not an act of resistance. "Access barriers" are naturalised; they exist "for various [unarticulated] reasons". Presents open access as non-threatening to commercial actors and achievable through market forces (Moore 2019b).
- Written in third person, but scientists are portrayed as active, able to adopt open access without systemic support – characterising grassroots phase.
- Moral and epistemic purpose are brought together the single heading "open access", but not blended: no suggestion that open science is better science.

5.2.5 Open access in research policy: the UK and Australia

Since the early 2000s the visibility of open access movements has extended further, in particular through the influence of major research funders, governments, and institutions who have made statements or policies encouraging or requiring that the research they support is open access. This has been an international phenomenon, but with different geopolitical foci promoting different approaches. My research is situated primarily in the UK and Australia, both in the global north, and with parallel, yet contrasting, histories of open access governance. The UK context is characterised by a relatively early and increasingly stringent pursuit of open access in research policy, led by charity and governmental research funders. It has been unusual in prioritising and funding a gold, APC-based – and therefore, publisher-centred – model of open access. In Australia, timings of major policy developments have tended to align with those in the UK, but the policies themselves have been less stringent, with fewer consequences for researchers who do not comply. There is also a preference for green, repository-mediated open access, with little funding available for APCs. Both of these contexts vary significantly in comparison with other regions: Latin America for instance is notable for an extensive "non-commercial and publicly-funded" open access infrastructure (Debat and Babini 2019:3, preprint) that was established early, from 1997, with the journal aggregator *SciELO* (Bulock 2019).

Open access in the UK

Of particular note for the UK context are the statements and policies of the Wellcome Trust and Research Councils UK (RCUK[22]), two of the most significant funders of publicly-oriented UK research. The influence of the Wellcome Trust here is notable, as a wealthy and influential biomedical charity which has proactively led the open access agenda – and subsequently other top-down "open" agendas – in the UK. The Wellcome Trust commissioned a report which showed, in 2004, that open access

[22] Research Councils UK (RCUK) was replaced by UK Research and Innovation (UKRI) during the course of this study, in 2018.

scientific publishing, supported by author fees, is economically viable and a preferable system to subscription publishing (Wellcome Trust 2004). Both Wellcome and RCUK announced and began implementing open access mandates for the researchers they fund in 2005-2006 (Suber 2006; Wellcome Trust 2008). The next major developments took place from 2012, when a government-commissioned working group chaired by Dame Janet Finch and composed of “individuals drawn from academia, research funders and publishing” (p. 2) released its report, *Accessibility, sustainability, excellence: how to expand access to research publications* (Finch 2012). This is **document #3, analysed in Section 5.2.6.** It recommended that “the UK should embrace the transition to open access” (p. 7), with an emphasis on gold open access paid for by APCs. It recommended that funds be put in place to allow researchers to cover these costs, an arrangement already pioneered by Wellcome. In the same month in 2012, Wellcome “strengthened” its open access policy to include sanctions for researchers who do not comply (Wellcome Trust 2012). Not long afterwards, in 2013 – and following the Finch report’s “broad acceptance” by the Conservative and Liberal Democrat government of the day (UK Government 2012) – RCUK introduced a more unified open access policy across its research councils, accompanied by a provision of block grants to universities to fund APCs (RCUK n.d.). These developments established a pattern in the UK of progressively more stringent open access policies that increasingly apply to a larger population of researchers.

5.2.6 Document #3: the Finch report (2012)

- Report recommending open access to the UK government, led by Dame Janet Finch, composed by a working group representing academics, research funders, librarians, learned societies, and publishers.
- Set the UK's policy direction towards journal-mediated "gold" open access including via hybrid journals, and led to the provision of public funds for APCs.
- Employs brief historical framing, establishing journal publication as a tradition for over 350 years. Does not re-frame past science as open, however.
- Open access portrayed as a single issue, not part of open science/research.
- "Open" terminology not emphasised: contrasts with the Royal Society report (2012) being produced concurrently. Perhaps a flattening of the disruptive connotations of "open" for a multi-stakeholder group.
- Emphasis on values with appeal for government: transparency; accountability; public engagement; stimulation of innovation with public sector benefits; economic growth; and increased returns on investment. These are placed alongside epistemic values: greater efficiency and innovation within research.
- Incorporates *libre* values in which open access also implies liberal re-use rights.
- Depicts tension, compromise and polarised interests: "Members of the group represented different constituencies who have legitimately different interests and different priorities" (p. 2). This is no exaggeration: the group includes corporate publishers whose businesses are disrupted by open access.
- Tensions manifest throughout: although "barriers to access…are increasingly unacceptable in an online world" (p. 5) there is a need to "sustain what is valuable in a complex ecology" (p. 3). Recommends "accelerating" open access, but warns that this must not be hasty. Ecological metaphors urge "balance" and "sustainability", arguably naturalising a corporate publishing hegemony.
- Critics of the report argued that its costly support of "gold" over repository-mediated "green" reflects publisher influence (e.g. Harnad 2012). Gold and hybrid APCs are an income for publishers whereas green challenges their role.
- Researchers positioned as one of many stakeholder groups involved in a systems-level transition. It is anticipated the researchers will lack awareness and require reasoned explanations and incentives to adopt open access.
- Slight, brief blending of the moral and epistemic purpose: "better, faster communication can bring better research" (p. 3).

In more recent years, two major open access policy developments have begun – or promise – to have an even wider-ranging impact on UK researchers, beyond individual grantee-funder relationships. The first of these is exceptional because it relates to the 2021 Research Excellence Framework (REF), an assessment exercise critical to the distribution of research funds among UK universities. It is a stage upon which the value of departments and universities are made measurable through reviews of research output. From 1 April 2016, research articles must have been made open shortly after editorial acceptance, by means of deposition in a "green" open access repository, in order to be eligible for REF2021 (HEFCE 2014). This policy has raised the stakes of open access acutely for research institutions, and in turn for researchers, during the period of my empirical data-gathering. Secondly, 2018 saw the launch of "cOAlition S", a partnership of national and charitable funding agencies centred in Europe and led by Science Europe, which proposes a major re-think to open access policy, termed "Plan S" (European Science Foundation n.d.). UK Research and Innovation (UKRI) – the new, transformed RCUK – and the Wellcome Trust are both members of this coalition, whose ten principles would add a new level of stringency to open access policy. Significantly, Plan S commits funder policies to aspects of the "open" definition established by the 2003 Berlin Declaration that have been treated loosely by existing policies. In particular, it requires that open access material is not just freely available online, but also that it has an open license allowing liberal reuse. Plan S thus tightens the criteria that journals and repositories must meet for their articles to be considered "open access" in policy terms. Notably, "hybrid" journals will be disqualified as a route to open access, having previously been major beneficiaries of research funder policies and associated block grant provisions in the UK. The empirical parts of my research were largely concluded before the announcement of Plan S, but it indicates a distinctive trajectory for open access policy in the UK that can be discussed in relation to my interview findings (see Section 6.8).

Open access in Australia

In Australia, open access has also grown in salience but to a lesser extent at present, and in qualitatively different ways. Australian universities were early to establish a green open access infrastructure: the Australian National University set up an "e-print repository" in 2001, and in 2003 the Queensland University of Technology was the first institution in the world to mandate open access for its research outputs (QUT n.d.; Steele 2013). Government funds were invested in institutional repositories, including by the Australian Scheme for Higher Education Repositories (ASHER) from 2007-2011. However, these funds were not necessarily used to make these repositories open access (Steele 2013) and no such support is ongoing (CAUL and AOASG 2019). The two major government-funded research councils began encouraging forms of open access from 2006-2008 (ARC n.d.; Suber 2009). Then in 2012 and 2013 respectively, the National Health and Medical Research Council (NHMRC) and the Australian Research Council (ARC) began to require open access (Steele 2013:284). Although this timing approximately aligned with major policy developments in the UK, the Australian policies were less stringent with regard to timing of, and exceptions to, open access. The policy preference in Australia is for "green" open access, and there is no central funding provision for the "gold" path that has been heavily supported in the UK. The policy momentum that is apparent in the UK, led by influential institutional actors, is not evident in Australia, where promises of an open access future are less apparent. At present, there are no sanctions for non-compliance with the major policies. Australia is on the periphery of the global north-centred Plan S initiative, with little indication that it will embrace landmark changes of that kind (Ross 2019). As a result, Australian researchers tend to encounter less demanding open access environments, unless their particular institution has developed strict policies, with fewer influences from the commercial publishing sector that has shaped the UK's "gold" policy emphasis. However, the global nature of open access discourse and scientific communities ensures that these national contexts are not independent.

5.3 Open research data

The use of "open" in relation to scientific research data has also became highly salient in research, policy and publishing discourses since the early 2000s. Like open access, and the open science label under which both have been incorporated, the open data approach is a plural phenomenon – indeed, it is arguably more diverse in its origins and practice than open access, and less conceptually unified. Most obviously in relation to the present study, it has foundations in scientific traditions of publishing evidence. According to the Royal Society (2012), until the mid-to-late twentieth century, it was a norm for published article formats to include all the data resulting from a study (p. 26). Since around this time, "the vast majority of scientific data [...] have only been accessed by small groups of experts" (Leonelli 2013:6). The Royal Society attributes this "data-gap" to new ways of collecting data in an increasingly digital, online age that produce huge quantities of data, going well beyond the capacity of a traditional article. Leonelli (2013) writes that the exclusivity of such data was normalised within a traditional view of scientific knowledge production as esoteric, technical and specialised, and thus the problem was not entirely a technical one. She argues that "the Open Data movement" arose in the twenty-first century as a challenge to this traditional view:

> *"The movement brings together scientists, policy makers, publishers, industry representatives, and members of civil society around the globe who believe that data produced by scientific research should be made publicly accessible online and freely usable to anyone." (p. 6)*

As well as being framed as a solution to a problem, an open data approach is seen as a technological opportunity to share, store and analyse data in new ways – an opening up of scientific evidence to a volume and depth that would have been impossible via traditional journal publication. As such it also represents not only a new kind of epistemic resource, but a new mode of transparency and accountability. This makes open data appealing to public institutions responsible for science, that are under pressure to foster public trust – especially in light of incidents that have challenged this trust in recent years, such as the University of East Anglia "ClimateGate" email scandal (Leonelli 2013:7). This is among the reasons highlighted by the Royal Society (2012) for its support of open data movements. Indeed, according to one of my interviewees who was closely involved with the Royal Society's 2012 report, ClimateGate was the catalyst for the Society becoming involved in open data. Leonelli (2013) also highlights how trends of these kinds – in scientific research and beyond – are associated with a re-conceptualisation of data:

> *"...data are increasingly conceptualized as inherently valuable products of scientific research, rather than as components of the research process that have no value in themselves." (p. 7)*

This means that research data are being treated in new ways. Open data movements encourage the publication of standalone "data papers" that might accrue prestige in the same way as research articles (Chavan and Penev 2011). And in a broader political economic context, this also means that data – newly mobile and valuable – can become commodities in a global market, with value to governments and industries in relation to challenges such as in human health (Leonelli 2013:9).

5.3.1 Forerunners to open data: databases and the Human Genome Project

Although the twenty-first century saw the birth of "open data", there were forerunners to this movement that also conceived of data as valuable in themselves – a communal resource for science beyond the research article. Strasser (2019) positions the first digital sequence databases in biology – EMBL library and GenBank – built from the 1980s, as formative examples of open science in action (pp. 227-254). Moreover, he argues that these databases follow a long historical tradition of collections in biology that that has been overshadowed by experimental traditions until recent decades, when collections have become more widely accessible online. The communal resources collected by model organism communities in the twentieth century, and their transformation into public, online databases, are notable in this context (see Section 4.9). Strasser argues that early sequence databases challenged the role of journal editors as the authorities on valid scientific data – as journals were no longer the only reputable scientific sources. An increased value of scientific data in digital form was thus already seeded in the 1980s.

Strasser's narrative about the development of nucleic acid sequence databases is instructive. It began with manual digitisation of sequences published in papers, proceeded to accepting electronic submissions, and eventually became a core infrastructure as journal policies enforced deposition of sequences associated with publications. He documents resistance to such policies by both journal editors and scientists that is highly reminiscent of pushback on open data policies today (see Section 7.6), and argues that it was an eventual top-down enforcement by journals that led to the success of early databases, and the acceptance of data sharing norms by scientists. It is notable that these norms were established specifically in genomic sciences, and in crystallography – similar culture change accompanied the establishment of the Protein Data Bank (PDB). Influences of disciplinary data sharing norms are evident in my interviews (see Section 7.4).

These developments all preceded the Human Genome Project (HGP), which spanned the turn of the twenty-first century (Cook-Deegan 1994; Maxson Jones et al. 2018). The public HGP drew global attention to the case for publicly available data in service of the public interest, through contrast with the proprietary approach to the HGP led by J. Craig Venter. The public effort established a norm in genomics of making large-scale digital datasets publicly available online for reuse by a wider community. Consensus around this practice was built through international agreements. The 1996 Bermuda Principles stated that large human sequence datasets must be "freely available and in the public domain in order to encourage research and development and to maximise its benefit to society" (HGP Information Archive 1997). The 2003 Fort Lauderdale Agreement (Wellcome Trust 2003) extended the same principles to all large-scale sequencing efforts. The public HGP effort emphasised speed of access: sequences were automatically released into public databases within 24 hours of production, and therefore – unusually – were accessible prior to the publication of any journal article that might analyse and interpret the data. To reassure data producers, the Fort Lauderdale Agreement proposed that users – while making free use of the data – should observe a community etiquette of allowing producers to publish their analyses first. Sarah[23], a senior biomedical journal editor, told me this is known as the "gentlemen's agreement".

These policies, agreements and infrastructures established data sharing norms and laid a path for twenty-first century open data movements – as well as open access movements, according to Strasser (2019:249–50). In fact, there are direct institutional links: the Wellcome Trust was a significant funder of the public HGP (Maxson Jones

[23] Interview: UK, Jan 2018.

et al. 2018) and convened the Bermuda and Fort Lauderdale meetings (ibid., pp. 247, 249), before going on to lead open access and open data policy. However, the earlier initiatives were not branded as "open" in the same way as today's movements and thus, I would argue, were part of a discursively distinct phenomenon. While data sharing may have helped set the scene for open access movements, open access set the scene for open data movements as they spread more widely across disciplines and acquired additional meanings associated with "open".

5.3.2 Crises of reproducibility

The rise of open data has occurred not just in the context of technological means to disseminate large datasets as shared resources. There has also been a growing concern that published scientific research may not be as reliable as it appears: that due to poor research practice, bias, mistakes, or fraud, many published findings may be idiosyncratic to the researchers and circumstances that produced them. This set of concerns has come to be known as the "reproducibility crisis", and while it is not entirely new – similar ideas can be dated to the 1950s (Bastian 2016; Stam 2018) – it has become more visible to research communities at large in the 2000s. John Ioannidis' article *Why most published research findings are false* (2005) had a particular impact, and since then there have been many high profile studies and lamentations on the topic. This has included a *Nature* report in which only 11% of 53 key preclinical cancer biology findings were confirmed (Begley and Ellis 2012); a *Science* report of a similar problem in psychology, with less than 40% of studies successfully replicated (Open Science Collaboration 2015); and a *Nature* survey in which many scientists across disciplines reported being unable to reproduce others' results (Baker 2016). This is in addition to a number of high profile cases of scientific fraud that emerge periodically (e.g. the case of psychologist Diederik Stapel; Callaway 2011).

Multiple causes have been linked with the reproducibility crisis, including failures of quality control by journals, an increase in scale and complexity of projects with a diffusion of responsibility, cognitive bias, problematic statistics, and a lack of transparency (summarised by Leonelli 2018). There have also been multiple critiques arguing that the "crisis" is not all it appears (Fanelli 2018), and that reproducibility is a limited lens through which to view research quality (Leonelli 2018; Stam 2018). The call for open data is nonetheless strengthened and given new meaning in light of this crisis narrative. Where early digital data sharing was framed as a resource to be mined for new results, sharing in the "open" era seems to be about integrity as well as utility. Errors, imperfections or fraud in the data or its interpretation are exposed, and those producing open data know this. Other kinds of scientific openness – of methods, code, laboratory notebooks, and so on – also connect with this intention to interrogate, and signal, integrity. The potential of these kinds of openness to signal trustworthiness to the public, as well as to other scientists (Grand et al. 2012) – and thus its appeal to institutions governing science – also connects with this theme.

5.3.3 Defining "open" in relation to data: grassroots movements

Prominent parts of open data movements have been concerned with establishing a definition of "open" in relation to data, as was the case for "open" access. Moreover, particular lineages of the two movements – access (to articles), and data – appear to be extensions of one another in the sense that within similar communities, attention appears to turn chronologically from the "opening up" of research literature to the data underpinning it. These communities have shared values, particularly the importance of *libre* licensing for any purpose, including commercial. In the UK, grassroots networks of researchers and open advocates began defining openness in relation to data not long after the landmark Budapest, Bethesda and Berlin statements on open access. The Open Knowledge Foundation (OKF), a UK-based global non-profit organisation, was established in 2004 and circulated and released an "open definition" that applies to data in 2005-2006 (OKF n.d.) (**document case study #4, analysed in Section 5.3.4**). In 2009-10 a related group produced the *Panton Principles* ("principles for open data in science", **analysed in the same section**), which refer to the OKF's definition and "state clearly what openness is in relation to public science" (Murray-Rust et al. 2010), with a focus on licensing data for reuse. These definitions and sets of guidance explicitly refer to free and open source software movements. Moreover, OKF's definition is "initially derived from the Open Source Definition", and is viewed as a continuation of Richard Stallman's "ideals of software freedom" (OKF 2015). This indicates that some "open" movements in science – particularly open data movements – are significantly influenced or led by grassroots communities with technological interests and expertise, and in particular, those versed in principles and practices of free and open source software (see discussion on this link in Sections 4.10, 6.3, and 7.8.1).

5.3.4 Documents #4: Open Definition (2006-) and Panton Principles (2010)

- Key cases of UK-centred grassroots advocacy of openness beyond the article.
- The not-for-profit Open Knowledge Foundation (OKF) released its first Open Knowledge Definition in 2006: an early instance "open" being defined as a concept and descriptor in its own right (OKF 2006, n.d.). The latest version, simply titled "**Open Definition**", was released in 2015 (OKF 2015).
- Aligns with a discursive establishment of "open data": the Wikipedia page was initiated by Peter Murray-Rust in 2006, who had an authorial role both of the documents discussed in this box (Anon n.d.).
- Some linkage with early open access movements: advocate Peter Suber, a BOAI author, was also involved in drafting the Open Knowledge Definition.
- An indication of influence: the OKF's definition is cited by UK Government-funded Open Data Institute on their page *What makes data open?* (ODI 2013).
- Explicitly linked with free and open source software. Indicates a lineage of open data with foundational ties to software cultures: "The Open Definition was initially derived from the Open Source Definition [...] Richard Stallman was the first to push the ideals of software freedom which we continue" (OKF 2015).
- A software-derived *libre* meaning of "open" is in the foreground: unrestricted access, reuse, modification, and sharing. Machine readability, open formats, and open licensing are required. The document is largely about licensing.
- Commercial purposes are explicitly included, placing the Definition closer to pragmatic, "neutral" open source values than anti-proprietary free values. The Panton Principles "STRONGLY discourage" limitations of commercial reuse.
- Concise, legal style, portraying "open" as precise, bounded, and operational. "Open" is portrayed as static, binary quality of an object, rather than a value, process, or behaviour. Scientists themselves are thus absent from this portrayal.
- To maximise decentralised, free sharing and reuse, the OKF acts as a centralised gatekeeper of "open", asserting the Definition as a "gold standard" that distinguishes between legitimate openness and "open-washing" (Pollock 2014).
- The **Panton Principles** are an extension of the Open Definition for a scientific data context, drafted by an OKF-associated group (Murray-Rust et al. 2010).
- Premised on a traditional openness narrative: "Science is based on building on, reusing and openly criticising the published body of scientific knowledge".
- Recommends that scientific data be licensed for *more* liberal reuse: i.e. placed in the public domain, with no attribution requirement. This asserts collective ownership values over individuals' attachments to data they produce.
- Implies that scientific data ought to be open, both epistemically and morally.

5.3.5 Open data in research policy: UK and Australia

At the same time as grassroots movements began defining "open" in relation to data, a top-down governance of openness in relation to research data was beginning to take shape. In 2004, the OECD issued a Declaration on Access to Research Data from Public Funding in which 34 countries – including the UK and Australia – committed to "the establishment of access regimes for digital research data from public funding" in accordance with several principles including an undefined "openness" (OECD 2004). In the years following, governments, research institutions, and publishers increasingly adopted stances that require or encourage openness of data from researchers, often as an extension of a stance on open access to research articles. In the UK, this tended to take the form of policy statements or mandates from institutional actors familiar in open access discourse: the Wellcome Trust and government bodies, particularly the research councils (overseen by RCUK/UKRI). In Australia, national-level policy is less strict, but there has been considerable investment in data sharing infrastructures (Shearer 2015:12). Researchers in both countries encounter international trends, such as the increasing adoption of data sharing rules by journals. In part due to the diversity of data types and how they are situated within any given research project, open data policies have to date been less prescriptive and trackable than open access policies.

Open data in the UK

In the UK, moves by major funders towards open data policy began in about 2007, when the Wellcome Trust began encouraging researchers to share data "as soon as feasible" and "with as few restrictions as possible" (Wellcome Trust 2007). In the same year the Medical Research Council (MRC) stated that "MRC research data…must be made available for new research purposes in a timely, responsible manner" (MRC 2007). By 2011, RCUK's *Common Principles on Data Policy* had extended this recommendation across the research councils (RCUK 2011). The Royal Society brought scientific openness into the UK policy spotlight in 2012 with its report *Science as an open enterprise*, which is primarily a manifesto for open data (**document #5,**

analysed in Section 5.3.6). The report coined the term "intelligent openness" as a way to define data that are sufficiently open: they must be "accessible, useable, assessable and intelligible" (The Royal Society 2012:12). It makes specific recommendations for how audiences including scientists, research institutions, funders, research assessors, learned societies, scientific journals, industry, and governments, should enact openness.

5.3.6 Document #5: The Royal Society report (2012)

- Landmark report advocating open research data titled *Science as an open enterprise*. Contemporary with the Finch report on open access (see p. 128). Unlike the Finch report it centres "open", including the use of "open science" as an umbrella term, in a limited sense: covers only open access and open data.
- Blends diverse values into "open": most salient are epistemic and technological advancement; public accountability; economic gain; and civic benefit.
- Histories in both reports co-locate origins of present science and 17th century journal publishing. The Society frames this as a history of openness in science:

  > *"Much of the remarkable growth of scientific understanding in recent centuries is due to open practices; open communication and deliberation sit at the heart of scientific practice." (p. 13)*

- Represents addition of a prestigious establishment voice to open data movements. However, it does not cite or align with the OKF grassroots definition of open data, and instead develops its own concept of openness.
- Coins the term "intelligent openness" that goes beyond "mere disclosure of data". Data must be "accessible, intelligible, assessable and usable"[24] (pp. 14-15). Intelligent openness is specific in different ways to the OKF's Open Definition; the former more contextually nuanced, the latter more technically operational.
- The Society portrays data openness as somewhat relational and socially embedded, rather than "open" being an intrinsic, technical feature of data. For example, they suggest that data communication should be audience-specific.
- Openness with the public realm is a defining feature: a precursor publication was titled *Science as a <u>public</u> enterprise* (my underline) (Boulton et al. 2011). The Society's entry into "open" discourse seems to have resulted from an alignment between open and public science, making its motivations distinct[25].
- Exceptions, including withholding of commercial, dual use, and personal data. Legitimises industry as non-contributing beneficiary of openness.
- Acknowledges labour of making data open, and suggests that only data relevant to the public needs opening. Comparatively early pragmatic insight, with potential to reinforce science-society boundaries that are otherwise challenged.
- Scientists are the primary audience, expected to need incentives to change.
- Openness is good for science and will stamp out "bad science": an epistemic virtue.

---

[24] According to a key associated interviewee, FAIR data developed from intelligent openness.
[25] The same interviewee gave reasoning: "open" was seen as a narrower and more achievable step on the way to a public science that "blur[s] the divide between professional and amateur in new ways" (p. 8). I was told the Climategate email scandal, and associated loss of public trust, was the catalyst for this report.

Major UK research funders and institutions – RCUK/UKRI, HEFCE, the Wellcome Trust and Universities UK – took a decisive and unified stance on the topic in 2016. Their *Concordat on Open Research Data* (HEFCE et al. 2016) (**document #6, analysed in Section 5.3.7**) sets out ten "clear and practical principles for working with research data", introduced by the first principle: "open access to research data is an enabler of high quality research, a facilitator of innovation and safeguards good research practice" (pp. 2, 7). It sets out expectations for researchers and their employers and funders, and is intended to guide data policies of all the signatory institutions. The Wellcome Trust, for instance, adopted a more specific, far-reaching policy in 2017: researchers "must make any data, original code or materials that underpin published research findings accessible to other researchers at the time of publication" (Carr 2017).

Open data in Australia

In Australia the two national research councils, the ARC and NHMRC, have taken less stringent stances on open data: in 2007 they co-developed the *Australian Code for Responsible Conduct of Research* (NHMRC, ARC, and Universities Australia 2007) which emphasises data management but not online sharing (Neylon 2017:10). By 2016, both research councils were encouraging data sharing. This became a "strong" encouragement on the part of the NHMRC, but not a requirement, in 2018 (Rowhani-Farid 2018:28). Australia is viewed as a leader, however, on a data infrastructure level: the federal government funded an Australian National Data Service (ANDS). In 2007 the ANDS invested in and hosted centralised data services including for research (Shearer 2015:12). In 2018, ANDS was incorporated into the Australian Research Data Commons (ARDC), that "pull[s] together industry, universities and the public sector to coordinate the sharing and open publication of research data" (Wallace 2019). The Australian government also co-founded the Research Data Alliance in 2013 with the European Commission and the US Government (RDA 2016).

5.3.7 Document #6: Concordat on Open Research Data (2016)

- Major UK policy development: brings together national institutions with power to enforce open data policies: HEFCE, RCUK, UUK, and the Wellcome Trust. Concise: does not specify exact rules for open data, but gives ten "Principles".
- Reference document guiding data policies. Data is what "underpins the answer to the research question, and can be used to validate findings" (p. 3).
- Open research data framed as a logical "next step" given progress towards open access, and positioned as part of "the UK's open science ambitions" (p. 1). Thus open science is portrayed as overarching policy agenda with a moving frontier.
- Again, diverse meanings are signified by "open": economic growth; epistemic efficiency; and public trust and support as a result of transparency; reputational gain. Legitimises closure for confidentiality, security, and commercial reasons.
- Like the BOAI, Finch report, and Royal Society report, the *Concordat* imbues "open" with *libre* values that go beyond access, without the technical precision of the Open Definition or Panton Principles – which are not referenced.
- Top-down dynamics are downplayed; the *Concordat* is positioned as a product of the research community itself, despite its institutional origins and working group composed of funders and university leaders: "It is not the intention to mandate, codify or require specific activities…." (p. 4).
- A centring of researchers and their communities continues throughout: "researchers are a diverse group of people operating in many different cultures and contexts […] there can be no 'one size fits all' approach" (p. 5).
- Marked departure from open discourses that frame researchers as a homogenous group to which universal requirements apply.
- But a transition towards data openness is still framed as a universal process with a single direction: some fields are at an "early stage" on a "journey in which the research community will participate over the coming years" (p. 4).
- Acknowledgement that open data carries a cost: for infrastructure and services, specialist support, training, and researchers' time. Where a financial framing is dominant for open access it has tended to be secondary for open data.
- "The right of the creators of research data to reasonable first use is recognised" (p. 12): echoes Fort Lauderdale Agreement (Wellcome Trust 2003) for genomic data. Acknowledges competitive pressure and researcher attachment to data.
- In contrast to Panton Principles' public domain approach, expects data citation.
- A moderated, contextual approach that aligns with concurrent discourse at European Commission: "as open as possible, as closed as necessary" (EC n.d.).
- Open data enables high quality, good research practice: epistemic virtue.

5.3.8 Open data in journal policy

In addition to policy stances from funders and institutions, and the provision of open data infrastructure, journals have had a key influence on the international open data landscape. Some journals have long-standing data availability requirements for certain data types, notably, genomic and crystallographic data; the earliest policy of this kind was made in 1988 by *Nucleic Acids Research* in relation to sequence data (Strasser 2019:188–89, 233–34). The adoption of data sharing policies by a wider range of journals, for a wider range of data types, has been more recent. The open access publisher PLOS announced a data access policy in 2014 (Silva 2014), framing it as an extension of their open access mission and formalisation data sharing expectations between researchers. According to the policy, “authors must make all data publicly available, without restriction, immediately upon publication” and must provide a “Data Availability Statement” to be published with the article. Other publishers announced similar policies afterwards, including larger, traditional, commercial publishers Springer Nature, Science, and Elsevier – although these varied in the extent to which they mandated unrestricted online access to research data (Federer et al. 2018:1–2).

5.3.9 A note on “open data”, Research Data Management, and FAIR data

Although the use of “open data” terminology took hold in the context of online data sharing in the 2000s and 2010s, it is not used as consistently as is “open access” to research literature. Earlier policy interventions – such as those from the OECD, MRC, and Wellcome Trust – do not use the phrase “open data”, alluding instead to data management, sharing, and access. “Open access to research data” remains a common phrase, perhaps indicating an origin of such policies in research literature context. However “open data” terminology has been used liberally by grassroots movements and, since the 2010s, by prominent UK and international organisations including the G8 (G8 2013; HEFCE et al. 2016; The Royal Society 2012).

Part of the reason for this discursive salience is a cross-pollination with movements pursuing data openness in a broader range of contexts: for instance, “open data” often implies the opening of government-held data, associated with public transparency agendas that date back to Freedom of Information Acts (Yu and Robinson 2011). Research data can fall into this category especially where they are government-funded. Thus the rise of “open data” is also associated with the pursuit by the UK Government of a broader “open” agenda for several years from 2010, with the creation of an Open Government Licence (compatible with CC-BY used for research articles; UK Government 2010); the launch of data.gov.uk (BBC 2010); the publication of an Open Data White Paper (HM Government Cabinet Office 2012); and the funding of an Open Data Institute launched in 2012 (ODI n.d.). The Australian federal government has likewise hosted data.gov.au since 2013 (Comprehensive Knowledge Archive Network n.d.) and it hosts a “node” of the ODI (ODI n.d.). With a statement in 2015, it “formalise[d]” its “commitment to open data and data-driven innovation” (Department of the Prime Minister and Cabinet n.d.). Developments like these reflect an intensification the discursive significance of “data” as well as “open”.

There is, however, some evidence that “open data” terminology is no longer, or has never been, widely preferred in academic research policy. Policies in the UK and Australia – as opposed to position statements – tend to emphasise research data management (RDM) and the creation of advance plans for data stewardship within which sharing is just one consideration. Neylon (2017) reports that such policies are increasingly common, and that they tend to require Data Management Plan (DMPs), although this is not necessarily enforced. Some policymakers believe this will raise awareness of data issues amongst researchers, where others worry it will create a “tick-box” culture (p. 2). Another development in terminology has been the coining of “FAIR data” – findable, accessible, interoperable and reusable – by a group drawn from academia, industry, funding agencies, and scholarly publishers (Wilkinson et al. 2016). According to Higman et al. (2019), “little in the research data field has gained such traction and universal acceptance as the FAIR data principles” (p. 2). Like RDM, FAIR can apply more widely than “open” – for instance, to access-controlled data. There are indications that policy emphases on RDM – and possibly in the future, FAIR

– may be strategically preferred, for several reasons. Firstly, they can apply to a broader range of datasets including access-controlled data. Secondly, FAIR in particular is more specific than "open" (Barbour 2016). Thirdly – of particular relevance – staff implementing research data policies at universities have adopted RDM language tactically. They have found that it represents a foundation and middle ground for engagement with researchers who may react negatively to "open data" or data sharing terminology. Danny Kingsley[26], a leader of open policy at a UK university, mentioned this strategic shift. It is also mentioned by LIBER, the Association of European Research Libraries (Christensen-Dalsgaard et al. 2012:2) and reported by Pinfield et al. (2014:17). The flexible meaning of "open" may therefore, in the case of research data, have become a liability for those attempting to implement it as policy, resulting in a retreat from "open" terminology.

5.4 Open science

As landmark recommendations in favour of open access and open data have radiated and intensified, they have been accompanied by and included in the growth of broader "open science" discourses and movements. Like "open access" and "open data", "open science" as a term was intermittently in use prior to the 2000s with links to its present meaning: a couple of instances just prior to the turn of the century are noted in Section 4.10. However, a concerted rise in its salience appears to have taken place from the mid-2000s and especially in the 2010s, led not by the emergence of a consensus definition but as a result of numerous movements – again, both bottom-up and top-down - pursuing far-reaching, online, digital changes to the way science is practised. In 2008, the chemist Jean-Claude Bradley and co-authors wrote of a "growing movement promoting more Open Science" as part of a funding proposal to support "Open Notebook Science" (Bradley et al. 2008). They described this approach, which Bradley had been practising since 2005, as a "logical extension" of open science trends that "expos[e] a researcher's complete record of progress to the public in near real time [online]" (p. 1). The Open Knowledge Foundation, which had by this time

[26] Head of Scholarly Communication, University of Cambridge; interview: UK, Feb 2018.

established its Open Definition in relation to data (see above, p. 133), started an "open science" mailing list in 2008 (OKF n.d.), bringing together a community under that heading. A few years later, Australian physicist and computer programming researcher Michael Nielsen (2011) wrote an influential popular book, *Reinventing Discovery*, in which he imagined a future, online networked practice of research that he referred to as "open science".

Reflections on open science as a single multifaceted phenomenon were in evidence a few years later (e.g. Delfanti and Pitrelli 2015; Leonelli, Spichtinger, and Prainsack 2015). For example in 2014, the online book *Opening Science* was published: a scholarly advocacy document advising "researchers, scientists, decision makers, politicians, and stakeholders" on the "*basics*, the *tools*, and the *vision* behind current changes we see in the field of knowledge creation" (p. v, original emphasis; the introductory chapter is **document #7, analysed in Section 5.4.1).** In a highly-cited chapter of this book, Fecher and Friesike (2014) characterise open science as an "umbrella term" covering "five schools of thought" (infrastructure, public, measurement, democratic, pragmatic). By this time, open science was often conceptualised in this way: as a catch-all that refers to a set of distinct but related practices, including not only open access and open data, but – for example – open methodology, open peer review, open preprints, open code, open citations, citizen science, and open notebook science.

5.4.1 Document #7: Towards Another Scientific Revolution (2014b)

- Introductory chapter of an edited book, *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing.* The book is an academic work, grassroots advocacy document, and a guide.
- The book's production reflects its "open" values: it was authored using a collaborative software tool, published under a Creative Commons licence, and invites "everyone" to "contribute to and reuse its content" (Preface, p. vi). The content of the book has in practice remained relatively fixed (Github n.d.).
- The introductory chapter is by the editors, Sönke Bartling and Sascha Friesike, researchers with medical and digital innovation/engineering backgrounds.
- Provides sweeping introduction to the concept of "Open Science", its historical context, and issues involved in transitioning to an open science future. "Obstacles" must be overcome to achieve "complex cultural change" (p. 3).
- Researchers are centred as an audience. They are framed as free agents – "every researcher has to decide for themselves which technologies and methods they will [use]" (p. 4) – who are also driven by incentives.
- Moreover, the involvement of researchers is framed as pivotal to the outcome of a digital revolution in science, despite an anticipation of reluctance (p. 11):

  > *"…much will depend upon whether researchers become the leading force within this transition, or whether they play a passive role driven by other stakeholders of the research process. (p. 12)*

- "Open" in science is thus constructed as a product of meaningful engagement by researchers; more than an abstract quality of information, or its processing.
- A long view of history frames the narrative, beginning in prehistory with a link between knowledge and civilization. The 17th century "first scientific revolution" is identified as the origin of open science, and the "groundwork" upon which a present-day open revolution is built (p. 6)[27]. Nielsen (2011) is a key reference.
- The history is strongly presentist: 17th century researchers are viewed as proto-professionals who were secretive whilst awaiting incentives – a system for "assessing the value of a contribution", i.e. journal publication (p. 6).
- The text alludes to complexity and contingency, but its narrative tends towards technological determinism (see Chapter 4), depicting cultural transformations of science as led by shifts from papyrus, to print, and now to the Internet (p. 8).
- Portrays difficulty in defining open science, Science 2.0, etc.: "All of these umbrella terms struggle to find a clear definition" (p. 10).
- Epistemic virtue is central: openness means immediate, complete, networked communication of ongoing science to all, bringing science into an ideal form.

[27] Non-technologically focused histories, e.g. of the commodification of publishing, are absent.

The European Commission played an important top-down role in solidifying this usage from 2015, when it adopted “open innovation, open science, and open to the world” (EC 2015a) as major policy themes. The term “open science” was adopted after a public consultation indicated it was preferred to the related term “Science 2.0”, used previously (EC 2015b:6). The EC subsequently launched numerous “open” initiatives: the Open Science Policy Platform (OSPP); the European Open Science Cloud (EOSC); and the Open Science Monitor which tracks global trends in relation to open access, open data, and “outputs such as open code, open hardware, the use of collaborative platforms between scientists and the ‘citizen-science’ phenomenon” (EC n.d.). This agenda has, with its associated terminology, had an impact in the UK. For instance, the League of European Research Universities (LERU) – which includes five of the most prestigious universities in the UK – recently published a “roadmap for cultural change” towards open science in universities (LERU 2018) (**document #8, analysed in Section 5.4.2**).

The terms “open research” or “open scholarship” are sometimes preferred to “open science” in English-speaking contexts, due to their broader disciplinary connotations. This also indicates a the growing influence of such movements, which have often been shaped in natural science contexts but are now being introduced at a policy level as universal. The Wellcome Trust, for instance, now has an “open research” division that attends to open access; to data, software and material sharing; and to clinical trials data (Wellcome Trust n.d.). The division also runs *Wellcome Open Research*, a publication platform that supports open practices including open peer review; administers an Open Science Prize and an Open Research Fund; and studies how researchers can be incentivised to take up open practices.

5.4.2 Document #8: LERU's Open Science advice paper (2018)

- Titled *Open Science and its role in universities: a roadmap for cultural change.* Report advising how universities can and should embrace open science.
- Produced by the League of European Research Universities (LERU), a network of 23 "leading" European universities, including five in the UK.
- Open Science as an overarching term and concept is given institutional weight and legitimacy – it appears 295 times. This is a European Commission (EC) usage; Open Scholarship is asserted as more correct in English, for its breadth.
- Recommendations based on eight "pillars of Open Science" identified by the EC, including: open access and FAIR data; training, incentives, and new metrics for open science; research integrity; and citizen science.
- For individual researchers, *libre* open access, open data and open software are expected; further, openness *during research* is encouraged (preprints and data). This begins to extend open practice as envisioned at an institutional level.
- "Open" represents epistemic efficiency, integrity, transparency, reproducibility, interdisciplinarity, public engagement, publicity, and professionalism.
- "Cultural change" is emphasised, bringing to the fore a marginal but persistent feature of open policy and advocacy discourse. "Cultural" tends to refer to complex, recalcitrant elements of a system being "opened" – often, researchers.
- Here, again, cultural change implies a shift in researcher behaviour. Change is framed as "not a top-down activity" (p. 8), in contrast with suggested strategies: e.g. senior leadership, policy interventions, monitoring, training and incentives.
- Dialogue and consultation are mentioned, but seem directed at persuasion: "Leaders should work with the community to explain why change is necessary and to support change" (p. 8). However, "trust and confidence" within the hierarchy is valued, and the difficulty of cultural change is not underestimated:

  > *"There are real dangers is trying to introduce new practices without carrying the academic community with the leaders [...]. It would be wrong to think that Open Science is simply a blueprint which can be introduced in a mechanistic way..." (p. 9)*

- An epistemic virtue framing is deepened by the assertion that researchers themselves are improved by openness, especially transparency, which is "good for researchers and good for research" (p. 5). Moreover, the paper begins with the phrase, "Open Science is not about dogma..." (p. 3), implying that "open" indeed wields an abstract discursive power, with moral overtones. The denial further evokes epistemic virtue, as it manifests in the case of objectivity: as the moralised rejection of value-laden scientific practice (Daston and Galison 2007).

Another example of this all-encompassing approach, but from an international grassroots perspective, is a document that was drafted online in a publicly accessible Google Doc, *Foundations for Open Scholarship Strategy Development* (Tennant et al. 2019, preprint) (**document #9, analysed in Section 5.4.3**). The document details short- mid-, and long-term strategies for implementing "open scholarship" at individual, laboratory/department, institute, and national levels. A profusion of diverse practices appear in the document, from those I have mentioned to: use of social media; use of open software and collaborative writing tools; seeking out open material for re-use; participation in online forums; making open achievements visible on a CV etc.; advocating and educating others about open approaches; adoption of an "open mindset"; refusal to engage with restrictive or proprietary platforms or publishers; building support networks for open scholarship; and so on. The document deliberately eschews a definition for open scholarship, recognising that open scholarship "encompasses many disciplines, practices, and principles" (p. 1). As such, it portrays open scholarship (or science) as a category actively under construction, fuelled by a kind of productive ambiguity. This ambiguity does not, however seem to obscure a sense of collective purpose and identity, and perhaps contributes to one. Moreover, "open" is constructed as an inward disposition ("mindset") and outward identity ("make individual contributions to openness that are visible in public", p. 4) to be reinforced in the self and created in others through reflection, education, training, and community organising.

Collectively, the growth in movements towards open access, open data and now an all-encompassing open science has been rapid and extensive over only two decades, with a particularly visible rise in policy agendas in the last five to ten years. Notable for the purpose of the present research is that the rise of "open" terminology and its associated practices and values will have fast become salient to researchers who were not previously engaged in this discourse, as the attention of funders, universities and publishers has turned towards governing openness.

5.4.3 Document #9: Open Scholarship Strategy document (2019)

- *Foundations for Open Scholarship Strategy Development* (Tennant et al. 2019) is a grassroots community advocacy document; the latest draft from January 2019.
- Its production reflects "open" values: collaboratively drafted online by 32 authors, with editable public drafts; available in several languages and digital formats; and shared as a preprint under a liberal Creative Commons licence.
- Extensive scope: characterises "open scholarship" as an international movement and strategically analyses its strengths, challenges, opportunities and threats. Identifies short-, mid-, and long-term goals from individual to national levels.
- "Open scholarship" is used as a more inclusive version of "open science" or "open research", reinforcing the idea that "many disciplines, practices, and principles" (p. 1) are encompassed by an expanding open movement.
- The term is not defined, allowing for continued expansion; indeed, the authors theorise open scholarship as a boundary object (drawing on Moore 2017). This encourages flexible interpretation, with some coherence for integrity (p. 15).
- The text emphasises plurality but asserts a shared identity and narrative: "…how we can more effectively work together as a global community" (p. 1). Attempts to build a coordinated focus, with "well-defined objectives" for success (p. 20).
- The opening up of "entire traditional research workflows" is envisioned (p. 20), and the expansion of policy beyond open access and data is encouraged (p. 4).
- The diversity of "open" values is accepted, theorised, and used to organise strategy, based on Fecher and Friesike's (2014) five schools of open science. A "community and inclusion" school is added (in brief), reflecting an expansion of open values as centred in the global north. Cognitive justice appears alongside e.g. equality, transparency, and collaboration (p. 17), and OCSDNet[28] is cited.
- A tension is acknowledged between community/non-profit and capitalist values of "open". This is a major fault line that elsewhere tends to be obscured.
- Researchers include advocates, but are mostly portrayed as resistant to change, busy, and in need of awareness, training and incentives. Questions why some would adopt open practices without a holistic "open scholar" approach (p. 30).
- Clear epistemic virtue framing: open "will result in a better (i.e. rigorous or fairer) research process" (p. 17). Moreover, the belief that open practices are "generally a *good thing*" is identified as the "most widespread commonality between Open Scholarship stakeholders" (p. 16): i.e. epistemic virtue anchors a collective identity. Also portrayed as practice of the self – a mindset, a pledge – and a virtuous identity to be promoted, role-modelled, and taught.

[28] The Open and Collaborative Science in Development Network is centred in the global south and has a manifesto advocating for inclusive, situated open science (OCSDNet n.d.)

5.5 Conclusion

This chapter addresses the moving target of "open" in science through the early twenty-first century through its expansion in meaning and impact, especially in the UK and Australia. My analyses of UK-relevant advocacy and policy documents, although they are a limited selection, illustrate qualitative features and patterns in open science discourse that address my research sub-questions.

5.5.1 Historical framings of openness

Some of the documents allude to – or centre – a long history of openness in science (theme of research sub-question a); this phenomenon is analysed in Chapter 3. However, this was not a universal feature in the documents, and the histories told are were always the same, nor were they told for the same reasons. For instance, the PLoS letter (2000) does not make an historical case. The BOAI (2002) does evoke an "old tradition", but not to argue that science has long been open, but to draw attention to a contingency: that scholarly publishing is not-for-profit, making open access possible. The Finch report (2012) evokes a long history of *publishing* rather than openness. But once open science is envisioned holistically, framings of science as enduringly or essentially open seem to become more prevalent (Bartling and Friesike 2014b; The Royal Society 2012), aided by historical accounts that place the origin of open science in the seventeenth century (David 2008; Nielsen 2011). Other documents lean less on historical detail, some referring to taken-for-granted traditions of openness (Murray-Rust et al. 2010; Tennant et al. 2019) or referring piecemeal to historical touchpoints for openness (16th century moveable type, LERU 2018). As the discourse moves towards "open research" and "open scholarship" to include the humanities and social sciences in Anglophone contexts, it seems like an origin story centred in the natural sciences will become less relevant.

5.5.2 Meanings of "open"

The document analyses illustrate an accumulation of meanings and values associated with "open" as it gains salience and impact (theme of sub-question b). This begins from a broad base. Even before the word "open" is applied (PLOS 2000), online access to the scientific literature is framed as an advancement of science with overtones of neoliberal productivity; an assertion of public ownership rights; and a protest against the power of publishers. The BOAI (2002) again emphasises speed of epistemic advancement, and public benefit, and underlines education, equality, cost efficiency, and visibility for individual researchers. Openness is also strongly tied to an embrace of technology, and *libre* re-use from software cultures, not yet named as such, makes an appearance. From this point forth, "open" tends to be defined according to technical and legal re-use permissions as well as access. With the entry of policy literature – the Finch (2012) and Royal Society (2012) reports – "open" is valued not only as a matter of universal and efficient resourcing, but as a check on science: a form of accountability and an opportunity to build public trust. This dimension intensifies as data – and subsequently, methods, notebooks, etc. – enter the frame, and openness comes to mean not only disseminating results, but showing your working: transparency and reproducibility. The public dimension also expands, suggesting a blending across science–society boundaries, although e.g. the LERU paper (2018) does so within existing institutional and scientific paradigms ("...for citizen science to be "science" it needs to adhere to scientific standards", p. 21).

In the late 2010s, an exceptionally wide and still expanding set of meanings is associated with "open", e.g. including the above as well as interdisciplinarity; data stewardship; standardisation for machine readability; reform of metrics and incentives; responsible, professional conduct; social inclusion; and cognitive justice. These latter critical reckonings with inequality seem to be on the fringes of open science in the global north. Earlier framings and many current ones assume that the universality of openness signifies equality. Openness as a form of epistemic advancement remains at the forefront of this diverse set of values throughout the 2000s, especially in bottom-up advocacy. Additionally, throughout, openness remains in tension with capitalist values. Its uptake in policy has no doubt been aided by its association with efficiency and economic growth (see Section 4.7), and acceptance of commercial re-use goes to the core of an "open" lineage established by open source software (see Section 5.3.4 above and Moore 2017; Tkacz 2012). However, multiple lineages of open science began in protest against corporate power and proprietary closures (Moore 2017), a dimension that tends to be obscured in both grassroots and policy documents tend towards pro-technology, business-friendly, apolitical framings. Sometimes, e.g. in the *Open Scholarship Strategy* (Tennant et al. 2019), this tension is articulated and it becomes clear that open science is increasingly co-opted by large corporate powers (as forcefully articulated by Mirowski 2018), a risk of its politically ambiguous character. Organisations such as the Radical Open Access Collective (ROAC) work to revive an alternative to market-driven visions (Adema and Moore 2018).

The increasing breadth and salience of "open" in the twenty-first century align with my conclusion, in Chapter 4, that the ambiguity of the term – and even the conflicts it encapsulates – allows it to accrue power and to hold together causes and constituencies with different histories and motives. Moore (2017) suggests that open access can be theorised as a boundary object (Star and Griesemer 1989) due to these sorts of qualities, and Tennant et al. (2019) suggest that the same applies to all of open scholarship. I concur, but also postulate that "open" is a different kind of boundary object or something else altogether: operating at higher level of abstraction, with *almost* no common ground internally except, as the latter authors observe, an agreement that openness is "generally a *good thing*" (ibid., p. 16, original emphasis).

5.5.3 "Open" as an epistemic virtue

The documents portray a gradual shift towards the construction of openness as an epistemic virtue (the theme of sub-question c, defined in Section 2.2.1). The earliest documents, portraying the rise of open access, depict openness as the right thing to do (e.g. good for society) because it makes science more accessible, integrated, and faster. Ethics and epistemology are arguably not combined here: the dissemination of science is transforming, but not necessarily science itself[29]. The Finch report refers obliquely to "better research" resulting from open access, hinting at epistemic virtue but not specifying how this research is "better". Something more significant happens when open data enters the picture: openness is portrayed as an improvement at the core of science. For example, the Royal Society (2012) portrays openness as an essential feature of good science: the very quality that "permits others to identify errors, to support, reject or refine theories and to reuse data for further understanding and knowledge" (p. 7). Open access, and particularly open data, are framed as extensions of that principle. Where openness connotes transparency, it can "deter, detect and stamp out bad science" (ibid., p. 8), improving epistemic integrity.

The LERU advice paper (2018) states that openness – transparency in particular – is "good for researchers and good for research" (p. 5), bringing together knower and knowledge, and evoking an integrity that is both epistemic and moral. An epistemic virtue framing is brought to full fruition by the *Open Scholarship Strategy* (Tennant et al. 2019), which not only depicts open knowledge as good knowledge, but depicts openness as a blended moral-epistemic practice of the self that is internalised by "open scholars" (p. 30). The document suggests scholars take on an "open mindset" (p. 4), adopt "an implicit or explicit (shared) open science pledge or code of conduct" (p. 7) and share and educate others on this principle. Moreover, a belief that open scholarship is "generally a *good thing*" (p. 16) is identified as the most widespread commonality in the movement, suggesting that – despite its almost complete flexibility and ambiguity – "open" may be anchored by its epistemically virtuous qualities, at least from a bottom-up perspective.

5.5.4 The role of scientists

Finally, the documents depict distinctive changes in the role of the scientist (or researcher, or scholar; theme of sub-question d) in relation to open science over the

[29] Open access could be seen as a qualitative transformation of science at a systems/community level. For a discussion of this idea in relation to epistemic virtue, see Section 9.3.1. However, these early documents do not emphasise such an interpretation.

course of the 2000s. The earliest of the documents (PLOS 2000) is written *by* scientists – in powerful positions – on behalf of scientists, and primarily casts free access to research literature as an epistemic project, with great benefits for science and scientific communities. Here, scientists have agency, political ambition, and make little distinction between themselves and the lay public – the democratic right to knowledge applies to all. The BOAI (2002) depicts scientists/scholars as active agents once more, but from the distance of the third person: nonetheless, any individual scientist is depicted as capable of practising open access without a need to wait for systemic change. These depictions change when open access, data, and science are framed in policy. In the policy documents I examined, researchers become one of many categories of stakeholder to be coordinated in a process of systemic change. They tend to be portrayed as passive – at most, motivated by their immediate research aims and competitive pressure – and are expected to require explanations, education, and incentives in order to embrace openness.

In this context, culture change becomes a policy interest in relation to openness. It develops from a passing mention in the Finch report (2012) to an overarching focus in LERU's advice paper (2018). Culture change evokes systemic change, but in documents often refers to research culture – and more specifically still, to behaviour change expected of researchers. This "culture" change is seen as slow, complex, and beset by passive resistance. The discursive shift from an active to a passive researcher has an obvious origin: earlier grassroots framings are made by, and in relation to, researchers who are on the leading edge of open movements; policy is instead concerned with shifting an entire population of researchers. Nonetheless, this shift may have profoundly shaped on how researchers are conceptualised within open science movements. The *Open Scholarship Strategy* is partly written by advocate-researchers, and portrays a struggle to be patient with the "apathy" of researchers who do not practise or commit to openness, and are "generally resistant to change" (Tennant et al. 2019:30, 34). This connects with a prevalent portrayal of researchers as under the control of incentives in a "publish or perish" culture[30]. Cultural change in

[30] Freese and Peterson (2018) similarly observe, in in the context of the reproducibility crisis, that an "economic view of the self" predominates, which "emphasizes responsiveness to incentives provided by institutions rather than […] morals or socialization to scientific norms".

this context is liable to be reduced the top-down alteration of incentives, in order to reward “open” behaviour. Many policies and discourses currently tend in this direction (e.g. Leonelli et al. 2015; “develop individual HR criteria for recognising and rewarding Open Science”, LERU 2018:18; "the EC [European Commission] should encourage the development of new indictors […] to measure and support the development of open science”, Wilsdon et al. 2017:16). This is accompanied by discourse acknowledging that metrics-based incentives must be used carefully to avoid counterproductive effects (Hicks and Wouters 2015).

This portrait of open access, open data, and open science as they have developed and manifested in advocacy and policy contexts is intended as a contribution in itself that addresses my research questions. It also establishes the context in which my interview study was conducted. Interviewees’ experiences have been shaped by the discourse and policy environment characterised here, and some advocate/policymaker interviewees have been involved in shaping this environment. The majority of interviewees are biological scientists, whose accounts align and contrast in instructive ways with the framings presented in this chapter. My findings from the interview study are presented in the next three chapters.

Chapter 6 | "We are obliged"

Open access to research articles

Molly was working towards the later stages of her PhD in behavioural ecology and evolutionary biology when we spoke. She had several papers written, lined up to submit to journals. They had long been awaiting sign-off by a group who gathered the data at an exclusive, remote field site. She had just received her first permission.

She had chosen a prestigious multidisciplinary journal, and prepared her article to meet its requirements. She only realised later that it charged a standard fee to publish, plus a fee for open access if she chose that option. "I have to find this money that I don't know where I'm going to find it from […] for the first paper finding money to get it published in the first place is enough of a hurdle."

Molly equivocated about open access.

"I know there are a lot of publications starting up that are entirely open access, which is wonderful. But also at this stage a lot of them have, are mostly very low impact factors. And I can't afford to publish my stuff in something that's a really low impact factor if I only have three papers and then I'm looking for a job. Then I need to come with these papers that are from quite high end journals potentially. So, as much as I love the idea of open access, I personally am not willing to put my career on the line for it sadly, which is probably a bit hypocritical, but yeah."

Molly, PhD student, behavioural ecology

6.1 Introduction

I set out to ask the overarching question: *how is the meaning of "open" (or "openness") being constructed in the context of science?* This chapter presents and discusses the first of three significant answers to this question as expressed by the biological scientists in my interview study: open access to research articles. This finding was not unexpected: open, online access to research articles, or other scholarly outputs, is one of the most salient categories of "open science" that can be found in policy and advocacy discourse; it has played a pivotal, pioneering role in the historical and continuing development of open science movements. Moreover, by 2017-2019 when these interviews were conducted, even scientists without connections to open movements might have become familiar with open access via multiple routes: the policy requirements of major funders in the UK and Australia; the visibility of open access publishers in the life sciences including PLOS, BioMed Central, PeerJ and eLife; and coverage in the popular and scientific press for a decade or more. However, I could not have been sure that open access would feature as a salient and distinguishable category of openness for scientists, and I could not predict how it would be understood, framed, and prioritised by scientists in relation to their practices and principles. Although open access is defined in specific ways in advocacy and policy literature, scientists may not encounter such definitions directly, and thus could be expected to form their own. This chapter unpacks a set of salient framings and patterns that characterised open access in scientists' – and sometimes advocates' and policymakers' – interview responses. These framings were: financial; journal-based ("gold", as opposed to "green", see Section 5.2.1); concerned with "impact", reputation, and sometimes quality; and centred around compliance.

6.2 The salience of open access in interview

Of the 40 scientists I interviewed, half mentioned open access, open publishing, or the accessibility of research articles, in their response to my initial, broad question about openness: "What first comes to mind when you think about openness in science?". More than three-quarters of interviewees [31] mentioned it spontaneously at other points in the interview, and all showed familiarity with the concept when I asked directly. At surface level, this indicates that open access is a common association with "openness in science" in many scientists' day-to-day professional lives. The converse should also be noted, however: half of the interviewees did *not* mention it in their initial response, and many [9] never brought it up of their own accord. This observation is noteworthy for a topic that is highly salient in policy and advocacy terms, particularly given that it was difficult to avoid priming interviewees on the topic. My past employment by an open access publisher was known to some interviewees, and sometimes a reference to this or to open access became necessary, for clarity, in an otherwise broad, non-directive interview introduction. Thus open access was salient, but was mentioned less than could have been expected.

6.3 Open access as a forerunner: earliest memories

Most of the scientists who I asked recalled a time before the term "open" was part of the language around science. Open access, more than any other category or concept, was associated with the emergence of such terminology: over half of those who were asked [17/30] mentioned open access among their earliest recollections of "open", and for a significant subgroup [10/17] it was the only topic of such recollections. This latter subgroup were mainly established scientists who had completed their PhDs in the 2000s or before. Most would thus have encountered the rise to prominence of open access (see Section 5.2) as already-established scientists. Roger[31], who completed doctoral training in the 1980s, recalled how "open" terminology "…somehow diffused into the language I guess maybe five or six years ago, more than that probably. When those first [open access] journals started to come out." Similarly Greg[32], who is of a similar academic generation – and went on to become Editor-in-Chief of a PLOS open access journal – reflected: "If I had to hazard a guess, I would say that the first time that I heard about it [open/openness in science] was open access. Yeah, so I'm gonna guess that it's a relatively new term. As apart from just a generic adjective." These interviews depict the transformation of "open" into a buzzword representing a movement and a practice.

Several interviewees specifically linked their memories with open access publishers PLOS and BioMed Central (see Section 5.2.2). My own association with PLOS – for those who knew about it – may be entangled with the emergence of such recollections, but they are notable nonetheless, and come from multiple national contexts. Jude[33] commented that use of "open" "…[is] a long-held thing in terms of

[31] PhD 1980s; interview: UK, Jan 2017.
[32] PhD 1990s; interview: US, Feb 2018.
[33] PhD 1970s; interview: Australia, Jan 2017.

open access journals, I mean I remember the start of the PLOS journals…". Adam's[34] earliest memories of "open" included a "young hotshot" and his students, who would go around wearing t-shirts supporting PLOS, having been trained in California where the publisher originated. Ernie[35] also remembered open access movements in California as the origin of "open". He recalled a crisis over Elsevier's bundled subscription charges to the University of California, around which:

> *"…lot of [people] became real activists for saying, 'Well we think this should be open, publishing should be open", but I think it was this community of people, primarily in the Bay area, who started pushing that. And because they were very influential, very powerful, very good scientists, they got noticed, and then other people joined in with the idea…"*

Amongst these interviewees there was a consistent understanding of "open" as a phenomenon that arose in a particular form (open access biomedical journals), at a particular time and place (California is prominent), and involved particular forms of activism (by influential scientists).

More recently trained scientists tended not to recount specific publisher- and activist-associated histories in their earliest memories of "open". Instead, there was an association in the UK context with top-down open access policies from the Wellcome Trust, the European Research Council, and REF (HEFCE 2014). Interviewees trained from the 2000s onwards often recalled encountering "open" at the start of their PhDs. Other "open"-related ideas tended to accompany open access in these more recent recollections. Julia[36] said of "open" terminology:

> *"I feel like it's more commonly used now than it was maybe 10 years ago, because I feel like there's a link with more recent developments. So obviously things like open access publishing, and all of the furore surrounding public money funding science that is then not open access, and things like DORA[37], this comes more in the community consciousness I think. And also there's various platforms where these*

[34] PhD 1980s; interview: Australia, Jan 2017.
[35] PhD 1970s; interview: UK, Oct 2017.
[36] PhD 2000s, interview: UK, Jun 2018.
[37] DORA is the 2013 San Francisco Declaration on Research Assessment (https://sfdora.org/read), that recommends the assessment of research based on its own merits rather than journal reputations or metrics, and frames online research publication as an opportunity for better assessment.

things are promoted or highlighted as an issue, like ResearchGate, kind of social science platforms I guess. So I feel like it's more in the vernacular in the last…but definitely related to internet-based sharing of science, yeah, and the preprint revolution and this kind of thing."

These responses suggested that from approximately the 2000s, scientists began encountering "open" through a wider variety of channels including policy requirements and newer online initiatives and platforms (e.g. DORA, ResearchGate, preprints) that increasingly encompass a wider range of "open" objectives. Lara[38], a PhD student, spoke of her encounter with "open" using the more recent, integrated term "open science". She recalled learning about it in 2017 by listening to the podcast *Everything Hertz* (Quintana and Heathers 2020), which features discussion of meta-science – the science of science – by two scientists who themselves trained in the 2010s. Encounters with open science movements via the social, platform-based web – especially Twitter – were mentioned more by scientists trained since the 2000s. Overall, scientists' memories of "open" tended to follow a pattern in which open access was constructed as a forerunner, but its history as a grassroots movement – and forerunner status – were increasingly obscured by the passage of time, as open science movements have broadened and acquired a top-down dynamic.

Not all interviewees linked their earliest memories of "open" with primarily open access or subsequent open science movements, however. For several – both scientists and advocates, across career generations – open source software was the forerunner. Elliot[39], who trained in the 2010s, spoke of being influenced by his older brother:

> *"…he was really into this kind of open source software movement [in the] very, very early days. I remember him getting copies of Linux on loads of floppy disks and stuff when I was really young. So he kind of indoctrinated me quite early into it."*

Elliot then described witnessing a movement of computer scientists with "open" principles into biology in the "mid to late 2000s" in association with a new generation of genome sequencing technology: "…that kind of seeped over into the science you were doing as well". This reflects a linguistic association between open software and open science (see Section 4.10) and evidences an early and ongoing flow of open

[38] Mid-PhD; interview: UK, Dec 2018.
[39] PhD 2010s; interview: UK, Jan 2019.

software enthusiasts and experts into scientific communities, shaping cultures of scientific openness. Mat[40] spoke explicitly of the analogy between open source and open science as the foundation for his laboratory's drug discovery approach (see Section 7.8.1, 7.9).

Some scientist interviewees distinguished the appearance of explicitly "open" language from an implicit (cultural or behavioural) openness in science. This latter openness was depicted as long-held presence or absence. Olivia[41], for example, commented that "the *implication* [of open/openness] has always been there. But I think the term, as openness in science, is quite a new thing", and Cedric[42] described the "jargon" as new, but the "notion" of openness as "part of what scientists do, the ethics of science [...] we took it with our mother's milk, as it were". Nicole[43] and Andrea[44], in contrast, implied an absence of openness in their communities. Many of these answers were part of an emerging category that I have identified as "interpersonal openness" and analysed in Chapter 8. Framings of openness as an enduring or essential to science also indicate an empirical relevance to the "traditional" narrative identified in Chapter 3.

6.4 Money, money, money: a financial framing of open access

Dominant among framings of open access by scientist interviewees was an understanding that it costs money, and is expensive: a financial framing. Interviewees in both Australia and the UK both commonly expressed this understanding, as did interviewees of all career generations. This is notable as from a policy perspective, as there are recommended routes to open access that are cost-free to researchers – the next section discusses the implication for "gold" and "green" routes. Molly[45], the PhD

[40] PhD 1990s; interview: UK, Apr 2019.
[41] Mid-PhD; interview: UK, Jul 2018.
[42] PhD 1960s; interview: UK, Feb 2018.
[43] Mid-PhD; interview: UK, Oct 2017.
[44] Mid-PhD; interview: UK, Oct 2018.
[45] Mid-PhD; interview: UK, Nov 2018.

student featured in the opening vignette, portrayed open access as desirable but prohibitively expensive. Cedric[46], an emeritus professor who remains active in research, offered a similar view: "I know about the issue [open access], but I'm not…I'm only affected by it insofar as, because I'm self-funded, I cannot pay the huge fees that publishers demand, so everything I produce, by and large, is not open access". Both Molly and Cedric positioned open access as out of their reach, and therefore of only theoretical relevance to them, owing to a lack of funding in these earlier and later stages, respectively, of their careers. Jason[47] described open access in association with publishing business models: "either I'm paying [open access] or you're paying [subscription access]". This financial framing was prevalent, and separated interviewees' responses according to their means and relationship with money.

Some interviewees portrayed open access as an expense they could habitually afford, either because of the financial privilege of their research group, or because of the block grants funded by UK research councils and charities including the Wellcome Trust (see Section 5.2.5). Kate[48] had an approach defined by the block grants system, as did many other UK interviewees:

> *"…at the moment that's [open access is] the easy part. The research councils require it. So at the moment we've published in whatever journal we've wanted to and then we've turned to the university and asked them to pay so that the article is open access and we've managed to achieve that so far."*

This idea that open access can almost always be bought, if one is able and willing to pay the price, suggested a widespread awareness that many normally "closed" subscription journals can be paid to arrange open access on a per-article basis – the so-called "hybrid" model of open access (see Section 5.2.1). Mark[49], who runs what he described as a wealthy laboratory, portrayed the ease with which he can now publish open access in a prestigious subscription journal. He contrasted the radical politics of

---

46 PhD 1960s; interview: UK, Feb 2018.
47 Mid-PhD; interview: UK, Nov 2018.
48 PhD 2000s; interview: UK, Jun 2018.
49 PhD 1980s; interview: UK, Feb 2017.

the open access movement represented by PLOS – an entirely open access publisher – to the practice of open access he is able to adopt through a hybrid model:

> *"Especially now - when PLOS was getting started it was radical and fantastic and very exciting, and everything's sort of moved on a long way, since then. So...yes, because, you know, we can publish in* Nature *and you pay extra five thousand dollars or whatever and it becomes open access, and that's great [laughs]."*

The "open" versus "closed" politics of journals and publishers could thus become insignificant in moneyed settings – including the UK's block grant system – which preserve the ability to publish in one's journal of choice, and reduce open access to a single – more or less expensive – transaction.

Alongside the question of affordability, a financial framing of open access added a framing of conditionality or precarity. For instance, in Kate's response above, there is tentativeness: she uses the language "at the moment" and "so far", implying that funding that may not always be available. This implication was also present on the administration side of block grants for open access in the UK. Danny Kingsley[50], who leads open advocacy at a different UK university to Kate's, described the complicated and precarious accounting required to administer open access block grants from different funders, under different and sometimes continually uncertain conditions:

> *"So we now have three different sets of policies that are all slightly different to each other, happen at different amounts of time, affect different people...may or may not have funding associated with them [...] last year, the RCUK said: transition period is going to finish, 31st of March, full-stop, that's it, you cannot pay anything after that. That caused a major problem for us [...] So, they then...said okay, well we'll continue the transition, that we're not now calling a transition period, for two more years..."*

She described how, towards the end of a budgetary period, it was sometimes necessary to change the conditions of funding availability: restricting funds available for open access in hybrid journals in particular. Scientists in the UK therefore – depending on their funder, institution, and publishing practices – sometimes experience inconsistent funding of open access.

---

[50] Head of Scholarly Communication, University of Cambridge; interview: UK, Feb 2018.

The financial framing of open access was associated with conditionality. Lena[51] spoke of "always" making her work open access, but made it clear that it was the provision of specific funding that enabled this, despite her lab's fortunate financial position. This was due to the degree of expense involved:

> *"...even for a lab that is well-funded, the amount that you need to pay for publications is not small. So it's a whole like, US or Japan type of conference that is cheaper than what you need to pay for open access. Yeah, so it's...this is big, right."*

Her financial framing placed open access in a cost versus benefit competition with a long-haul flight to a conference and challenged the value of the former. She also described the cost of four or five open access papers – "twenty-five thousand pounds" – as an amount that would otherwise keep a laboratory in her field running for a year: a major rather than incidental expense. Steve[52] , similarly, was "very pleased to publish our work open access", but disclaimed responsibility for the cost: "I don't think I have ever paid to have open access when it wasn't free to me. Because it is quite expensive...". Like Lena, Steve imagined the weighing up of open access costs in his laboratory budget: "If I've got that thousand pounds in a grant, I'd much rather spend it on laboratory consumables...or indeed on the salary of the person doing the work, than on open access." Melanie[53] made enthusiastic remarks about open access, but again, her commitment was conditional on financial provision:

> *"...the only time I would ever say it [open access] would not be something I would do, is if it was money limited right? So now - it used to be that...they [funders] would give you publishing charges as part of a grant, and now they want that out of the grant, and they give the university a chunk of money [the block grant]. And so, if we didn't have money for it, it would be the only time I wouldn't do it."*

The conditionality of open access through a financial lens demonstrates its tenuous position in current scientific practice in the UK. Moreover, it shows that when scientists are asked to place a financial value on open access and weigh it against other research expenses, it is often considered less than essential.

[51] PhD 2000s; interview: UK, Apr 2018.
[52] PhD 1970s; interview: UK, Nov 2017.
[53] PhD 2000s; interview: UK, Jul 2018.

In the Australian context of this study, where specific funding for open access tended not to be available, Jude[54] spoke of a non-hypothetical case of budget compromise:

> *"Okay, this [journal name] paper [...] I didn't pay for open access, because I figure, most of the people who are interested in that field will have a good university library, so what's the point? You know, [collaborator] shelling out a couple of thousand dollars to make it open access, because I figure most people that want it can get it."*

Jude's financial reckoning on a per-article basis shows that a lack of dedicated funding makes open access publishing a weightier decision. For those with research budgets – but not block grants for open access – paying for the latter is not necessarily out of reach, but requires compromise, and is only "easy" at the highest level of laboratory wealth. This scenario also makes cost of open access publishing more visible and transparent. It is perhaps remarkable that some interviewees in Australia did claim to publish open access habitually; however, they were a very small minority.

Finally, a financial lens on open access conditioned particular emotional and attitudinal reactions, including scepticism, bemusement, suspicion, and resentment about the way money is used. Jason[55] expressed indignation at the idea of being required to pay a publisher for open access:

> *"You have to pay the publisher that money, and that could be thousands of pounds, which I find offensive [...] I do feel that all of this material should be open and unrestricted of course. But now I have to pay for it and I don't have a lot of money. But I have to pay for it, because I have to have a job. I want to get a job after this in academia and research, and I need to have publications to do that. And that's that whole publish or perish model."*

His anger is connected to a broader sense of injustice and vulnerability about his dependence on publications for career security; along with scepticism about academic publishing as whole: "I think the whole publication thing is a bit of a scam".

---

[54] PhD 1970s; interview: Australia, Jan 2017.
[55] Mid-PhD; interview: UK, Nov 2018.

Scientists in more secure positions in their careers also had objections to the movement of money around open access: there was a sense that unjust power dynamics were playing out behind the scenes, and that science and scientists were losing out to publishers and bureaucrats. David[56] positioned open access and the ability to "do the science" as a zero-sum game in budgetary terms, perhaps reflecting a wider sense of grievance and scarcity in relation to top-down governance:

> *"…it's extremely annoying in this country, open access drive when they basically force us to do it but don't give us the money to do it. [laughs] So it basically just reduces the amount of money available to do the science […] So basically it's just another way that they took money away from the science budget without appearing to cut the science budget."*

Ernie[57] stated that he always publishes open access using the "pot of money" [block grant] available through his UK funder. He nonetheless remarked, "I don't see necessarily why anybody should have to pay", and was suspicious that PLOS might "make a tonne of money" from "big publication fees". This suggests the construction of resentment about the cost of open access alongside broader resentments towards academic publishing. Julia[58] expressed a broader concern about academic publishing with amused exasperation:

> *"…ethically, [there is] something a bit tricky about people making even more money out of something being made open access, like demanding a large fee for that, when they're already… People outside of science think it's hilarious that you publish your work and you pay to do it [laughs] rather than the other way around…"*

Lena commented drily that "science publishing is just amazing" for similar reasons: charging fees to publish and to read; asking researchers to review for free; acquiring copyright. Ironically, objections like these have long been *motivators* for open access and other open science movements, but through a financial lens open access can stoke the same kind of resentment it was intended to dismantle. And through this lens, hybrid and subscription publishers with exceptionally high profit margins may not be distinguished by some scientists from non-profit open access publishers.

[56] PhD 1970s; interview: UK, Dec 2017.
[57] PhD 1970s; interview: UK, Oct 2017.
[58] PhD 2000s, interview: UK, Jun 2018.

6.5 Going for "gold": journal-centred views of open access

Open access is conceptually and practically differentiated into categories in open advocacy and policy settings: journal-mediated access ("gold") and repository-mediated access ("green") are two of the most commonly recognised such categories (see Section 5.2.1). "Gold" and "green" each have a complex history and politics within open access movements; feature differently in institutional policies; are enabled by different economic models; and in practice require different sets of practices and interactions. The framing of open access by many scientists through a financial lens – it costs money, and is expensive – aligns with another framing, which is of open access as a "gold", journal-mediated practice. This is because the "gold" route often, but not always, requires payment by the scientist, their funder, or institution; whereas "green" is cost-free to researchers. The "green" route was invisible or marginalised in most interviewees' accounts, which is significant given that the UK's rigorous REF2021 open access policy centres a repository-mediated "green" route (HEFCE 2014), despite an early and continuing emphasis on "gold" by funders including the Wellcome Trust (see Section 5.2.5). "Green" is also the policy approach favoured by the major research councils in Australia. There is therefore a noticeable disjunction between many scientists' constructions of open access, and the modes of open access that are available, and perhaps required, in their research context.

The main indication that scientist interviewees held "gold"-centred views was the dominant financial framing of open access, including familiarity with the hybrid journal concept. Secondarily, language associated with journal publication rather than repositories was used to frame the topic. Jude, for example, raised the topic with this comment: "Okay, probably open access journals would be one [kind of openness]", and Lena saw open access as being "all about the journals". Talk of publications and journals does not indicate an exclusive focus on "gold": "green" provides repository access to published articles. However, combined with a financial framing, this way of talking does indicate that journals and the academic publishing industry – not institutional or disciplinary article repositories – are at the forefront of scientists' minds in relation to open access.

The Australian context was noteworthy, as – despite national funder policies requiring "green", a similar policy at university level, and a lack of specific funds to support "gold" – interviewees still primarily spoke of open access as a practice associated with journal choice and payment. None of the interviewees in Australia mentioned, by name or implication, having followed a "green" repository route. Adam[59] did mention the terminology, however:

> *"...green is when there is a version of the paper that is almost the same as, but maybe the non-typeset version or whatever of the paper in a, typically an institutional repository and [university] has such a thing, but it's not the culture here to put papers on that, I don't know anyone who's put papers on the [university] one, so I don't think...I like the green idea but I don't think it actually works."*

This explicitly confirms a journal-centred view of open access which seemed implicitly widespread amongst interviewees in Australia, and suggested a lack of engagement with national funders' recommended route to open access. This emphasis on "gold" in a "green"-centred policy environment is striking and perhaps indicates an influence of global open access discourses and influences over national and local governance and infrastructures.

Repository-mediated open access was mentioned more by UK interviewees, perhaps because this was a larger group. However, it tended to be a secondary framing, for example emerging through follow-up questions, suggesting that scientists may indeed be familiar with "green" without considering it a meaningful mode of open access. Yvonne[60], who said that open access "usually means paying the open access charges, for the journal", commented as follows when asked about deposition in the university repository:

> *"So we do that - we're obliged to do that [...] For the next REF [...] You have to have put your publication in a public repository within three months of acceptance. And so what we're told now is that we have to put that final accepted version in the repository. Some of the journals won't let it be open access. So some of the journals will not let you*

[59] PhD 1980s; interview: Australia, Jan 2017.
[60] PhD 1970s; interview: UK, Nov 2017.

> *actually make that accessible to the public, but it has to be there in the repository. So yeah, we do do that."*

Yvonne's response showed that she was knowledgeable about the university and REF arrangements for arranging "green" repository access, whilst simultaneously constructing open access as a "gold" journal-centred practice. Her comment suggested that this may partially be due to the mixed levels of access – not always "open" – that repositories allow.

Other UK interviewees constructed "green" as a form of open access, but assigned it a lower or less desirable status than "gold". For instance, Erin[61] told me that the two top journals in her discipline are not open access unless you pay expensive fees, so she "...will publish it there [in a top journal] and make it green open access. But not pay the fee to make it gold". She then commented:

> *"But when I can, you know, intersperse that with things that are actually gold open access. And moving more towards that. So I think we're trying to publish more in PeerJ so that people in our discipline will know that you can find stuff in PeerJ for example."*

Her use of "actually" in this statement suggested that "green" repository-mediated open access is a less realised or valuable form of open access. Luke[62] similarly constructed "green" as a second-class option: he described it as a "baseline" in comparison to "the highest open access you can...gold". In his account, "green" is a work-around:

> *"...I think [current university] wants you to make it as open access as possible, but they have ways that they get around it. So like, I give them the copy of my work before it's copyedited, and that gets around some of the open access issues. They can release our version - it's just not the published version of your work."*

This account suggests that repository-mediated access may be viewed by some as a compromised version of open access – or not open access at all, but an imitation that satisfies certain technical requirements.

[61] PhD 2000s; interview: UK, Aug 2018.
[62] PhD 2010s; interview: UK, Nov 2018.

There are many possible reasons for the secondary placement or backgrounding of "green" open access. In some cases there may be a lack of awareness, as suggested by the contention that a lack of funding unequivocally places open access out of reach. However, even in these cases it may be that "green" is not considered an adequate form of "open access", implicitly or consciously. As Luke identifies above, a "green" open access article is not the final, polished, typeset ("final published form" in some policy language: Wellcome Trust n.d.) unless the publisher allows use of their final version – it is in that sense a "way to get around" certain publishers' rules. It may be that this peer-reviewed but slightly less finalised article is considered *less of an article,* or *less open*. Given the importance that interviewees attributed to the reputation and audience of their chosen journal, it may make a critical difference not to have the article openly accessible from the journal website, and for it not to be visibly authorised by the journal's branding. And given the importance of publishing as a threshold to authorised and attributed knowledge – constructed as such throughout the interviews – an article that appears slightly less than published, even though it *is* published in almost identical form, may be felt to lack crucial value. It also emerged in a subsequent section (6.7) the centring of "gold" could relate to scientists' agential connection with the practice of journal publication; repository deposition may feel like more work, and less under scientists' control. Thus there may be multiple overlapping explanations for the backgrounding of "green", but many relate to the cultural importance of publication in academic science.

Alternatively, scientists may not consider "green" sufficiently *open*. Erin's support of the open access journal *PeerJ*, above, and the surrounding context, suggests that for her, "actual" open access means publishing in a fully open access journal, rather than choosing a subscription journal with open access via hybrid or "green" routes. Her valuing of "gold" implies a relationship with the politics of open access movements in addition to open access as a technical state. A few other interviewees voiced similar considerations, but it remained rare for the open access status of a journal to be considered above its audience, reputation, or impact factor in publication decision-making (see next section). Immediacy of access via the "gold" route did seem to be part of its appeal: Steve thought that his work would not get "such an airing" without open access via journals, and Melanie commented, "…I love that [open access] because everyone has access the day it comes online". The value of licensing for re-use, which "gold" open access can more easily allow and is strongly upheld in dominant open science movements (see Section 5.2), was not raised by interviewees except those involved in open advocacy, and seemed not to be influential to the backgrounding of "green". This was a clear point of disjunction between pro-"open" and broader scientific communities.

There was a notable exception to the centring of "gold". Lara[63] was being trained in a lab with a strong commitment to many kinds of "open" practice, and she had become committed to such practices herself. When it came to open access, however, she expressed some of the scepticism and resentment that was characteristic of financial framings:

> *"Gold open access is where you pay a lot of money to the journal to publish your manuscript. So, I don't know. I think I have a lot of opinions on the publishing industry. I don't know. I think these journals are making a lot of money out of something that should be just available for free. So that's why I'm obviously a really big fan of green open access. So just like making your manuscripts available online."*

Unlike other interviewees who expressed such resentment however, she distinguished this as a problem of "gold", and framed "green" as a cost-free and open solution. Her mechanism of choice was preprints: making articles openly accessible online in a preprint repository. Other interviewees too, mostly of a more recently trained generation, and were adopting preprints as a means to make their work visible and attributable prior to publication. This suggested a trend towards the co-option of preprints as a form of open access that combines the cost-free element of "green" with the immediate visibility, accessibility, sometimes open licensing, and even journal-platform brand recognition of "gold" (the preprint server *bioRxiv* was commonly mentioned by interviewees). Preprints might be considered indistinguishable from "green" open access, especially if an editorially accepted version is made available: this is the implication of Lara's account. However, preprints are not necessarily considered

[63] Mid-PhD; interview: UK, Dec 2018.

to be open access in policy terms. For example, the REF2021 open access policy in the UK only counts preprints if they are both "accepted for publication" and uploaded prior to publication; and, *bioRxiv* discourages posting of accepted manuscripts (ASAPbio 2019). Thus, a generational trend towards preprints arguably represents an embrace of open access – much desired, theoretically, by policymakers – but in a form not recognised within policy and compliance framings (see also Section 6.7).

6.6 Quality, reputational, and "impact" framings of open access

A further significant facet of open access in scientists' constructions is related to quality, reputation, and "impact". Some interviewees portrayed a systematic link between one or more of these characteristics and open access; others did not. An understanding of open access as costly and journal-mediated – shown above – shaped these portrayals. One type of response, from a minority of interviewees, associated open access with concerns about quality in the production of science. For some, this was a concern about the editorial practices of open access journals. Jude commented that "one of the big dangers" of open access was that "a lot of the editors don't have intimate knowledge of the field": he associated this sense of unease with receiving invitations from open access journals to review papers from outside his field, and conversely seeing papers in his field of expertise "being reviewed [at an open access journal] by, you know, associate handling editor…someone I've never heard of". His contrast case is with society journals whose editors:

> *"…are normally in their fifties or sixties so they've got a lot of experience and they know who hates who, who might be…best able to give a fair review. I mean obviously by their choice of referees they can bias the outcome, but…"*

Jude's view is thus shaped by a journal-centred view in which the reputation and trustworthiness of open access is tarnished by particular observations of editorial practice. These stem from a lack of community-embedded leadership, seen as important to quality refereeing, and in turn the quality of published knowledge.

Similarly, when I raised the topic of open access with Jacqueline[64], she told me that she was most concerned about:

> *"...making that sure that there has been appropriate peer review. And so if you're talking about journals that are just going to publish...anything, I'm not in favour of that."*

Like Jude, she was preoccupied with an associated loss of "human judgement" and the use of "computer-generated list[s] of reviewers". These are not direct, inevitable or systematic consequences of open access; nor was concern about them widespread in the interviews. But Jacqueline and Jude – trained in the 1970s and 80s – expressed a poignant worry about loss of community that had become, for them, a connotation of open access. "I find it really scary if we're going to move to a situation where we don't trust our colleagues", Jacqueline told me.

Although Jude and Jacqueline's social-epistemic quality concerns were not held by the majority, they did have echoes in other responses, including of more recently trained interviewees. Jason, a PhD student, keenly advocated open access whilst adding:

> *"... if you are making free and open access to publications, it needs to be quality publications. It can't just be anything that someone put out. I still believe in peer review."*

Jason depicted an association between open access and poor quality – the idea of publishing "anything" – as a tendency to guard against. When I asked him about repository-mediated open access, he again defended quality:

> *"I'm for that as long as it goes through quality peer review [...] There are journals out there that you can submit to without peer review and that stuff gets published. And that is a huge problem."*

This concern as well as those from Jude and Jacqueline is likely influenced by the phenomenon of "predatory publishing", originally characterised by Beall (2012) as a corruption of author-pays models of open access, in which researchers are spammed and "duped" into paying for publication with poor quality or non-existent editorial quality control. The phenomenon was later acknowledged to be broader, incorporating misleading, indiscriminate, and low quality journal practices that are

[64] PhD 1980s; interview: Australia, Jan 2017.

not necessarily linked with open access, and do not necessarily distinguish predatory journals from those deemed "legitimate" (Cobey et al. 2018:3, 15).

While most interviewees did not voice open access quality concerns, the spectre of predatory publishing continued to haunt open access, even where the two were distinguished. Mark[65] phrased it as follows:

> *"...there are some very unfortunate, pernicious, secondary - unintended, secondary consequences from open access publishing. Which has led to this spawning of a whole series of low quality predatory journals. Which I think has, yeah, it's not the open access movement's fault at all [...] but it's terrible thing that's happened in scientific publishing, a really, really bad thing."*

Ben[66] expressed a similar sentiment, and observed that although "high-end academic institutions" tend not to suffer from the confidence tricks of predatory publishers, researchers in less privileged environments can be systematically affected. Notable in Mark's and Ben's accounts was a sense of profound damage to scientific and publishing systems. It is important to acknowledge this as a meaningful loss or concern for researchers that shapes understandings of open access, including for those who express support of open access practices and movements.

Quality framings were entwined with reputational framings of open access that concerned the prestige or "impact" of published research (where "impact" tends to refer to the journal impact factor, see Section 4.6). Where only a few interviewees expressed explicit quality concerns about open access, many more associated it with lower "impact" work – again, a journal-centred framing. Open science movements have long made the argument that reputational shortcuts to research assessment – and in particular, the journal impact factor – are problematic as they do not meaningfully reflect the value of individual research articles: "We conclude that science is currently rated by a process that is itself unscientific, subjective, and secretive" (The PLoS Medicine Editors 2006:0707).

[65] PhD 1980s; interview: UK, Feb 2017.
[66] PhD 1990s; interview: UK, Jan 2018.

However – unsurprisingly – my interviews showed that journal reputations and impact factors were still commonly used to make publication decisions, whether in the pursuit of prestige, quality, or a mixture of the two. Steve expressed a common stance, which to priority a journal's reputation over its openness: "…the open access option has not been the decider […] - it's the good old impact factor!" Steve otherwise expressed wholehearted support for and adherence to open access practices, as did most interviewees who nonetheless considered journal reputation – including community factors like audience – before the possibility of open access. Reputational orientations were not necessarily the preserve of senior scientists: Molly, a PhD student, associated open access journals with low impact – and poor career prospects – at the start of this chapter. The conclusion that high-impact publication is mutually exclusive with open access again reflected a journal-centred view, and one in which hybrid open access was either unaffordable or not considered a true form of open access. Julia valued publication at fully open access – as opposed to hybrid – journals, but observed a structure to the availability of such options in relation to prestige:

> *"So there's a kind of mid-range level where there's several really good options for open access publishing. So for example one of the last papers I published was in* eLife*, and that is a kind of mid-range impact factor. […] The issue is if you go a bit higher up, then you start to lose such good options for open access I think, which is when I would consider to go to paywall journal and make it open access with the funder."*

Thus, even as a keen supporter of open access, Julia associated open access with compromises when it came to the most prestigious forms of publication. An alignment between prestige and open access did sometimes occur, in disciplinary context. Gavin[67] told me that "a lot of the open access journals are some of the top journals in our discipline [computational biology]". Even in this case it appeared that the reputations of such journals drove loyalty, with open access constituting a secondary part of that reputation. As Michelle[68] put it: "I have published in open access journals, but it wasn't because they were open access journals."

[67] PhD 1990s; interview: Australia, Jan 2017.
[68] PhD 1980s; interview: Australia, Jan 2017.

Some interviewees did not question the legitimacy of impact factors and other reputational shortcuts; others took care to distance themselves from such approaches, whilst complying with them nonetheless in practical terms. Ian[69] commented:

> *"I would like to publish everything in* PLOS ONE*, or* F1000 Research *or something [fully open access journals] [...] But, it's not fair on young people [his trainees/co-authors] to do that, because they need to get the kudos from publishing in good journals."*

Adam expressed almost exactly the same desire and justification, describing this approach as "operat[ing] knowing how the system works whilst railing against the system at the same time", with associated with feelings of tension and hypocrisy. Accounts like these suggested that reputational factors often influenced publishing behaviours not because they were intrinsically valued, but due to a conviction that such factors are systematically upheld by *others* in professional assessment contexts. This observation aligns with a with the findings of a survey by Niles and colleagues (2020) in a North American context, who found a disconnect between what academics most value when choosing a journal – readership – and what they expected their peers and evaluation committees to value – prestige-related metrics.

The pursuit of high-impact publication by senior scientists on behalf of junior colleagues – which often came at the expense of open access, in journal-centred and financially constricted contexts – seemed to be a pattern. Adam also observed it as such: he felt sure that some senior scientists used such reasoning as a pretext to advance their own reputations, whilst others genuinely sought to assist their juniors. However, it was not clear that junior scientists were allowed agency to form or pursue their own dissemination preferences. When I asked Rory[70], a postdoc, whether he was in the driving seat when it came to publication decisions, he said: "No. I'm normally sitting in the back. Someone tells me, I think we should go here, go there." Thus I observed a potential for reputational systems – often negatively associated with open access – to be maintained by a cycle of expectations, including by senior scientists

[69] PhD 1980s; interview: Australia, Sep 2017.
[70] PhD 2010s; interview: UK, Jan 2019.

notionally or actually protecting their juniors. Meanwhile, junior scientists may not be given the authority to instigate alternative systems – or may already have internalised the same expectations as their seniors.

These generational dynamics are notable because of evidence that some scientists – those more recently trained– configure open access differently in relation to reputation. Several interviewees did not make any association between open access and quality, reputation, or “impact”. This tended to occur where multiple routes to open access were viewed as possible and legitimate: for instance if hybrid, “green”, or preprint-mediated open access were felt to be salient, accepted, and affordable forms of open access. In other cases, “open” principles themselves actively drove journal choice, and the presence of a “paywall” or expensive hybrid fees was considered harmful to a journal’s reputation. For instance, Olivia was determined to publish the first paper from her PhD in an open access journal: “…I’m really pushing to get it in *PLOS ONE* because it’s a journal that’s known for its open values”. She framed open access publication as aspirational in the way that other scientists frame high-impact publication – strikingly so, in comparison with accounts above from senior scientists eschewing publication in *PLOS ONE* for the sake of their students. Erin also actively pursued open access (see Section 6.5 above) and, moreover, consciously sought to raise the reputation of the open access journal *PeerJ* in her disciplinary community, by publishing there to increase its visibility and relevance to her colleagues. It was not clear how widely valued openness was as a reputational feature of journals or platforms: it seemed to be emerging rather than entrenched like “impact”. However, the demographics of the group upholding this feature suggest its reputational pull may increase over time. Thus, the adoption of open access was only partially driven by the abandonment of journal reputations as markers of quality publication; the growth of alternative brands (e.g. *bioRxiv*) and reputations representing open qualities also seemed significant.

6.7 Compliance framings: open access as distanced and bureaucratic

A final salient construction of open access arose from its position in top-down institutional policy regimes. Most notably, UK interviewees tended to practise open access via administrative mediation that prioritises policy compliance. This featured particularly in the accounts of research group leaders, and was likely connected with their employment at UK universities that have significant open access budgets and teams of professional services staff. In Australia, where funding for open access was not routine, interviewees did not speak of administrative mediation for open access, but they did report bureaucratic monitoring from funders. Nonetheless, it was a prevailing experience of interviewees in this study – especially in the UK – that in everyday life, open access took on the form of an administrative task or interaction; a responsibility displaced to others that still requires passing attention; a tick-box task that is also a black box, its contents relatively opaque to researchers.

When interviewees in the UK felt that the administrative process worked smoothly, they spoke in positive terms about their experience of open access. Ernie, for example, commented:

> *"So sometimes it's quite - sometimes it is a drag [the open access process]. But in general, it - actually, we're blessed, it goes pretty smoothly, and the guy [library staffer name] who's in charge of doing it, he's great. I hope he retires after I do [laughs]."*

Ernie's experience of open access was strongly influenced, perhaps even defined by his interactions with a particular staff member who performed the involved tasks of open access, which could include: knowing precise definitions of "open access" in licensing terms; ensuring access is achieved in a way that meets specific funder policy requirements; and managing any financial transaction required. Ernie described an occasion on which "it took poor [staffer name] about 20 or 30 emails back and forth" to negotiate payment for open access at his chosen journal, which had only recently introduced an "open" option. Scientists in Ernie's position thus experienced open access as relatively effortless – "all I have to do is…make sure that that option gets ticked" – with moments of friction from which they are partially shielded. Simultaneously, they are shielded from practical knowledge of open access, and an associated agential connection: a sense of responsibility and control. Steve, for instance, stated that open access is "quite expensive", but added: "I'm not 100% sure [what it costs], because I never see the bill…". Scientists in this UK context are strictly required to practise open access – this is at least part of the reason for administrative mediation – which compounds a lack of agency, as the decision is not theirs, nor are they required to carry it out proactively. In many cases, this disconnection was experienced as a convenience, allowing open access to become assumed and unremarkable: according to Henry[71], "…we just have to publish open access so it's not really […] a topic to discuss" and similarly Andrea said "there's no discussion about it".

The Australian context was markedly different because open access requirements were less stringent, specific funding was generally not available, and administrative mediation systems were not mentioned. Amongst the small interviewee sample – mostly at one university – a pressure to practise open access was not in evidence: only two of eight interviewees (Gavin and Jenny[72]) spoke of habitually pursuing open access, and both described exceptions to this practice. Moreover, both seemed to pursue open access for reasons unrelated to policy rules: Gavin mentioned the "top" quality of open journals in his field, and Jenny mentioned the benefits of immediate, wide accessibility – particularly for her contacts in India and Thailand. Australian interviewees thus seemed to feel greater agency in relation to open access, through which a minority became personally attached to the practice. They did not, however, have financial freedom to pursue the "gold" route, which was most salient in their accounts, and their overall prioritisation of open access was low.

Curiously, there was at least some awareness that the national research councils in Australia required open access – and recommended a repository route – but the associated compliance procedures were viewed as ineffectual. Adam told me:

[71] PhD 2000s; interview: UK, Dec 2017.
[72] PhD 1970s; interview: Australia, Nov 2017.

> *"...we're asked to keep it [database] up to date, put papers in there and one of the questions...so what's the title of your paper, date of publication, page numbers etc., which grant did you use to fund this work, and is it logged on a...open access thing. And for most of them the answer is no, and as best I know there's no pushback on that, they don't - I've never, I've not heard of anyone being pinged for not having their stuff on an open access site. Even though I'm quite sure most don't."*

Ian portrayed a similar meaninglessness to bureaucratic monitoring of open access, especially in the context of time pressure:

> *"...so [funder], for example, requires you to deposit your publications in a, in an open place, I went - last grant I put in, I was up against a deadline, and I came to that question, I just ticked all the boxes. I have no idea whether they're openly accessible or not - nobody looks at that..."*

This suggested that both extensive, effective policies and bureaucracies – as well as those perceived as minimal and ineffectual – struggle to connect scientists meaningfully with open access. Whether through a sense of unaccountability and futility, or a convenient displacement of personal labour and responsibility, scientists were at an arm's length from open access – at least, as defined by policy compliance. Through a compliance lens, open access was part of an administrative workload. This may have led scientists to associate it with their feelings about administrative work and research bureaucracies generally.

Scientists were not inevitably disconnected from open access as a policy compliance activity. Where they did actively seek to comply, the bureaucratic context shaped levels and experiences of engagement. In the UK, Yvonne had a conscientious knowledge of open access options and routes to compliance (see Section 6.5), but this appeared to reflect resigned rather than enthusiastic engagement: "We're obliged to do it, so we do it". Miriam[73], meanwhile, had actively brought members of her research community together to petition a top journal in her field to comply with the REF2021 standard for open access. However, her construction of open access remained dominated by a sense of distance and exasperation, created not only by the administrative load of such engagements – which she described as a "pain in the bum"

[73] PhD 2000s; interview: UK, Jan 2018.

– but also by a sense that compliance requirements were obscure, inflexible, and punitive, even for those like herself who were proactively committed to open access.

In particular, she referred to an occasion on which she "got into trouble" for failing to submit an open access article she had co-authored to a university repository. The article was openly accessible through a journal but not registered through the "green" mechanism by which the university monitored open access. She found this experience to be intensely frustrating:

> *"...you're just thinking...ugh, lads...just go out there, see, can you access it, you don't need to have to have it documented in that exact format [...] people can search it on the website and find it there, it's this open access paper that's been written by [Miriam], and look she's one of 30 authors, and she's right in the middle, so it's really not an important part of her work! But isn't she awful, because we can't count that in our [laughs]..."*

Erin, who was also committed to practising open access, described an administrative struggle to arrange payment at her chosen open access journal, *PeerJ*, because it operates on a membership fee rather than fee-per-article basis:

> *"...the university or the granting agency would rather you pay single fees that cost way more money than to actually just say, can I buy a membership that – yes – will last longer than a duration of my grant, but will be way more economically feasible and a better use of grant money. And that seems to be – even though it's totally logical it seems to be a real difficulty."*

Experiences like these suggested that bureaucratic procedures associated with policy compliance and financial management can create distance from a wholehearted embrace of open access – associating it emotionally with limited agency and frustration – even, and perhaps especially, where scientists have proactively sought to practise it. Even Ernie, who described a positive experience of open access made easy by delegation, initially labelled it as a "drag", alluded to the "pain" of arranging open access for papers co-authored across institutions, and lamented his funder's approval of "certain types of open access license" and not others due to a "technicality". The particular qualities of systems and individuals managing open access thus have a powerful role in constructing scientists' relationships with this form of openness.

The experience of agency may be one reason for salience of "gold" over "green" routes to open access for interviewees: the former is a publication practice with which they have already traditionally engaged, and the latter may feel like new administrative work primarily associated with compliance. The appeal of preprints may also be associated with agency and an escape from bureaucracy. Julia, for instance, happened to be submitting to *bioRxiv* for the first time on the day of our interview, and told me:

> *"I'm quite excited actually [laughs]. It almost feels like you get a paper for free, right, without the pain of like - it's going to have to go to an editor sooner or later. But you almost feel like I'm going to publish a paper today [laughs] without having to wait painfully for the review process, and the editorial process."*

Scientists are accustomed to the bureaucracy of publishing, so the removal of these layers combined with the separation of preprints from compliance framings of open access imbue the practice with a sense of freedom.

In the UK context, scientists' experiences with bureaucracy around open access – both of convenience and frustration – are given context by interviews with UK university staff working to advocate, implement, and facilitate open access policies. Individuals in these roles portrayed a complicated policy landscape: Paul[74], Pro-Vice-Provost at UCL and advocate of open science at university leadership level, told me that there were "over 400" open access policies across Europe (see also Danny Kingsley's account in Section 6.4 of administering multiple funder policies). Laura[75], who was working in administrative facilitation of open access, described compliance with open access policies as a complicated and confusing process which she would not expect researchers to navigate themselves. The procedural intricacy of compliance was evident in Laura's need to refer to specialised digital tools and databases – synonymous with the UK organisation SHERPA (Jisc n.d.) that produces them – in order to navigate, for example, the exact forms of open access required by each funder or offered by each journal. These tools are publicly available but the only scientist interviewee to mention them was Thomas[76], who had existing expertise and

[74] Interview: UK, Feb 2018.
[75] Interview: UK, Jul 2018.
[76] PhD 1990s; interview: UK, Oct 2017.

motivation having turned his career towards research and advocacy of open practices. Compliance with open access policy thus manifested as a practice of such esoteric complexity that it required mediation by both human and machine specialists.

The development of bureaucracies to implement open access compliance reflects not only its complexity but a sense that scientists will not engage proactively. Laura described mediation as necessary partly because open access is "the second stage, or not the priority" for researchers. In advocates' accounts, there was tension between two reactions to scientists' disengagement. On one hand, disengagement was viewed as increasingly irrelevant given that systems of compliance successfully implement open access without necessitating scientists' engagement. On the other hand, it was viewed as part of a cultural problem in which scientists are complicit, and which calls for fundamental culture change. Danny Kingsley[77] described her own struggle with this tension. She portrayed with regret her view that open access was failing to "win hearts and minds"[78] in the research community:

> *"…some people get it [open access]…some people do it because they think, oh well, it's easy enough, I don't really care. There are others who do it really reluctantly, and there are others again who have actually refused to, completely refused to engage. Full stop, they just go: this doesn't apply to me […] my feeling is if all of this stuff [policies] went away tomorrow… we probably would have actually sort of only incrementally increased engagement. So, the voluntary engagement before this was around the 10 to 15% of material was being made open access voluntarily. I would say if we took it all away, we'd be lucky to be hitting 25%. Really lucky. So I think that the researcher attitude to openness is reluctance…"*

Her account was striking in its portrayal of widespread open access as reliant on policy compliance systems rather than scientists who value the practice. She critiqued the "compliance line" as a narrow framing that prioritises quantitative reporting of open access over any qualitative purpose:

> *"The original goals of the policy were not that we'd have 81% compliance, presumably it was […] that we wanted to share our*

[77] Head of Scholarly Communication, University of Cambridge; interview: UK, Feb 2018.
[78] The phrase "winning hearts and minds" is notable for its use in military strategy, where it refers to the winning over of a local population, although the consent of that population is a matter of interpretation (Dixon 2009). It seems telling in this context, where it implicitly frames the introduction of open access as a battle or conflict.

research with the world, or we want to improve mankind, or we want to…whatever it is. Whatever that is - is that being measured? Have we achieved that? 'Oh, well that'd be quite hard to measure.' It's like, yeah, well maybe [laughs] you should have thought about that five years ago!"

Danny's account suggested that compliance framings of open access are set at the very top of research governance structures – the UK research councils, in the case above – and subsequently manifest in university approaches encountered by scientists.

Despite evidence of critical reflection on a compliance framing, university staff implementing open access policies seemed caught in an ongoing bind between attempts to engage "hearts and minds", and strategies that compel open practices regardless of researcher engagement. Danny described how, "ironically" and out of a feeling of necessity, her team was reverting from a values-based framing of open access ("this is a great idea and this is a positive thing to do") to a compliance framing "if you do not do this, you will not get any money. Do it."). This position was reflected by an open policymaker at a different UK university, who implied that he now cares more about researchers' actions than their attitudes and values: "I don't want them [researchers] to *want* to do it [open practices] any more, I just want them to do it." These views seem to stem from a persistent difficulty in engaging researchers, accompanied by growing fatigue and frustration over a period of years. Many open advocates in different roles had stories about their struggle to be heard by researchers, and including instances of outright hostility. Stephanie Dawson[79], who leads the platform *ScienceOpen,* told me that a professor invited her into his office just to deliver "a ten-minute tirade about how terrible he found open access". Accounting for the lack of engagement, Danny said that "[researchers'] eyes are on the novel results in high impact journals - don't care about anything else". A policymaker different UK university expressed frustration that researchers are "sleepwalking", complicit in maintaining – rather than challenging and taking ownership of – a system where publishers have a high level of control.

[79] Interview: Germany, Jan 2018.

This is an illuminating counterpoint to scientists' constructions of open access, and suggests a cycle: policy compliance approaches breed bureaucratic distancing of researchers, with associated loss of agency and frustration; in turn, disengaged researchers frustrate the efforts of advocates; and the advocates feel compelled to reinforce compliance framings. In this way compliance framings of open access may both create, and result from, disengagement amongst researchers. Similarly, scientists, policymakers at research organisations, and open advocates all make observations of a research culture seemingly driven by the pursuit of high "impact", at the cost of various forms of openness, including open access. However, they tend to feel that others are in control. Danny portrayed researchers as stubbornly impact-driven, while Jenny, a scientist in Australia, felt the same about institutions: "…the only interest most institutions have in publication is that they want more. […] I mean it's all become very much scores, and impact factors and stuff like that…". And while some scientists that I interviewed expressed a personal valuing of high impact publication, many who followed this approach were motivated by the expectations of others – particularly assessors of research. Thus, my findings suggest a collectively felt loss of agency, in which cycles of frustration and distancing drive a compliance framing of open access – fuelled by cycles of expectation which have been constraining research publication and dissemination practices for a longer time. This may be a version of the "cultural problem" that is taken to be a barrier to openness in science, and which "culture change" policies in open science increasingly seek to address. Compliance approaches ("I just want them to do it") sit awkwardly with the goal of culture change, as they may aggravate the "cultural problem". Moreover, they involve top-down governance, whilst culture change implies the movement of a whole system or community, presumably including the hearts and minds of researchers.

6.8 Generational change and future orientations

In much of this chapter I have attended to the distance, ambivalence, scepticism, and even resentment that often features in scientists' relationships with open access.

However, this should not be mistaken for a report of widespread disapproval: the vast majority of scientist interviewees were positive about open access at least in principle, along with a range of more complex reactions. It is also important to note that scientists trained from approximately the 2000s onwards had a slightly shifted profile of reactions to open access, as noted throughout, which tended to be more positive and engaged. Financial, journal-centred framings of open access were prominent in all career generations, but open journals and preprint repositories were often actively embraced by recently trained interviewees. These generational dynamics suggest the potential for an acceptance of open access actively led bottom-up by scientists rather than top-down by compliance measures. However, this is uncertain, as for any prediction of generational change outside of a longitudinal study.

There is also major policy change on the horizon in relation to open access: Plan S, centred in Europe, is due to be implemented by many research funders worldwide, including in the UK from 2021. Plan S (European Science Foundation n.d.) makes a number of policy changes with relevance to the themes in this chapter: including limits to open access charges, a removal of support for the "hybrid" route, and stricter open licensing expectations. My interview observations suggest that scientists may welcome – in theory – a number of these changes: especially cost reductions and checks on the power of "hybrid" publishers who claim income from both subscriptions and open access. In practice, where scientists are already distanced from open access procedures, cost reductions may not be felt; and any cost of open access can arouse suspicions of injustice. Emphasis on licensing arrangements and on cost controls, may reinforce financial, journal-centred framings. Compliance framings may be reinforced as the definition of open access becomes more stringent, and as funders promise to "monitor compliance and sanction non-complian[ce]" (European Science Foundation n.d.). However, any policy unification arising from Plan S may ease bureaucratic friction.

Many of the anxieties in interviewees' accounts are associated with a lineage of open access that has become dominant in advocacy and policy, especially in the sciences: one which prioritises reusability and friction-free movement of knowledge (*libre*) alongside free access (*gratis*), has neoliberal resonances, and become big business for commercial publishers (Moore 2017, 2019b). This is not the only type of open access, and Moore (2019b) argues that scholar-led, not-for-profit, open online publishing – owned and run by research communities, and providing a space for experimentation, collaboration, and self-sufficiency away from market forces – is an important alternative. Moreover, such operations have been present but on the fringes since the advent of online scholarly communication, and are currently re-emerging, for example in the form of the Radical Open Access Collective (ibid.; Adema & Moore, 2018). This form of scholarly communication has a stronger tradition in the humanities and social sciences than the natural sciences, and was absent from interviewee accounts – except in the form preprint archiving. However, it has the potential to address deep preoccupations expressed by scientists: financial strain and injustice of steep publication fees; a loss of community feeling and possibly quality in publication; a tense allegiance with "impact" as configured in commercial publication; and a rising tide of bureaucracy that feels distanced from scientific priorities.

6.9 Summary

In this chapter I have presented the first of three significant categories of openness as constructed by scientists in interview: open access to research articles. This category is also highly salient in policy and advocacy discourse. The task of the chapter was to make sense of continuities and contrasts in scientists' portrayals of open access, identifying themes and patterns that together illuminate the relationship between scientists and open access in the context of this study.

In alignment with policy and advocacy histories, scientists with careers spanning decades tend to remember open access as the earliest explicitly "open" practice and movement in science – a forerunner to other forms of open science. More recently trained scientists linked their early "open" memories with open access policies and expanding "open" visions and technologies, suggesting a generational shift in scientists' understandings as the phenomenon itself evolves. Several scientists across generations positioned open source software as their earliest memory of “open” – reflecting an early and ongoing cultural influence from software communities.

The clearest framing of open access was financial. Across contexts and generations, scientists associated a cost with open access, and portrayed it as expensive. For scientists without access to funds, open access was felt to be out of reach. Many UK scientists had access to block grant funding for open access at the time of this study: in those cases, or when funding was otherwise in generous supply, open access was experienced as easy and reduced to a transaction – enabled by the hybrid publishing model. However this framing primed conditionality: scientists suggested that their practice of open access was dependent on the – sometimes precarious – availability of funding. A financial framing also primed negative and ambivalent emotional and attitudinal reactions to open access associated with feelings of scarcity and injustice.

Linked to a financial framing was the construction of open access as a “gold” journal-mediated practice, which tended to be associated with publication fees. The cost-free “green” repository route, where it was mentioned at all, was framed as a less valuable or actualised form of open access – despite its policy relevance. Although this indicated an awareness of the “green” route on the part of some interviewees, others seemed to construct repository-based open access as a lesser form of scholarly communication. There was evidence that more recently trained scientists may – outside policy framings – see preprints as a form of open access that combines the benefits of “gold” and “green”.

A minority of scientists associated open access with concerns about quality and rigour in scientific systems: this was linked with the rise of predatory publishing, and perhaps a deeper sense of loss – for established generations - of order and community, coincident with the rise of open science. Many more scientists associated open access with a compromise in journal “impact” despite attempts of open movements to challenge this thinking. However, impact was often pursued for its apparent value to *others* who make career-defining judgements, in opposition to internal values. Some –

especially those more recently trained – did not link open access with reputational concerns, and some framed open access as a reputational asset. However, junior scientists had less agency over publishing decisions, and were influenced by others' reputational concerns.

Scientists' experiences of open access were shaped by a policy compliance framing through which open access was distanced and bureaucratic. This was especially the case in the study's UK context where professional services staff mediated the practice of open access, managing compliance at a university level by shielding scientists from the labour and procedural expertise involved in matching practice with policy. As a result UK scientists sometimes experienced open access as easy, but lacked an agential connection with the practice. In other cases, bureaucratic friction arising from the compliance framing led to frustrating experiences of open access. Advocates of open science within universities gave responses that matched these patterns of behaviour, one describing a struggle to "win hearts and minds" on open access. The frustration caused by this struggle seemed, ironically, to reinforce compliance approaches – creating a cycle of frustration and distancing.

Generational differences in the construction of open access suggested the potential for a gradual scientist-led embrace of the practice, as more junior scientists upheld the reputations of open access journals and supported preprints. However, any such prediction is fraught with uncertainty, and does not apply across all scientific communities. The policy changes soon to be expected with Plan S may result in an open access landscape quite different from the one encountered by scientists in this study. On the other hand, some of the patterns I have identified may be reinforced.

Chapter 7 | "You feel quite vulnerable"

Opening up research data

Luke was a few months into his first post-doctoral position when we spoke. His manner was cheerful and laid back, but he was also feeling the pressure of his new career stage: it was in the same lab as his PhD, but "a big change" – with a renewed focus on producing papers.

When I asked him about openness in science, he began diffidently – perhaps due to modesty, or a real sense that his views may be atypical – before making his position clear:

"So I might be a bit warped in my views here but I think openness in science should be that everything is available. So if I publish a paper, I put all my datasets, all my R scripts, I put everything I can into it so that it's completely open. Which makes you feel quite vulnerable…"

Luke was not to be alone in holding these commitments, especially amongst interviewees of his generation. He was more unusual in articulating the vulnerability of data openness. His account helped me to understand the emotional range that characterised this topic.

Luke, post-doc, evolutionary biology

7.1 Introduction

This chapter is about openness of research data, the second of three significant categories emerging from my interviews with biological scientists, and thus a second interview-based answer to my overarching research question: *how is the meaning of "open" (or "openness") being constructed in the context of science?* The prominence of data openness, like that of open access, was unsurprising due to the salience of open data in policy and advocacy discourse (see Chapter 5). Norms and policies that expect digital research data sharing have been in place since the 1980s and 1990s in some disciplinary niches relevant to the life sciences: genetics, genomics and x-ray crystallography in particular (Strasser 2019). In the 2000s, and especially in the 2010s – following shortly in the wake of open access – "open" research data was increasingly advocated across disciplinary contexts. Advocacy has come from both grassroots (e.g. Murray-Rust et al. 2010) and top-down (e.g. HEFCE et al. 2016; The Royal Society 2012) sources. By the time of my interviews in 2017-2019, biological scientists in the UK and Australia may have become familiar via multiple routes with the concept of open research data, including major funder and journal policies that increasingly require or encourage data sharing upon publication. Once more, however, I could not be sure that open research data would be a distinguishable, salient category of openness for scientists; nor could I predict how they would understand, frame, and prioritise the concept. Open research data is less bounded than open access, as it may refer to many different types and hierarchies of data with varied epistemic and temporal relationships to a research project. Nonetheless, as conceived in advocacy and policy terms, open data has certain consistent features: digital, online, publicly accessible, licensed for reuse, machine-readable, curated with relevant metadata, readily discoverable, and securely archived. Not all of these features may be reflected in scientists' accounts as they form their own concepts of data openness with varying degrees of closeness to advocacy and policy discourse. I use the broad term "data openness" in this chapter to reflect an awareness that scientists' constructions of scientific openness in a data context may be varied, and formed at some distance from "open data" movements. In this chapter I present salient themes and patterns that emerged in relation to data openness during a grounded analysis of interview responses.

7.2 The salience of data openness in interview

Of the 40 scientists I interviewed, approximately half [19] mentioned data openness – or related concepts such as data sharing, data access, or data transparency – in their response to my initial, broad question about openness: "What first comes to mind when you think about openness in science?". Nearly three-quarters [28] mentioned it spontaneously at other points in the interview, and nearly all [37] showed familiarity with open data practices when I asked directly. This indicated a similar, or marginally lower, salience and familiarity than with open access [respective figures: 19, 31, and 40] in the interviewee group. Despite this quantitative similarity, it was not necessarily the same individuals who raised the two categories: 7 individuals mentioned data openness but not open access in their first response [6 vice versa], and over the course of the interview, 4 raised data openness whilst never raising open access [4 vice versa]. For the purposes of these indicative figures, I was looking for contextual indications that interviewees were referring to online sharing of digital datasets with a wide audience rather than more abstract interpretations such as communicating one's "data" via a talk or conversation (this is examined in Chapter 8). Superficially, this indicates that data openness is commonly associated with "openness in science" in biologists' day-to-day professional lives. Again, it is also notable – given its prominence in advocacy and policy – that a high proportion of interviewees did *not* mention data openness in their first response, and a significant number [12] never brought it up. Several [5] raised neither of the two categories most primed by policy – data openness and open access – and had constructed different meanings of openness, discussed in Chapter 8. Data openness thus emerged as a significant category of scientific openness in the context of my study, but again less significant than some may predict.

The ways in which interviewees constructed openness in relation to data, and the range of associated attitudes and emotions that they expressed, differed from the case of open access. Open access was widely approved of in principle; in practice it was associated with a degree of frustration towards costs, bureaucratic processes and power imbalances, but it was also often accepted, impassively or with a degree of enthusiasm. Data openness was associated with a wider and distinctive range of reactions. Those who approved often expressed principled attachment and passion, focusing on the topic for large portions of the interview. Negative or unsettled emotional reactions also featured more strongly: fear, defensiveness, and vulnerability arose alongside frustration and scepticism. Together, these reactions suggested a deeper connection to scientific practice and professional or scientific identity in the case of data openness. Openness in the context of data was also clearly constructed as an epistemic virtue in several cases. It became evident that understandings and attitudes were highly contextual, entangled with factors such as data formats and infrastructure; the timing of openness; resourcing; and social support networks.

In the following sections I explore interviewees' constructions of data openness through four characteristic reactions that I observed: familiarity and acceptance; scepticism and unease; passion and principle; and vulnerability. I will consider each in relation to the material, social, and temporal data contexts that seem to mediate attitudinal disposition, as I gradually build a conceptual understanding of how such diverse dispositions arise. I begin with some groundwork on the language of data openness used by interviewees, and the variety of practices to which it referred.

7.3 Terminology: do scientists talk about "open data"?

Given the breadth of data-associated terminology used in open advocacy and policy discourse – including open data; open research data; open access to research data; data sharing; data access or availability; research data management; and FAIR data (see Section 5.3.9) – I paid attention to language used by interviewees as an indication of how settled or wide-ranging biologists' discourse on these concepts may be. I occasionally introduced the term "open data", but this was rarely before the interviewee had framed the topic themselves. Overwhelmingly, I found that interviewees used descriptive terms rather than any settled jargon to refer to online sharing of research data. For example:

- Jude[80] raised the issue of "data deposition in large databases";
- Gavin[81] talked of an instance in which "all our data had to be open";
- Jaqueline[82] spoke of "obligations now to archive data";
- Michelle[83] raised "the sharing of knowledge and data", including "...the issues of big data, and you know, all the 'omics stuff, and...when people generate such a huge amount of data like that, and then making it available to other people";

[80] PhD 1970s; interview: Australia, Jan 2017.
[81] PhD 1990s; interview: Australia, Jan 2017.
[82] PhD 1980s; interview: Australia, Jan 2017.
[83] PhD 1980s; interview: Australia, Jan 2017.

- Ben[84] spoke of “complete transparency in sharing of data”;
- Arthur[85] mentioned requirements “that you publish the data, the raw data”;
- Jason[86] referred to “free to access to data”;
- and Madison[87] talked about “a really big trend towards sharing the data”.

Only scientists substantially involved in advocating openness appeared to use of the phrase “open data” routinely. This included Thomas[88], a biologist who works in research data management, and Greg[89], who leads an open access journal. Similarly, only Thomas referred to FAIR data or “research data management”. The heterogeneous, descriptive language of data openness used by biologist interviewees suggests that associated data concepts and practices are diverse, lively, and under negotiation. Where open access has become familiar jargon in research communities, data openness has not been reduced to an equivalently recognisable label. The data practices represented by this diverse language were often implicit. Where further context was given, or emerged upon questioning, it became clear that multiple practices of varying significance were invoked: for example, data sharing was assumed to occur at different times in relation to publication, and the constitution of “data” differed according to context. I have taken this into account, and provided contextual information where available, in my interpretations.

7.4 Familiarity and acceptance: disciplinary data sharing norms

A significant proportion of scientist interviewees [about one-third[90]] discussed online data sharing with familiarity and acceptance, indicating that it is uncontroversial in their fields. These interviewees’ disciplinary associations – genetics, genomics, and x-ray crystallography – indicated the continued salience of particular digital data sharing norms for sequence and crystallography data that pre-date the rise of “open

[84] PhD 1990s; interview: UK, Jan 2018.
[85] PhD 2010s; interview: UK, Jul 2018.
[86] Mid-PhD; interview: UK, Nov 2018.
[87] Mid-PhD; interview: UK, Jan 2019.
[88] PhD 1990s; interview: UK, Oct 2017.
[89] PhD 1990s; interview: US, Feb 2018.
[90] The boundaries of this category are subjective; my count is conservative and does not include those categorised as principled or passionate in their approach to data openness.

data". These norms were established as far back as the 1980s and 1990s (Strasser 2019), both by journal policies and – in the case of large-scale genetic sequence data – by the international Bermuda Principles and Fort Lauderdale Agreement (HGP Information Archive 1997; Wellcome Trust 2003) in the context of the Human Genome Project. For instance, Gavin[91], who was involved in a publicly shared large-scale HIV sequencing project, commented that "…the enforcement by journals to publicly deposit all your sequences in GenBank etc." was "a great leveller" in terms of the resources available to a community for reanalysis. GenBank was one of the earliest online databases involved in the establishment of data sharing norms (Strasser 2019). Other interviewees working in genetics and genomics also described certain kinds of data sharing as routine: "Yes, so with microarray data or RNA sequencing data…we do try and do it routinely" (Yvonne[92]); "all of our sequence data goes into the sequence data archive, ENA, European Nucleotide Archive" (David[93]); "Oh yeah like when you upload to NCBI, yeah - I do that, I mean you have to anyway" (Henry[94]); and "so for sequence data - yes, everything goes online. It always has done" (Ben[95]). This familiar relationship with data openness tended not to be highly salient in interview: interviewees in this category tended to highlight other dimensions of openness before data, indicating that this practice was not central to their construction of openness.

Jenny[96], who was working in animal genetics and genomics, described sequence sharing that was not only routine but immediate:

> *"In genomics of course, it happens pretty much automatically, any sequence we get goes immediately open on the website, so - you know, we can't publish any sequence without it being accessible. And that's been policy in genomics for a long time, and of course it has enormously advanced the field, I don't think anybody would tell you otherwise. So we appreciate that and of course all the genome data is immediately published. Everything that we've got on the [specific animal] is already online. So people are already, even though we haven't published the data yet, people are already accessing it…"*

[91] PhD 1990s; interview: Australia, Jan 2017.
[92] PhD 1970s; interview: UK, Nov 2017.
[93] PhD 1970s; interview: UK, Dec 2017.
[94] PhD 2000s; interview: UK, Dec 2017.
[95] PhD 1990s; interview: UK, Jan 2018.
[96] PhD 1970s; interview: Australia, Nov 2017.

This mode of data sharing is in line with the agreements made around the turn of the century for genome and other large-scale sequencing projects. Outside this disciplinary context, pre-publication data sharing was very rarely reported by interviewees, and never reported as a disciplinary norm. The Fort Lauderdale Agreement (Wellcome Trust 2003) asserts that in this context, data producers should be allowed to publish about their data first, even though others can freely make use of it. Sarah[97], a senior journal editor, told me that in her experience, this "gentlemen's agreement" is mostly observed. This offers a level of reassurance, in the form of community etiquette, that data sharing will not be competitively risky.

7.5 Contextualising familiarity and acceptance

These examples firstly suggest a persistence of specific data sharing norms established twenty to thirty years ago via journal policies and international agreements. It may be that this persistence partly arises from the first-hand experiences of interviewees in this group, who were largely trained prior to the 2000s, making them the generation in which such norms were established – or their immediate scientific descendants. A relative absence of younger interviewees from this group does not necessarily suggest an abandonment of these norms; it potentially suggests a reconfiguration of such norms as part of a broader openness concept amongst recently generations (see Section 7.8). This finding may also be shaped by the disciplinary backgrounds of senior interviewees in my study, a high proportion of whom were trained in genetics.

I found that these routinely accepted forms of data sharing were conceptualised by interviewees as specific, both in relation to timing and data type. Each culture of data sharing in this category seemed to have its own locally developed expectations, whether these were of immediate data release, release upon publication, or release after an acceptable embargo period. What was viewed as open behaviour in one context would likely be viewed as closed in another, but within each context

[97] Senior biomedical journal editor, prior career in biological science; interview: UK, Jan 2018.

acceptability had developed. Moreover, each expectation applied to a specific standardised data type, e.g. DNA or RNA sequence data, microarray data, or Protein Data Bank files. Interviewees in this category only brought up these kinds of standardised data types in relation to data openness, but upon questioning it emerged that some of them produced other data types that they did not expect to share. For instance, alongside their open genome data, Jenny's lab produced "tonnes and tonnes of mapping data" – physical mapping of sequences to chromosomes, in the form of photographs. Jenny was positive about the idea of publicly archiving this data, partly due to its "voluminous" nature which made it difficult for her to store; indeed, a collaborator had investigated archiving it in processed form. However, she did not view release of the original mapping data as feasible due to the absence of a specialised repository. Moreover, she described feeling no top-down pressure to do so – "quite the reverse, nobody seems to care at all". Likewise, Henry routinely uploaded sequence data, but also – it emerged – produced image data: histology results. Henry reflected that "it would be really cool if that kind of thing could be shared", and had briefly looked into it, but had become too busy to go further.

Jenny and Henry's accounts illustrate the specificity of their disciplines' data sharing norms; non-standard data were invisible in the interview discussion until I asked. This in turn illustrates the constitutive role of digital infrastructure in data openness: the norm of sequence data sharing was built upon the availability of sequence databases (Strasser 2019). Although generalised data repositories now exist, they clearly do not play an equivalent role, especially for scientists habituated for many years to format-specific data-sharing. Jenny and Henry were in part positive about sharing other data types because they envisioned practical and epistemic benefits, such as being able to offload storage of huge files, and visualise their data in new ways through sharing platforms. This suggested an attachment to data sharing grounded in contextual functions and conveniences – like those that arose from sequence databases – rather than a generalised embrace of open principles.

Additionally, scientists who accepted sharing norms for certain data types were not necessarily comfortable with sharing a layer of more "raw" underlying data. This emerged in an interview with Ben[98], who described the reaction in his genetics-associated field when journals began requiring greater depth in data sharing:

> *"So in genetics now a lot of the larger journals are requiring you to not only deposit your DNA sequences on GenBank and public-access databases, but also all of your raw data, your genotype data, so every measurement from every individual at every DNA marker, to be put online [...]. And that was a bit of shock to a lot of people - it was like, you want my data as well?! That's my data."*

Ben indicated that this detailed form of data sharing was now routine in his field, albeit avoidable in certain journals. However, his account highlights the conditionality with which data sharing was accepted in his field; extension of the principle across context was uncomfortable. Neil[99] alluded to a similar transition in crystallography: "once upon a time people would perhaps make the PDB ['just a list of coordinates'] available but wouldn't bother uploading their diffraction data ['the raw…the experimental data']". It is possible that data sharing norms established in the 1980s and 1990s to some extent eased the acceptance of extensive data openness in certain fields in the 2000s, but Ben shows that such transitions may not be easy. This may be linked with a rise in the interrogative function of sharing increasingly "raw" data that unveils more of one's scientific process, and imperfections. This reflects shifting connotations of data openness in the 2000s, as transparency, reproducibility, and trustworthiness come to be seen as functions of openness, alongside resource-building (see Section 5.5.2).

Moreover, a routineness of data sharing did not eliminate friction or competitive tension. Interviewees reported deviations and ambiguity in relation to apparently long-accepted norms. For instance, Ben indicated that for genome sequencing projects, epistemic ownership and the readiness of data for release were negotiated boundaries. An example of this is what he called a "typical land-grab":

> *"If you say you're sequencing the genome of a species, a wildlife species. And everyone will benefit from having access to that genome once it's published, and everyone's thinking: for my research it would be really helpful to get a genome. If someone says, "well I'm sequencing the genome of this", that's basically saying - I'm doing it,*

[98] PhD 1990s; interview: UK, Jan 2018.
[99] PhD 2000s; interview: UK, Nov 2017.

so don't bother [laughs]. You're stamping your…you know, you're claiming some territory."

He went on to describe how a genome project, once announced, may be delayed for years for reasons that are opaque from the outside. The promise of openness acts as an ownership bid, granting a period of closure during which the data producers make private judgements about the readiness of sequences for sharing. According to Ben, an extended approach of this sort risks frustrating the community and may lead to duplicate attempts. It is, however, "open" compared with an unannounced approach: "some people are very quiet about what they do, and you don't know anything about it until they've published a paper." This negotiation of openness as a social as well as a material act has echoes through centuries of the history of science (see Section 3.4.2.1, *Historical example of openness and closure in practice*). Today's data sharing norms are not, therefore, free from the social context and tensions that operated in science even prior to contemporary competitive pressures.

Additionally, some large-scale genetic projects involve more explicit forms of closure: David[100] spoke of consortium projects that "have no choice but to share data" but nonetheless insist on "complete secrecy" prior to this point due to competitive and commercial interests ("we actually had to sign documents […] undertaking not to disclose information to any outside group"). On a smaller scale, Roger[101] was involved with the sequencing of a genome that was embargoed by a collaborator, who "felt strongly that we would lose out" due to competitors "siphon[ing] off the low-hanging fruit". This suggests that data-sharing norms established around the turn of the century are applied flexibly across communities producing genomic and related data types – and in some cases they have broken down, or never been established.

[100] PhD 1970s; interview: UK, Dec 2017.
[101] PhD 1980s; interview: UK, Jan 2017.

7.6 Scepticism and unease: defensive reactions to data openness

In contrast – upon first impression – with the familiarity and acceptance shown by interviewees in genetic and crystallographic fields, there was a smaller group of scientist interviewees [4, or 10%] who reacted defensively to the idea of data openness, showing frustration, scepticism, and unease. This did not mean avoidance of the topic: most of this group highlighted it early on, indicating salience. For instance, as part of his initial response to my openness prompt, Ernie[102] commented:

> *"There's this new thing coming along where - wanting people to make all their data open. But that's, in my opinion, particularly groups of people who sort of have an axe to grind and don't think about - really about how what they're arguing for will impact many other people..."*

Ernie's reaction was strong, immediate, and strikingly defined in relation to an opposition group. Jude[103] similarly raised the issue early, expressing disquiet:

> *"Okay, so the general public's paid for these data, but I have real problems with, you know, so a lot of grants and so on say okay you've got to dump your data in a central database...immediately you get them. And I think that does no one any good because you need to know a lot more about acquisition of the data, the care with which it was acquired, and the way the systems works..."*

Jude followed with other concerns about data openness: that a "big lab" will mine the data faster than his "two-man operation in Australia"; and that online data may not be provided in good faith – he knows researchers who hold back selected sequences to maintain a competitive advantage. He occupied an embattled position on the topic, defending on both epistemic and personal fronts: his concerns about the online data's completeness, integrity and validity – *what use is it anyway?* [paraphrased]– come alongside a threat to his own scientific enterprise – *this feels unjust and risky.*

[102] PhD 1970s; interview: UK, Oct 2017.
[103] PhD 1970s; interview: Australia, Jan 2017.

Jacqueline[104] reacted with a more quizzical tone, but similar concerns:

> *"It's very hard work, getting your data and whatnot, and I find it curious on some of the things they've…obligations now to archive data, for example. And I find that really fascinating. Because most of the time, you're doing an experiment, you're going to have very particular conditions that you impose or that you use, so the data are not necessarily going to be all that useful in a direct sense, to somebody else, unless they're doing the exact same measurements […] So I don't have a problem with publishing it [the data], however, here's the rub: we spent the last three years getting these data together, working very hard, we now see that this set of data can be used in several ways to address these questions and we want to get it out there. Once we've published it, we've lost control of it."*

Once again, personal and epistemic objections were combined: the injustice of working hard only to lose control was expressed together with doubts over the scientific purpose and validity of data sharing. Conceptually, Jude and Jacqueline's accounts imply that open online data may lack integrity and usefulness to others, and also pose a competitive threat to themselves[105]. A conceptual interpretation, however, may miss the point. Ernie, Jude, and Jacqueline's accounts show how some scientists perceive the idea of data openness not as an opportunity or solution, but as a nuisance and a threat. Competitive, even ruthless behaviour is expected from others; protection of data is normalised; and trust in a wider community is low – sharing is "losing control". Epistemic objections are perhaps rationalised secondarily to these existential concerns (which is not to invalidate them). In particular, Jude and Jacqueline indicate that their data are difficult to separate from their context of creation, a circumstance that is likely to occur when research settings and data types are less standardised. Together with Ernie's objection that advocates of data openness do not think about impacts on "other people", there is a hint that scientists who react defensively feel that "open" expectations are being imported, without empathy, from data contexts in which sharing is easier.

---

[104] PhD 1980s; interview: Australia, Jan 2017.

[105] This seemed like a conflict to me, at face value: if the data are not useful to others, how can they be provide a competitive advantage? I then realised this was not a binary: interviewees made it clear that shared data or ideas need not be complete to provide clues or nudges to competitors.

The observations in this section, although they apply to a small number of interviewees, are consistent with other evidence that a significant proportion of researchers react negatively to the prospect of data openness. Some advocacy documents suggest these ideas should be approached cautiously, perhaps with language that de-emphasises openness and sharing (see Section 5.3.9). Moreover, Danny Kingsley[106] spoke vividly on the topic:

> *"We do not use the term "open data" anymore. Because it's very, very poorly received by the research community. When we first started talking to the research community about open data, my colleague who has a wonderful turn of phrase [...] she said: we were speaking to them, and they were throwing the rotten vegetables at us! [laughs] It was true! I mean, not literally. But they were yelling, a lot of them. Like, very emotional."*

She goes on to explain that researchers were troubled by the organisational state of their data, having not expected it to be widely visible. To address this, her team adopted a "research data management" framing, reasoning that data well managed from the outset is not embarrassing to share: opening up is just one small, extra step. However, the interviewee accounts above suggest that this could be just one of several uneasy aspects of data openness.

7.7 Contextualising uneasy reactions

Further dimensions emerged as interviewees expanded on their uneasy reactions to data openness. As Ernie's account continued, it became clear that the practice he objected to took a particular, relatively radical, form:

> *"...people who are arguing that all - basically lab books, all lab books should be electronic. At one extreme there are people who are arguing that the content of those electronic lab books should be made accessible to anybody at any time - to everybody, basically. Right from the moment they're created."*

Ernie's concern was thus not necessarily about mainstream, commonly mandated data openness practices, but as he put it, the "extreme" end: what is sometimes labelled open notebook science (see Sections 1.2 and 5.4). He envisioned requests for

[106] Head of Scholarly Communication, University of Cambridge; interview: UK, Feb 2018.

all his data – and upstream ideas, notes, and other lab book contents – to be public; and for it to be released in real time, before publication. It became clear that more conservative and regularly practised forms of data openness – release upon publication of associated datasets – might be perfectly acceptable to him. Indeed, he talked about having uploaded his raw data into a journal's data viewing tool, after publication, and his comfort with sharing data underlying a specific published result:

> *"I don't have any problem with giving people access to the raw data, but I think that people should have access to the raw data that's in the publication, and that's backing up what's in the publication."*

With statements like these, Ernie's discomfort with data openness was contextualised, and his attitude could be viewed as one of conditional acceptance.

However, it remains notable that Ernie began with the extreme case, and his strongest possible objections to data openness. Viewed cynically, this could be interpreted as hyperbole that avoids engagement with more limited forms of sharing. Notably, however, the extensive openness which Ernie opposes is in line with the techno-utopian visions of some advocates (e.g. Bartling and Friesike 2014b; Nielsen 2009). Ernie, and other interviewees, seemed to conceive of and experience open science as a rolling frontier in the same way as it has been framed, and has developed, as an overarching movement (see Chapter 5). Although open access and open data are the only practices currently widely mandated in UK and to an extent, Australian contexts, recent policy and advocacy documents suggest that interest in extending open requirements further, including to unpublished articles and data, and methods (LERU 2018; SSAC 2019; Tennant et al. 2019). It may be that the expansiveness of "open" – both conceptually, and in policy – has a role to play in the oppositional stances of some scientists, who expect to be asked for more and more (data) openness, and perhaps eventually for unlimited access to ongoing scientific processes and ideas. Ernie likened extreme openness to the conditions of a "totalitarian society".

Similarly, Jude's concern that a large laboratory could take advantage of his data was associated with the idea of *immediate* data release. When I asked him about data sharing with, or after, publication, his attitude changed: "I think with publication, yes, there's very few things I would say shouldn't be shared [...] as long as the lab that did it has first access to the data." Jude was even happy to consider data release prior to publication, upon acceptance, if publishing priority for the data producers could be arranged. This latter condition shaped his views strongly. Receiving this credit in "the system" was seen as crucial in a competitive environment where data, and the funding to produce it, is hard won. Like Ernie, Jude placed himself in opposition to those who ask for data openness not because he objects outright to making data available on publication, but because this request may presage demands for immediate openness – and the need to defend his "little lab that has worked its butt off for a year". Jude's expectation of immediate data sharing might have been derived either from utopian visions constructed by open movements, or more established expectations of immediate sharing for large-scale sequencing projects – his field overlaps with genomics, so he may have directly encountered the latter. However, he did not mention the community etiquette – asserted through the Fort Lauderdale Agreement – that was intended to address concerns about priority in a genomic context of immediate data sharing (Wellcome Trust 2003). It may be that, inside or outside genomics, Jude's community does not uphold any such norm. His views underline the importance of publication as a boundary before which data openness is seen as a particular threat, especially in the absence of a community consensus protecting a sense of fair play for data producers. It is notable that the UK *Concordat on Open Research Data* (HEFCE et al. 2016) suggests that such a consensus should apply across disciplinary contexts, but how it can build the trust required in communities is unclear.

Jacqueline's position is also interesting to examine in context. It emerged that one of her research projects already required data sharing. She was therefore expecting, after completing data collection, to upload a "massive dataset" to a project-associated archive that allows gradual, publication-associated data release:

> *"...it's a simple matter of putting an embargo on it, as we're working through papers in various areas and then as we submit those, as those papers become published then we just open up the portal."*

Jacqueline framed this practice as appropriate and non-threatening. Although she remained sceptical that her data would be useful to others working in different contexts, she gives consideration to the possibility:

> *"You'd probably get more out of reading the paper than you would of looking at these numbers in an archive [...] But, I don't know, maybe someone will find that useful. I don't have a problem with it."*

Again it was striking that Jacqueline expressed unease about data openness before her conditional acceptance – and actual practice – of data openness emerged. As for Ernie and Jude, her conditional acceptance hinges on the post-publication timing of data release, and the practice she adheres to seems designed to ensure that researchers that they will not "lose control" – that they will be able to gain credit for their hard work. Despite the emergence of conditional acceptance, initial negative reactions signal a lack of trust in advocates' and policymakers' visions for open science. For these scientists, data openness is constructed as a promise of something more radical: that science will change in ways that transgress familiar boundaries, threaten epistemic integrity, and remove a sense of control.

Ernie, Jude, Jacqueline and Steve[107] (whose account also fits this category) have in common that their disciplinary identities – cell biology, zoology, ecophysiology, and plant biochemistry – are outside, or only partially overlapping with, domains in which online data sharing traditions have been established. It seemed that in these fields, data sharing is not supported by a community consensus that might make it easier for individuals. Moreover, these interviewees seemed to deal with different modes of knowledge production, and different data types: data that may be more laborious, proportionally, to produce; larger in size (thousands of microscope pictures, in Ernie's case); non-digital; and more difficult to interpret outside its context of production. Ernie expected that sharing his data "in a comprehensible way" would take months of extra work that he could not imagine his funder supporting. Jude and Jacqueline portrayed their data production circumstances as so specific and entangled that data lose integrity and meaning as they travel. Although open and FAIR data discourses emphasise the need for rich metadata (Wilkinson et al. 2016:4), they also emphasise

[107] PhD 1980s; interview: UK, Nov 2017.

standardisation, and not all contexts can be communicated in this way: subtleties such as “the care with which [data] was acquired” (Jude), for instance. As a result “open data” is likely imbued with assumptions from settings in which both research conditions (e.g. laboratories) and data types (e.g. nucleic acid sequences) are more standardised and digitised, with commonly assumed context, lower costs of curation, and more digital infrastructure support. Scientists working in less standardised settings not only encounter greater barriers to data openness, but face the implicit suggestion that their way of working is suboptimal – even less scientific – because it is more difficult to “open”. Advocacy, policy, and academic literature has reckoned to an extent with this diversity: there are repeated warnings that a “one-size-fits-all” approach to open data must be avoided (HEFCE et al. 2016; Levin and Leonelli 2016; The Royal Society 2012). My interviews underline the importance of these warnings, and show that scientists anticipate one-size-fits-all before its enforcement. Ernie, Jude and Jacqueline anticipated norms from other fields being applied to them, and expressed resistance.

The four interviewees in this category were all senior and well-established in their fields, having received their PhDs in the 1970s and 1980s. The unease in their reactions might partially be about a dislike of disruptive change, but it cannot be reduced to this. These scientists’ views were based on experience accumulated over decades in particular community contexts and interwoven with knowledge of how data interact with local competitive and collaborative dynamics. The following sections show how their reactions differ from those of many other interviewees, but also how they are not alone in experiencing data openness as unsettling.

7.8 Passion and principle: embracing data openness

A third category of reaction to data openness is best characterised neither as familiar/conditional acceptance, nor as unease: a significant group of interviewees [about one-third] expressed actively principled or passionate attachments to data openness. These interviewees tended to bring the topic up early on, and to construct openness in science largely in relation to data practices and related methods. They tended to discuss the topic extensively: for several, it shaped the whole interview. The meaning of data openness as constructed by these interviewees reflected the breadth of post-2000 open science movements (see Chapter 5), rather than the more specific ambitions of data sharing as developed in genomic or crystallography contexts. Although passion or principle united these "open" attitudes, they seemed to emerge from diverse sets of values, both within and between individuals. Unlike acceptance of data sharing in genetic fields, which was context-specific, the interviewees in this group tended towards a holistically positive attitude, not only towards different forms and contexts of data sharing, but towards a wide range of additional "open" practices. This did not mean an absence of sensitivity to timing and extent of data sharing in relation to publication. Some of these interviewees were speaking more of hypothetical than actual practice, due to limitations in circumstance, or an early career stage. And, strikingly, most of these interviewees *were* early in their careers: they tended to have been trained since the 2000s, and many were PhD students. The few more established scientists in this category were also established as open science advocates.

7.8.1 Software principles

One route by which an attachment to data openness seemed to develop was software culture. Elliot[108] spoke about being "indoctrinated" into the ideals of open source software by his brother, and then seeing its benefits in his field of molecular and computational biology. A new generation of sequencing technology was introduced in his area that required computational expertise for data analysis; computer scientists were recruited and brought open source software and methodologies with them. According to Elliot, the analogy between software and data then became compelling:

> *"So then people were like well if all the software's open, why shouldn't the data be completely open as well? Why shouldn't everyone be able to use all of the data that's being produced in all these papers?"*

[108] PhD 2010s; interview: UK, Jan 2019.

He was far from alone in making this connection. Mat[109], an organic chemist and advocate of open science, had built his laboratory's entire approach around the analogy: "we're trying to mimic the way people make open source software and we're trying to apply that to the discovery and development of new medicines". Mat's approach includes, but goes far beyond, open data: it is open notebook science, in which ongoing details of experiments, and collaborative discussions, are online and public-facing. The historical and ideological influence of software cultures upon developing open science movements is discussed in Sections 4.10 and 5.3.3, and seems particularly significant in the data context.

7.8.2 Transparency and reproducibility principles

A software-influenced framing of data openness is in turn associated with a valuing of transparency and reproducibility. As open source software inherently reveals its inner workings and allows them to be reproduced and modified, it resonates strongly with certain scientific ideals: that good science should be reproducible, in part because its processes are transparent. These ideals have received particular attention since the 2000s, as high-profile concerns have been raised that swathes of scientific results may not be reproducible – the "reproducibility crisis" (see Section 5.3.2). This discourse shapes visions of open data and open science, which are seen as remedies. The value of transparency and reproducibility was also upheld by interviewees strongly attached to data openness. Elliot's focus on software and data, for example, was introduced through the frame of transparency:

> *"What first comes to mind [re: openness in science] is the transparency of everything that you do. So that's transparency in all the methods that you carry out, and transparency in the results that you publish, about what you show to other people."*

Madison[110] also spoke predominantly about data openness, introducing it in relation to transparency and reproducibility, with connections to code and method sharing:

> *"… what comes to mind for me [re: openness in science] is that it's very transparent, that scientists are quite upfront about sharing what*

[109] PhD 1990s; interview: UK, Apr 2019.
[110] Mid-PhD; interview: UK, Jan 2019.

> *they did and more and more like their actual data. So I think that's how I've thought about it a lot recently because there's a really big trend toward sharing the data and sharing code and sharing methods and that kind of thing, rather than – and also reproducibility as well, I think probably, or with an aim toward reproducibility."*

For Arthur[111], reproducibility shaped his earliest encounters with "open" in science, which also occurred as he moved from a chemistry background towards biology:

> *"…the other thing that came across when I first got into life sciences was reproducibility, like the reproducibility process, so it's openness in the context of that. I have a friend who's actually at [an open science institute] […] And they write a lot about openness in science and how you make science more transparent and how to make it more reproducible. So yeah that's where I really first encountered it."*

This mode of attachment to open data, which extends into broader "open" visions, was framed as a principled pursuit of good, rigorous science. This finding has clear implications for my sensitising concept of epistemic virtue (see Section 2.2.1), and I discuss it further in the final chapter (see Sections 9.3.2 and 9.4.2).

7.8.3 Public duty principles

Another factor motivating strong attachment to data openness was a principle of equal, democratic access for all, including the lay public. This perspective overwhelmingly shaped Jason's[112] self-declared passion for openness in science:

> *"So when I think of openness in science I'm thinking of free access to data. So data generated from a project. I'm thinking, and by free I should say free of cost and free of restriction[113]. You don't have to be a member of a university research group to access it. You're a member of the public and you can access this material."*

For Jason, open data and open access blended together as an information domain that should be public-facing due to taxpayer investment. This also connected with what he

[111] PhD 2010s; interview: UK, Jul 2018.
[112] Mid-PhD; interview: UK, Nov 2018.
[113] Notably, where "unrestricted" in an open advocacy context usually refers to *libre* licensing, Jason defined it as access across institutional boundaries. This was consistent with near-absence of licensing for re-use as a feature of openness in scientists' accounts: Mat and Elliot's accounts, which explicitly discuss software–science analogies, were telling exceptions. This is striking given the prominence of licensing in advocacy and policy literature as well as in interviews with advocates.

describes as his “naïve” and “idealist” perspective, developed from childhood, that science should be accessible to the public e.g. through museums. These values blend with Jason’s conviction that scientists produce better knowledge when their access to resources is improved:

> *“I believe that the more scientists especially that have access to research data, resources, publications, the better your science will be, the more efficient it’ll be and the more innovative it will be.”*

This, again, brings epistemic virtue to the fore, but based on ideals of cumulative, progressive science rather than reproducibility (see discussion in Sections 9.3.1 and 9.4.1). Jason brings reproducibility into his view as well: he runs a website that hosts publicly, freely accessible datasets for researchers in his field, and he only includes datasets he thinks meet standards for reproducibility.

Publicly oriented attachments to data openness were less common than those focused on scientific integrity. However, another notable case was Greg’s[114]: as Editor-in-Chief of an open access journal, he had witnessed an extension of “open” from publications to data, and highlighted both reproducibility and the public interest as reasons to pursue data openness:

> *“…since then [early days of open access] it's also grown into this idea about data accessibility, and open data. And I think that that's enormously important as well, I think that that's helping combat the problem in rigour and reproducibility that we're facing in science. And I think that it lets people who aren't necessarily part of this big, expensive machine that we have at universities and companies, participate, by looking at data, and analysing it post-hoc, […] it includes more insight by broadening the community of people who use it.”*

His response also hints at “opening” of science–society boundaries, in which scientific data is not only accessible, but begins to be shaped and diversified by interests outside of institutional science. Openness as an orientation to publics framed Greg’s entire response, which was rare in the interviews. Significantly, a public orientation – in both accounts – led to the categorisation of data openness as an obligation only for

[114] PhD 1990s; interview: USA, Feb 2018.

publicly-funded science. Private science was viewed separately, a distinction that did not feature in other modes of attachment to data openness (see also Section 8.5.5).

7.8.4 Collaborative and personal principles

Attachments to data openness also emerged in connection with collaboration and personal scientific integrity. These aspects seemed not to drive attachment to data openness in themselves, but to be important secondary aspects for some interviewees. In this case, data openness was valued not only as an act of opening or sharing, but as an ongoing practice, integrated into scientific life, that anticipates openness. Data was imagined through others' eyes – the eyes of the public, collaborators, and even one's future self. For example, Madison[115] told me that whenever she thinks about her datasets, she tries to think about "how would I feel about just making them public". She goes on to describe how this kind of mind-set shapes her data practices:

> *"...[data openness] makes me be much more stringent and reproducible of my own data and my own code and everything [...] I try and design the datasets themselves so that they're not very specialist just in my eyes, that someone else would understand them [...] as transparent as possible and as transferrable as possible [...] Like, even for myself, if I needed to do another type of analysis on the data, it's much more helpful if they're [...] very easily formattable and very clean data [...] I think that's been the biggest thing is thinking about well what would it take for me to just put this in a publicly digestible format, and try and work backward from there. And also generally just make sure that I'm setting up my workflows so that my data is both for me and potential collaborations..."*

This practice of reflective self-discipline is reminiscent of a Foucauldian "technology of the self", a concept that Daston and Galison (2007) invoke to explain how an epistemic virtue – objectivity – operates at the level of the scientific self, integrating into habits of the body and mind. Thus, a clue is offered as to how openness as an epistemic virtue might operate on this level, connected with the pursuit of transparency and reproducibility, as well as accessibility. Madison portrayed openness with collaborators as on a spectrum with public openness, and had become familiar with collaborative data sharing via her lab's regular practices. Her data practices

[115] Mid-PhD; interview: UK, Jan 2019.

occurred in the context of feeling ownership over her data, which gave her confidence to initiate and negotiate collaborations.

Olivia[116], another PhD student, described going to extra lengths to prepare her data for sharing. She wanted to publish in an open access journal with an open data policy, and her data required anonymization. Like Madison, Olivia framed this not as an imposition but as an opportunity for professional self-improvement:

> *"…I think on a personal level it's good for people to be able to reproduce your work easily and see the work that you've done and be able to criticise you. If people can get access to your code they can get your data and they can say: oh well you could have done this better, and you can learn from it and definitely improve as a scientist."*

Again, this framed openness as an epistemic virtue, operating at the level of the scientific self: simultaneously an improvement of knowledge, and knower. Olivia also described, during an undergraduate project, voluntarily keeping an extensive lab book because of a sense that if her work was not recorded properly, she would not be able to share discoveries. She described having an "engrained *feeling* about it" [her emphasis], driven by "*really* wanting to become a scientist". Olivia thus constructed scientific identity according to a sense, both rationalised and subjectively felt, about good knowledge-making, which was expressed and reinforced it through disciplined data practices, with a view to making knowledge shareable. This is vividly reminiscent of Daston and Galison's (2007) portrayal of epistemic virtue in the context of objectivity, which also attends to scientists' note-keeping practices (pp. 244-245); in Olivia's portrayal, objectivity and openness are related (see discussion in Sections 9.3.2 and 9.4.2).

7.9 Contextualising principled and passionate reactions

Principled and passionate attachments to data openness, operating via diverse sets of values, were often expansive in scope, sometimes connecting with overarching visions

[116] Mid-PhD; interview: UK, Jul 2018.

of openness in science, including of open software, methods, and laboratory notebooks. However, these attachments were not independent of context. Most notably, immediate data release was only rarely considered definitive of openness: openness in the final stages of a project, following publication, or even several years later, was more typically assumed. For instance, the data openness that Jason[117] passionately advocated would occur, his view, well after a project is finalised:

> *"…I won't release any of my data until after a year, until after one year from my analysis. So I give myself a year to do the work on my own stuff, then I release it, which I think is reasonable."*

He went on to reason that an embargo of about five years would be appropriate for a five-year project. This would be viewed as an extended delay in many fields, but in Jason's field of palaeoanthropology, embargoes applied by senior scientists sometimes extend for decades. Molly[118], who was also in favour of data openness, distinguished between data directly supporting a publication and a wider dataset. The former she expected to be open with or soon after publication; the latter could be released a couple of years later, "to make sure you've published all of the things that you were aiming to publish from it". The latter was a necessary compromise, in her view, between furthering her career, and furthering science via data openness. All of this was hypothetical to Molly, however: her PhD data were generated by others in remote field contexts; her access was conditional, achieved through long-term negotiation, and did not allow re-sharing. It was from this context that Molly envisaged any form of data release – even after a couple of years – as opening up.

Elliot[119] also strongly endorsed open data, code, and methods. He saw these as means to "good science" – an epistemic virtue framing – on two counts. Firstly, having himself learned from others' openness, he saw holding back as "stopping some actually other good work being done". Secondly, he saw holding back as "like saying you're never going to do any good work again" – a reliance on exclusivity of past work for one's scientific reputation, rather than ongoing work. Again, this embrace of openness was anchored by an assumption that any data released were already

[117] Mid-PhD; interview: UK, Nov 2018.
[118] Mid-PhD; interview: UK, Nov 2018.
[119] PhD 2010s; interview: UK, Jan 2019.

relatively finalised and readily attributed to their producer, thus not at significant risk of "scooping" by competitors: "you're already kind of finished the project once you're required to be quite open about what you're doing". In contrast with Jason and Molly's accounts, however, this did not mean to Elliot a delay of years or even waiting for publication. It meant online dissemination, e.g. posting a preprint, or putting up software on Github. However, he considered the effect to be similar: "...your name is already attached to them. So you can already claim that it was your work if that's important to you." His confidence contrasted with the worries about scooping that he saw as prevalent amongst more established colleagues, for whom openness about process could mean "throwing away" work. An acute generational–disciplinary shift may be at play here. Elliot describes the arrival of computational expertise and culture into his field as recent; his established colleagues may value data, code, and methods differently, having trained in a different epistemic culture. Elliot's confidence that he will receive recognition for online and pre-publication modes of dissemination may be shaped by his immersion in a computational field with high digital literacy and open source values.

In contrast to these cases, Mat[120] took a fully "open source" approach: his lab's data, and the process and discussions surrounding it, were made publicly available "in real time", prior to publication. Many other interviewees viewed such an approach as prone to scooping, and Mat acknowledged that others "don't want to reveal their competitive advantage [...] I get that". However, he seemed unconcerned himself about this type of professional threat as he described to me his open principles, and their epistemic and societal benefits. As the interview went on, it became clear that Mat had, as a side effect of his unusually open and innovative approach, changed the competitive dynamics of the field in which was operating:

> *"...if you're working on a problem and you manage to attract in a few people who are able to contribute to that project, that you generate a momentum which is quite tough to beat."*

Instead of competing, Mat gained a secure position by collaborating so extensively – his project was open to any interested party – that the traditional element of

[120] PhD 1990s; interview: UK, Apr 2019.

competition was virtually eliminated, or made ineffective. Through his approach, Mat had established networks, including in industry, and a professional and public reputation that was invulnerable to the competitive dynamics negotiated by the majority of scientists. This demonstrates not only how certain contexts make open principles and practice more comfortable; but also how open principles can participate in creating such contexts.

The cases above show how principled, passionate attachments to data openness are closely intertwined with contexts that tend not to be initially apparent in interview. Just as defensive responses could become conditionally accepting as context changed, principled attachments seemed to be eased by the contexts in which they were held. Timing was clearly critical: those who were defensive were reacting to expectations that data openness would be required before publication; those were principled and passionate tended to assume that post-publication release – even years later in some cases – was acceptably open. Underlying this dynamic seemed to be an issue of control and security: defensive accounts anticipated a loss of control as data openness was imposed; but positive accounts implied a sense of control and choice. Although all responses were rationalised, the affective qualities of both positive and negative reactions suggested to me that a sense of control manifested at an emotional level and could be a primary influence on the embrace or rejection of data openness. I conceptualise this observation further in the following section, and in **Table 6** and **Figure 6**.

A final set of interviewee responses made a connection between the poles of defensive unease and principled attachment: some interviewees described fear and vulnerability in their practice of data openness, but pursued it nonetheless.

7.10 Vulnerability: feeling the fear and doing it anyway[121]

Melanie[122], a principal investigator who trained in the 2000s, recalled being "super worried" about releasing data on long-term ecological projects she had worked on. This was especially the case for her PhD data:

> *"...I mean I think at the beginning, you're just like...[laughs] so the first dataset I collected was a big one, as a PhD student – and I worked* so *hard on those, like multiple years, out every weekend, and then like...when I published the first paper on that series, they were like "okay well we'd like you to publish all the data". And I thought w- ...I had like three or four other papers on that dataset, and so...it was just something that we hadn't talked about openly as a part of our lab. Like... I wasn't ready for it, right? I didn't know, and so I completely freaked out and I thought like well...I didn't want someone else to take the data and analyse it in these ways."*

The emotion depicted in this response is acute: feelings of shock, not being ready, "completely freak[ing] out", and a need to be protective. This is comparable to the uneasy, defensive category of interview reaction, but is Melanie's reflection upon experiences and feelings in her scientific past. She described having "evolved" to a position of principled acceptance of data openness: "...there's such a big push to make things open and accessible, and I think the downsides are really like, small compared to the positives." She also described data openness as "better for the field and for science", and a duty in the context of public funding. She pointed to embargoes of a year or two, or the release of limited datasets for individual articles, as acceptable ways of managing openness for large, long-term datasets.

Melanie's account is notable not only because she portrayed a transition from a highly defensive position to an accepting one. Her initial and later experiences reveal contextual factors that may have primed her defensive feelings initially, and those which later eased her embrace of data openness. Firstly, her fear about releasing her PhD data was linked to her personal and scientific relationship with the data. From an instrumental perspective, her data resulted from an extensive investment of time and

[121] A reference to Susan Jeffers' (1987) self-help volume: *Feel The Fear And Do It Anyway*.
[122] PhD 2000s; interview: UK, Jul 2018.

effort and was highly valuable professionally – the basis for a whole series of papers. And at that particular moment in time, during her PhD, its value was heightened by her reliance on it for establishing an academic career. Because of the way data are gathered in her discipline, over long periods of time in the field, a new dataset would not be just around the corner. This is very different from Elliot's scenario, in which he could suggest that reliance on past "locked up" work was like "saying that you're never going to do any good work again". Moreover, more was at stake than instrumental value. Melanie had invested herself personally in the data with years of embodied labour, her leisure time, her care and attention, and even her love; she described her relationship with her discipline as a "match made in heaven" and her "scientific love". Especially at a formative time in her scientific life, this investment – that closely blends the personal with the scientific – imbued her data with exceptional value. The loss of these data to a competitor would not just be a professional blow, but a personal wound. Finally, Melanie encountered the requirement for data openness – via a journal policy – unexpectedly, without supportive communication from her seniors: "it was just something that hadn't talked about openly…". This might have contributed to her panic, as she could not draw on time or social resources to gain a sense of control.

Melanie's eventual acceptance of data openness seems to have been aided by key contextual changes. Firstly, she became established in her career: this alone would make her less reliant on any one dataset, and more resilient to professional risks. Secondly, she began to collect data in different ways: not only from the field, but from laboratory experiments that yield smaller and shorter-term – less personally entangled – datasets. Thirdly, she described having supportive discussions with other scientists in her current institution, including those who collect long-term ecological data. There was some indication that these conversations allowed risks of data openness to be articulated, put into perspective, practically managed, and emotionally processed. She had somewhat detached from her former worried self: the concern about scooping started to seem like "a bit of a crazy idea", even "self-centred".

Luke[123], an evolutionary biologist in his first postdoctoral position, also spoke of overcoming fear in relation to data openness (see the opening vignette, p. 190). However, he articulated a different category of data-related risk, and one that was ongoing rather than in the past. From the first time I asked about openness, Luke underlined his dedication to open practices, following immediately with:

> *"Which makes you feel quite vulnerable: like whenever I publish a paper [accompanied by open data/code] I'm terrified I'm going to get an email within like a month with someone saying you've done this all wrong and you need to retract the paper…"*

His identification of this experience as vulnerability is illuminating, because it frames openness as a state of emotional and professional exposure – as well as an exposure of scientific information – that leaves one susceptible to emotional and professional harm as well as growth. This again depicts an entanglement of data with the person producing it, evoked in Melanie's account.

The harm the Luke feared – as well as the benefit he pursued – was not scooping, but the exposure of his scientific process, and the unveiling of potential errors:

> *"…it is scary and it is, like sometimes when I'm falling asleep I'm like oh shit, what if someone finds a huge mistake? Because I'm talking 2000 lines of code and if I've got one letter wrong it could completely change the result […] Like that paper came out last week and I think as long as I'm doing something I don't worry but there is always this fear. And then it always flashes up, so if I've had people email me about my work to ask a question and sometimes about the analysis and that's when the fear comes again. You're like oh crap, are they emailing me to say this all wrong?"*

His recollections were viscerally frank: *vulnerable, terrified, anxiety, interrupting sleep, huge mistake, fear flashing up, is this all wrong?* Luke's disclosure of these emotions could be a window into the experiences of many scientists as they consider opening their data to the world. Feelings like these may well contribute to rationalised objections and defensiveness, which could also be understood as attempts to avoid vulnerability. Luke's account was remarkable as a case of conscious reckoning with vulnerability: he knew how he would feel; he sometimes wanted to retreat; yet he engaged with it repeatedly – and had done so in the week prior to the interview. He

[123] PhD 2010s; interview: UK, Nov 2018.

saw openness as a route to the improvement of science and scientists: better for a mistake to be corrected than for his anxiety to prevent that. He was also motivated by having conducted a meta-analysis through which he encountered refusals to share data by scientists who seemed to have everything to gain – "they would have got full acknowledgement, they would have been cited, and it would have furthered the field that they work in, yet they still wouldn't help". This experience prompted Luke to embrace a whole spectrum of open practices, suggesting that his commitment to openness had become identity-defining: a distinction between himself and those whose closure had been epistemically and socially uncaring. He also described rewarding consequences of openness: forming new professional connections, being able help other researchers, and understanding his own work better.

Luke's attachment to data openness was clearly deep and principled, as indicated by his willingness to endure intense episodes of vulnerability. But again, contextual factors add another layer to the story. Firstly, Luke was not fearful about scooping: in fact, he viewed preoccupations with scooping as "ridiculous". He described with frustration attempts to coordinate research plans with a former supervisor – "to stop us doing the same thing" – that were blocked due to the latter's fear of being scooped (including by Luke himself). However, Luke described his own experiments as "so massive that no one could reproduce them": they take a year or more, even with several scientists working together. Additionally, they involve some resources that are "only available" in the local area. Thus, Luke's principles took form in a research context that was protected from some competitive pressures. In contrast, he described his former supervisor as having been "scooped in the past", trained in a competitive environment, and "really unlucky with grant proposals". Although such differences in attitude and behaviour cannot be reduced to contextual factors, the context is illuminating. It suggests that career experiences – particularly early, formative experiences or adverse events – as well as career security, shape whether scooping comes to be seen as genuinely threatening, or as a preoccupation of the paranoid. Impressions like these became detached from context in scientists' accounts, and seemed to be part of a long-term, internalised feeling of security or insecurity that eases or inhibits openness.

A second context influencing Luke's account – in common with Melanie's case – is social support. Luke described feeling strongly supported by the principal investigator (PI) of his group and his local academic community. Moreover, openness was modelled and actively pursued in these social relationships. Luke's PI was a close and established support base; he also supervised Luke's PhD, in a research group that began with just two of them. Luke described the group atmosphere as laid back and social: a place where everyone can "just chat", but also a place where his career ambitions are actively recognised and facilitated. Luke described his PI as the opposite of his previous (Master's) supervisor: open (interpersonally rather than online; see Chapter 8) and unworried about scooping, having established his career studying long-term data from a remote field site. Luke also described his local academic community as exceptionally social and supportive:

> *"Every day at [time] we go for coffee, and there's a communal coffee pot. So anyone who's salaried buys bags of coffee and we make cafetières and share it and you just sit for half an hour and chat science [...] And then [...] everyone goes for lunch as well. And it's really social. Yet [nearby institute] [...] they don't come to coffee and stuff. So it's a real community within [building]."*

Help and collaborative opportunities were offered cross-hierarchically in his building; it was the opposite of feeling "on your own", which Luke had experienced in a previous (Master's) research environment. Luke also observed an influence of this "community spirit" on his behaviour in other contexts, such as conferences. Additionally, he was close to his PhD cohort – they would discuss issues including data openness on WhatsApp – and was part of a journal club that reviewed preprints. This exceptionally strong web of community support could be expected to anchor Luke's openness in myriad ways. Firstly, the collaborative, interpersonal openness practised by his PI and community seemed to provide a foundation for, or at least to align with Luke's online openness practices. Secondly, Luke himself identified early-stage interpersonal openness as a way to *avoid* scooping, by allowing competing research plans to diverge, or even to become collaborative. Finally, even if openness caused him harm, Luke would know he is not "on his own": his PI and community are likely to build his confidence, allowing him to embrace vulnerabilities that may feel dangerous to the isolated. And should anything go wrong – a mistake is exposed, or he is scooped – he knows that he has a personal and professional safety net, a community that is likely to affirm his scientific integrity rather than shame him for being open to critique and competition.

7.10.1 Caring for data with closure and openness

There is an interesting difference between Luke and Melanie's accounts that relates to affective and caring dimensions of scientific knowledge-making: these themes have been studied in various settings (e.g. Kerr and Garforth 2016) and theorised in STS as "matters of care" (Puig de la Bellacasa 2011). While both Luke and Melanie spent formative years of their careers gathering large datasets, Luke did not portray the scientific labour and data itself in affective terms, where Melanie strongly evoked her labour of care. The narrative differences are of course influenced by the particular vulnerabilities that Melanie faced and Luke did not, and their personal – perhaps gendered – modes of expression in interview. Luke did, however, portray the affective labour of openness vividly. An intriguing interpretation would be that Luke cared for his data *through* his practice of openness, where Melanie cared for her data by protecting it – an extension of her care in creating it. Conceivably, Luke's internalisation of openness as an epistemic virtue meant that the best way to care for his data was by opening it to the world for feedback and correction, despite the anxiety this could cause, rather than caring for it alone. This lens might be applied to other accounts as well, both from my interviews and elsewhere, wherever affective or other forms of labour are undertaken in order to uphold openness as an epistemic virtue; and conversely, where data are cared for through an avoidance or delay of openness. The latter case may not only arise from a need to control risk, or due to practical barriers: other epistemic considerations might conceivably be cared for by delaying openness, such as a cautionary rigour (e.g. Madison: "I worry about overstating my own research as well. So I think I try and check myself a lot…").

7.11 Conceptualising scientists' relationships with data openness

Identifying vulnerability in scientists' accounts helped me to understand relationships with data openness as complex and entangled with the self: not simply a matter of a positive or negative attitude. Unease and fear were experienced by those who embraced data openness as well as those who resisted it. I have also shown that context is key in these differential responses, which are not just a matter of emotional processing. Circumstances differ widely: some projects are especially vulnerable to competition, where others come with unique materials, resources and networks. Some scientists occupy secure positions, where others are fighting to build or maintain their careers. Data openness can not only create vulnerability to scooping, but introduces the vulnerability of exposed scientific processes. Exactly which vulnerabilities apply in any circumstance is variable, and multiple dimensions of vulnerability interact: e.g. those who feel competitively secure may be better placed to accept other forms of vulnerability. A feeling of security and control can be gained in the short term – e.g. by protecting data until publication – but can also be a deep-seated disposition shaped influentially by e.g. experiences of community, or formative events, in early training periods. Interviewees tended to dissociate their sense of security or insecurity from the contexts that shaped them, leading to the surface impression that attitudes to data openness are individual preferences, or at most, alignments to disciplinary culture – rather than complex, specific responses to past experience, present circumstance, and future expectation.

To make sense of these interacting influences, I present in **Table 6** a summary of the contextual factors that seemed to promote or lower a sense of control or vulnerability, based on the interviews. I made these observations not only in association with data openness, but with the topic of Chapter 8, interpersonal openness, since the two categories evoked related forms of vulnerability. In the table I distinguish short-term factors (controllable by actions, e.g. waiting until after publication to open data); medium-term factors (often relating to current research, e.g. resourcing); and long-term factors (often relating to community and the shaping of the self, e.g. social support). Medium- and long-term factors were not clearly separable and played different roles according to circumstance.

Table 6 | Based on interviews, factors that seemed to be linked with higher and lower feelings of control or security, influencing whether data openness was experienced and/or tolerated as a vulnerability.

Higher control (lower vulnerability)	**Lower control** (higher vulnerability)
Short-term factors *– often actions/behaviours*	
Openness *after* credit for work is secured, usually via publication	Openness *before* credit for work is secured, typically pre-publication
Only sharing data that underlie a specific published work	Sharing extensive datasets that could underlie several publications
Only sharing specific data types, e.g. sequence data, that are routinely shared according to community norms	Sharing every type of data produced in a research project, including those to which community norms do not apply
Sharing processed or summary data	Sharing "raw" data that reveal more process and potential imperfections
Medium-term factors *– often related to current research project*	
Highly-resourced data production and analysis that are difficult for others in the same field to achieve	Low resourcing, e.g. smaller team size, less access to equipment or facilities, lower research budget
Privileged or exclusive access to a research setting, contextual knowledge, primary data, or methods	The research setting, contextual knowledge, primary data, or methods can be easily accessed or replicated
Low reliance on any one dataset; low level of labour per dataset; ease of obtaining new data	High reliance on each dataset for recognition; high level of labour per dataset; difficulty obtaining new data
Limited affective tie with data, or minimal caring labour required	Strong affective tie with data, or extensive caring labour required
Long-term factors *- often related to community, and shaping the self*	
Trust in fair recognition by community, especially if openness is pre-publication (related to community norms, e.g. Fort Lauderdale Agreement)	Lack of trust in fair treatment by community, expectation that others take advantage of data access to gain competitive advantage
Low competitive pressures, possibly due to extensive or exclusive collaborative networks	High competitive pressures and/or limited collaborative networks
Social support by supervisors, PIs, peers, and/or local and extended communities, especially where openness is discussed and practised in these relationships	Lack of social support by supervisors, PIs, peers, and/or local and extended community, or relationships in which openness is not discussed or practised
Secure, established career or reputational position	Insecure career or reputational position, or junior status

I propose that issues of control and vulnerability are fundamental, and predispose scientists to highly variable reactions to data openness. Two other dimensions also seem to have a key influence. The first (i) relates to the contextual ease or labour-intensiveness of making data open (see Section 7.7). The second dimension (ii) is the extent to which data openness is constructed as an epistemic virtue (see Sections 7.8 and 7.10). In **Figure 6**, I visualise possible interactions of these influences that condition scientists' likelihood of accepting or resisting a practice of data openness: this could serve as a starting point for further sense-making beyond this research.

I suggest that combined medium- and long factors shape a scientist's overall sense of control and vulnerability: a feeling of control predisposes acceptance of data openness practices, and vulnerability predisposes resistance. Certain data-specific factors (i) seem to modify these predispositions, but not radically: e.g. acceptance may be lowered, or resistance strengthened, by labour-intensiveness or lack infrastructural support. Reducing labour and adding technical support, however, cannot be expected to overcome vulnerabilities deeply ingrained in individuals or research settings. Individuals can gain control in vulnerable situations through short-term actions that limit their openness in specific ways (e.g. later timing of sharing). If they lose agency over short-term factors – e.g. due to policy requirements – these factors become vulnerabilities, and combine with other vulnerability factors. Scientists may also work *around* policy requirements in these situations to maintain control, e.g. by complying to a minimum required extent, or withholding contextual detail (see Section 9.3.2)

Finally, I suggest that when data openness is pursued as an improvement to knowledge and knower (i.e. [ii] as an epistemic virtue), the interactions above are altered. Control/vulnerability remains important and may shape an individual's likelihood of relating to data openness in this way to begin with. However, once such a disposition is internalised, it may change or even reverse the significance of other factors: a higher degree of vulnerability in the service of openness may be tolerated, and both vulnerability and technical labour may be re-cast as acts of care that reinforce an open epistemic identity.

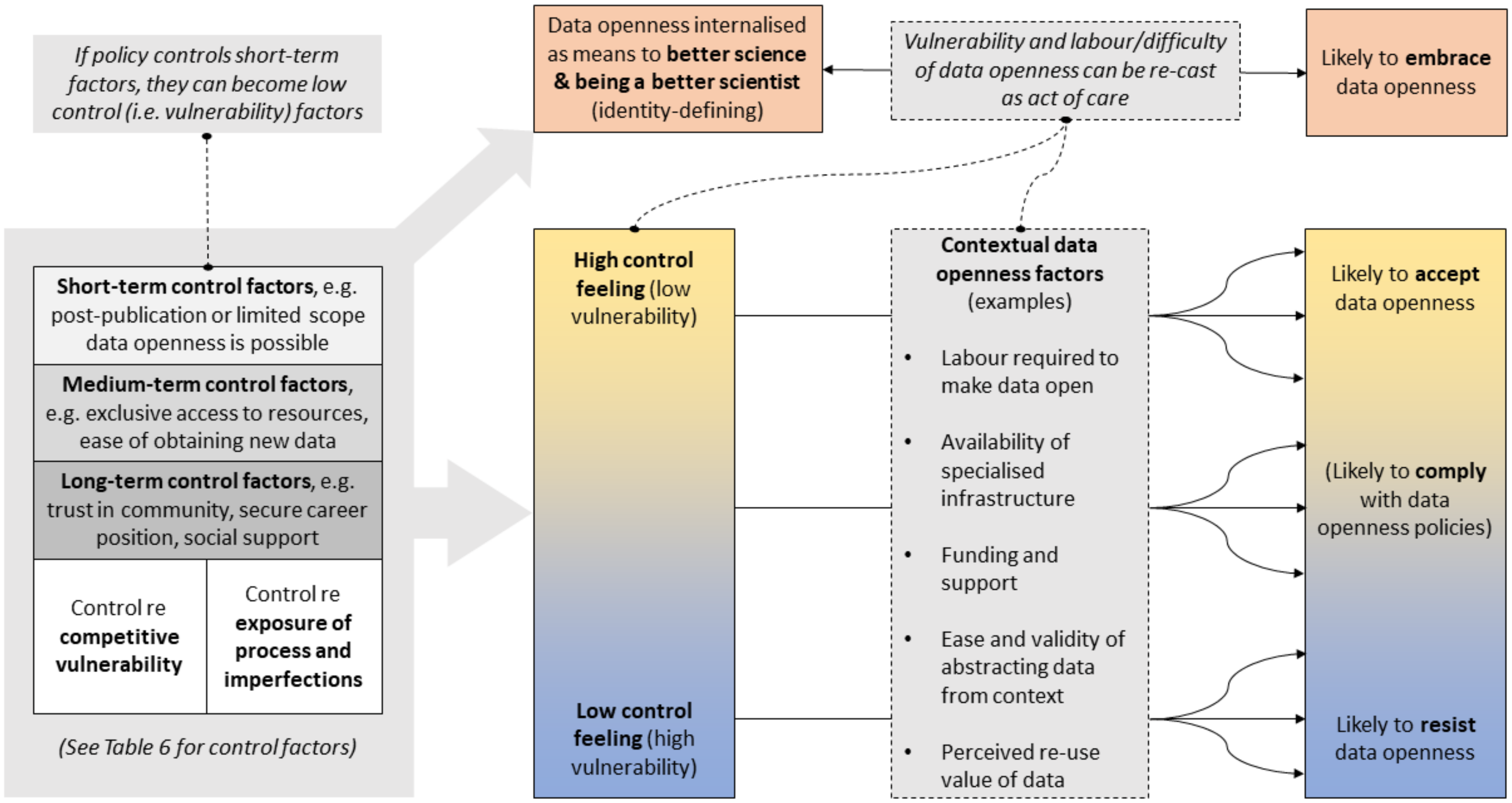


Figure 6 | Visualisation of how the conceptual elements identified in the observations made in this chapter may interact. See previous page for commentary.

7.12 Conclusions

In this chapter I have analysed the second of three significant openness categories constructed by scientists in interview: data openness. This category and open access were comparably salient in interviewees' accounts, and both were linked with a high salience in policy and advocacy discourse. The aim of this chapter was to make sense of the great diversity of practices, attitudes and emotions associated with data openness in scientists' accounts, and to identify common patterns.

One sizeable group of interviewees was familiar and accepting in their attitudes: they portrayed data sharing as routine and uncontroversial. The topic tended *not* to be a highly salient one for this group: although many brought it up, it was not usually their first openness-related interest. They tended to work with genetic or crystallographic data, and to be trained in the 1990s or before. I concluded that the concept of data openness, for this group, was probably shaped by digital data sharing discourses and norms dating to the 1980s and 1990s and applying specifically to sequence and crystallographic data (HGP Information Archive 1997; Strasser 2019). This was not a holistic orientation towards data openness, as is encouraged in more recent discourses.

A second group of interviewees reacted with a distinctive unease and defensiveness to data openness. The topic was salient for this group, and raised early on. They presented a wide range of objections that brought together the epistemic and the personal. Data openness was variously seen as lacking in integrity due to its abstraction from context; not useful for recipients; excessively laborious; and/or competitively risky. However, beyond first impressions, it emerged that all of the interviewees in this category would be happy to share data under particular circumstances: when the data was associated with a particular publication, and – crucially – when that publication credit, was secured. Despite the potential acceptability of data openness to these scientists, they expected to have to defend themselves, and their scientific interests, against expansive or one-size-fits-all "open" requirements. They did not expect to be shown empathy or given contextual accommodations to ease data openness in practice. Their research contexts likely shaped their objections: these interviewees tended to produce data that were less standardised and digitised, without links to decades-old data sharing norms.

A third, sizeable group of interviewees showed a principled or passionate attachment to data openness that could be distinguished from familiar acceptance. Data openness was highly salient for these interviewees. The principles or passion driving their attachment were varied and included connections with open source software culture; transparency and reproducibility, and public duty. Many of these accounts framed data openness as an epistemic virtue. Interviewees in this category tended to view openness as expansive in a positive way: applicable to multiple data types as well as articles, methods, code, and more. In this way, the principled and passionate group of interviewees clearly reflected the influence of twenty-first century open science movements: their motivations were as diverse, their visions as expansive, and their attitudes as enthusiastic, as these movements. Moreover, nearly all of these interviewees had trained since the 2000s. This suggests that open movements are gaining deep traction with more recently trained generations of biologists, even though this trend has not yet stood the test of time. Moreover, data openness seemed to inspire this generation more profoundly than did open access to articles. However, the embrace of data openness was not context-free. Where those who reacted uneasily to data openness assumed the "worst" in terms of the difficulty and risk, principled and passionate interviewees assumed the "best": commonly, that data release would occur after securing credit for work – sometimes after a significant embargo period. I identified a feeling of control and security (or lack thereof) as an underlying influence on ease or unease with data openness, and linked it with contextual factors such as trust in one's community; resourcing; and career security.

A final set of examples revealed a common thread between uneasy and enthusiastic relationships with data openness. Several interviewees reported worry or fear about data openness, but subsequently or simultaneously embraced it. Luke's account helped me to make sense of data openness as a form of vulnerability: a state of personal exposure, as well as an exposure of data, that entail susceptibility to emotional and professional harm as well as growth. The two main vulnerabilities scientists faced were competitive threats – the risk of scooping – and exposure of the details of scientific process. Both also carried opportunities for growth: e.g. through collaborative rather than competitive interactions; feedback; and corrected mistakes. In both cases, the scientist and their data were closely entangled in a relationship of care. This could mean – as in the case of Melanie's PhD – that as data came to represent a labour of love, the prospect of its use by a competitor was felt to be both emotionally and professionally threatening. But when openness was embraced as an epistemic virtue, as in Luke's case, it seemed to become part of a labour of love: a way to care for data – and to cultivate one's scientific self – that may be difficult, but is ultimately worthwhile. Context – particularly a sense of security – made a crucial difference between these two scenarios. I identified a range of influencing contexts in **Table 6** and conceptualised their interactions in **Figure 6**.

An overarching finding from this chapter is the extent to which context matters in scientists' relationships with data openness – and not only in the obvious ways such as ethical or commercial sensitivity, anonymity, or even generalised "disciplinary norms" (HEFCE et al. 2016). It also matters in ways that are unseen even by scientists themselves: complex interactions between the epistemic and competitive positioning of a research project; the networks and resources of a research group; characteristics of social relationships on individual and community scales; and the career-long experiences of each scientist that have helped them feel secure about opening their data, or have warned them to be protective. It does, however, appear that more recently trained scientists are profoundly influenced by twenty-first century open science agendas, an influence distinct from earlier discipline-specific data sharing norms. Some are internalising data openness into their scientific identities, suggesting a potential for deep and lasting bottom-up culture change. However, scientists who are not supported through the vulnerabilities of data openness may remain isolated and disadvantaged in cultures that increasingly value data openness, both bottom-up and top-down.

It remains an "open" question whether polarised reactions to contemporary data openness will settle into a familiar acceptance, as have early data sharing norms. Strasser (2019) suggests that this is likely: he argues based on historical precedent that a general acceptance of data openness only develops when top-down rules are enforced and eventually become norms. Only time will tell, but my study suggests that contemporary data openness could be different. Its expansive character within open science lacks the boundaries which may have made earlier data sharing acceptable; and norms that support openness – e.g. etiquette about use of unpublished data – would need to be developed in many, varied epistemic cultures. Moreover, contemporary data openness is aimed not only at resource-building, but also at checking integrity: a far more vulnerable form of openness. However, an increasing embrace of data openness as an epistemic virtue may provide grassroots momentum, and a tolerance of "open" labour and vulnerability, that is difficult to create through top-down intervention: a changing of "hearts and minds" with which open access, by comparison, has struggled. The topic of the next chapter, interpersonal openness, adds understanding about the support or privilege that might be necessary in scientific communities for vulnerable forms of openness to flourish.

Chapter 8 | “Feeling free to talk”

Interpersonal openness

Adam was the first person I interviewed for this research. When I asked him about “openness in science”, his responses were confident and relayed to me at speed, but reflective. Perhaps it was the time pressure of his executive-level academic diary; there also seemed to be a focused concern and care about the meaning carried by those words. My question seemed to fit into a narrative he understood implicitly. Perhaps because the conversation flowed smoothly, I didn’t immediately realise that he was saying something unexpected to me. Something that is not typically articulated in policy and advocacy visions of “open science”. I heard it more clearly once I had listened to several other scientists talk this way.

“...when you ask me: what do I think of when you say openness?
I think about the many scientists I have known
and the way they approach how they communicate their work
and when they communicate it
and how communicative they are
and what sort of discussions you can have with them.
That's what I think about.”

Adam, professor and lab leader, parasitology

8.1 Introduction

This third interview-based chapter is distinctive because it features a category of openness without foundation in advocacy and policy discourses. Nonetheless it was raised by scientists with a prominence and consistency comparable to open access or data openness. Interviewees had no shared, concise name for it. I have called it “interpersonal openness” to capture its clearest feature: it occurs in personal interactions between scientists. Typically these interactions are face-to-face rather than online, although they include socially bounded online settings such as email. They range in intimacy, interactivity, and formality from one-to-one conversations, to conference presentations, although they tend towards the intimate, informal and interactive. They occur in any space in which scientists mingle and communicate about science with those beyond their immediate research group: meetings, conferences, cafeterias, visits to other laboratories, and social events. Interpersonal openness was distinctive partly because of its mundane, micro-scale qualities: it was integrated with everyday professional–personal conduct, and governed not by top-down policies but by the instincts or intentions of individuals, and the norms of small groups. It emerged as a mode of openness entirely owned and operated by scientists.

Interpersonal openness had characteristic features that emerged repeatedly in scientists' accounts. Firstly, it meant communicating about science-in-the-making that had not yet been published, or perhaps even funded: current work, recent data, and especially *ideas*. This made it an acutely vulnerable, but also valuable, form of openness. Secondly it ideally meant *talking freely* or *feeling free to talk*: a feeling of honesty, freedom, flow, and excitement. Thirdly, interpersonal openness was associated with *reciprocity*: ideas were shared generously, but in the long term there was an expectation of epistemic, social, and professional rewards. Finally, *trust* was key: it both enabled and resulted from interpersonal openness, and structured the social networks through which privileged scientific information flowed. I explore and evidence these features of interpersonal openness in the first part of the chapter. The middle sections of the chapter explore interpersonal open-*ness* as a spectrum of intertwined attitudes and practices that is never entirely open or closed. A significant group of interviewees upheld interpersonal openness nearly unconditionally, arguably as an epistemic virtue with a close connection to their sense of scientific self. Because of the vulnerability it entailed, however, the majority of interviewees described making contextual judgements about how much to share – and especially, with whom. And for some, protecting ideas was the norm.

The final chapter sections explore factors that enable and constrain interpersonal openness, many of which have also featured in relation to data openness. These factors include both circumstantial and systematic forms of security and insecurity. In the conclusion I reflect on the insights that interpersonal openness offers for open science movements and for my research questions. To begin, I describe the salience of interpersonal openness in my interviews with scientists.

8.2 The salience of interpersonal openness in interview

Of the 40 scientists I interviewed, just under half [17] referred to interpersonal openness in response to my initial prompt: "What first comes to mind when you think about openness in science?". Just over half [21] brought it up spontaneously at some point in the interview, and most [36] showed recognition of this mode of openness, whether they saw it as something to promote, negotiate, or avoid. These figures were somewhat lower than those for open access and data openness [respectively: 19 and 19 raising at the initial prompt; 31 and 28 raising spontaneously at any time; and 40 and 37 showing familiarity]. However, this was partly because those raising the topic tended to embrace interpersonal openness; those who recognised its dynamics while viewing it with ambivalence or caution were less likely to nominate it as a form of "openness". **Figure 7** shows how spontaneous mentions of the three main categories of openness overlap in interviewees' accounts. It shows that interpersonal openness is somewhat set apart: a significant group [8] mention interpersonal openness but not open access or data openness in their first response, whereas there is a tendency for the latter two categories to be mentioned together.

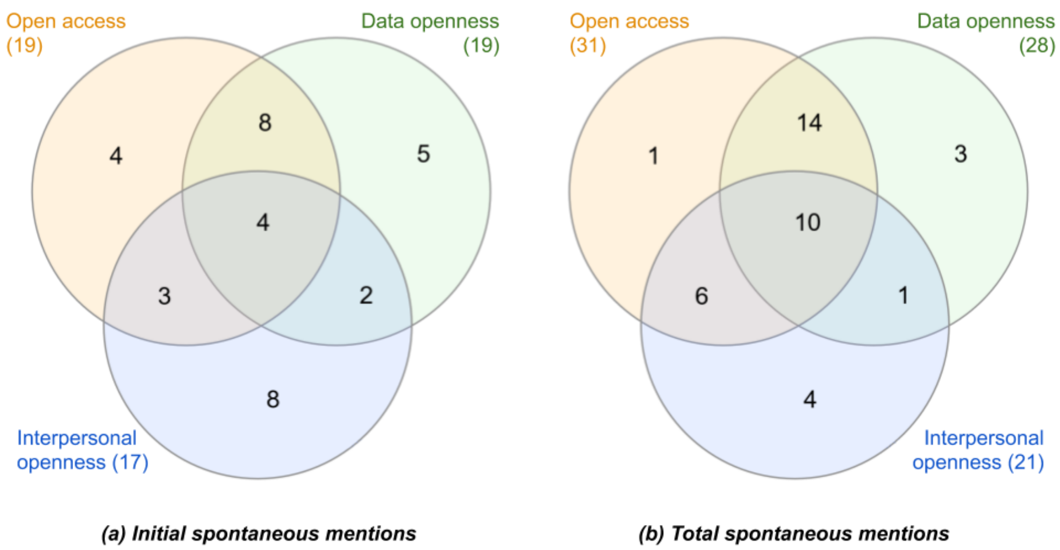


Figure 7 | Venn diagrams showing how the three main categories of openness overlap in interviewees' accounts. The counts are of interviewees, according to whether they mention a given category of openness either (a) in their initial response, or (b) at any time prior being directly asked. Those who did not mention any of these three categories spontaneously are not shown, but are: for (a) 5, for (b) 1. Other categories also arose, but were less salient than these three. Made with a tool by Heberle et al. (2015).

8.3 "Openness in science" versus "open science"

An important context to this chapter is that my choice of words in interview seemed shape the emergence of interpersonal openness as a salient category. I asked interviewees about "openness in science" because I reasoned that it would be less prescriptive than "open science", a common phase in policy and advocacy discourses. This proved to be the case: there were a broad range of responses from inside and outside such discourses. But as the interviews progressed, I observed ways in which "openness in science" was not a neutral prompt. As well inviting many interviewees to talk about topics like open access, "openness in science" seemed for many to actively resonate with the phenomenon I now call interpersonal openness. I initially hesitated to make a feature of this, worrying the category might somehow be too loose, obvious or generalised. I was convinced otherwise by its quantitative salience, almost uncanny qualitative consistency, and captivating quality in many accounts. Those who raised it often spoke as if encountering a rare opportunity to articulate a deeply-held principle or hope. Several took pains to distinguish it from open access or data openness as a more profound form of openness. Lest I conclude otherwise, Richard[124] closed his interview with this reflection on "openness", which he had defined interpersonally:

> *"My feedback to you is that [...] I think that the general concept of openness is much broader, and much more important, than the concept of open access itself. I think open access is a little thing. Openness is much more important."*

Thus although interpersonal is at face value a broad, common-sense category, I began to take seriously its value to interviewees, and to explore its meaning.

I also began asking interviewees about their impressions of the difference between "openness in science" and "open science". The outcome was illuminating, and confirmed that the former tended to connote interpersonal or behavioural openness; and the latter, "open science" movements, practices, and policies. There were several other enlightening thematic distinctions. For example, "open science" tended to be associated with an extreme, ideal, or complete state of openness, whereas "openness

[124] PhD 1980s; interview: US, Mar 2018.

in science" was more associated with a process of becoming open, a spectrum, or openness possible within current realities. Kate[125] captured this: she described open science as "when you're there, when you've managed, when you're achieving science that's more open than what I'm currently achieving". Openness in science she characterised as "when you're trying your best".

"Open science" was also associated with narrow or specific concepts or technologies, where openness was linked to an expansive set of values and behaviours. Greg[126] put it as follows:

> *"I think that "open science" sounds like a specific set of rules or, a specific…maybe for lack of a better word, movement. Right? I mean it almost sounds like CC BY[127], or something like that, right? […] "Whereas "openness" is a perspective. It's an attitude. It's a set of behaviours."*

Some interviewees connected open science with information outputs such as articles or data, where openness was associated with embodied behaviour and internal experience or identity. This meant a difference between performing open practices, and *being* open in an intentional, principled way. Erin[128] commented:

> *"…open science to me is: here's the data […] it's like something that you have to do. […] Openness suggests a process in which people are actively engaged and willing to be open […] you can still be someone who's not an open researcher but that still puts their data online and still publishes their papers in an open access way but doesn't ever engage in that openness."*

These findings have implications for the construction of "open" or "openness" as an epistemic virtue, and for scientists' engagement with these concepts: I discuss this further in Chapter 9. In sum, my choice of words was a fortunate methodological contingency because it gave interviewees space to talk about interpersonal openness: a less prescribed, more embodied form of openness that happens "when you're trying your best". In the following sections I characterise interpersonal openness as constructed by those who spoke passionately about it.

[125] PhD 2000s; interview: UK, Jun 2018.
[126] PhD 1990s; interview: USA, Feb 2018.
[127] CC BY is a licence widely asserted as a standard in open movements (see Section 5.2).
[128] PhD 2000s; interview: UK, Aug 2018.

8.4 Characterising interpersonal openness

Interpersonal openness emerged from the very beginning of my interview process. Adam's[129] first thought about openness was not of articles or data, but of scientists and how they behave differently to one another: "how open scientists are at discussing their own work", as shown in the quotation that starts this chapter. Adam was well aware of open access, and even of open notebook science, but highlighted interpersonal openness first – and tended to return to it after discussing other modes of openness. He portrayed interpersonal openness as a spectrum:

> *"…you see a huge variation in that, you see schools of thought where you don't tell anybody anything until you've published it, you don't give talks about it in public, you don't discuss it with other scientists, you go as far as instructing your students and people not to discuss it elsewhere, because it is really important to you to be first […] and get maximum credit for it […] and there are those at the opposite extreme who will talk about everything."*

Here Adam drew attention to features that would emerge repeatedly: pre-publication timing; settings such as public talks and private conversations; and group dynamics in which research leaders set norms for their students. He embedded qualitative, relational features in his characterisation: scientists' *approach* to communicating; *when* they communicate; *how* communicative they are; *what sorts of discussions* you can have with them. Adam saw himself as open ("I'm someone who does communicate very freely"), and framed this as a virtue with both moral and epistemic overtones, closely tied to his sense of professional self and identity. He was motivated by a Mertonian-like "collegial, collaborative" view in which "we are working side by side, shoulder to shoulder with our fellow scientists, pushing back the barriers of science". Adam described open interactions as enriching for both science and scientists, yielding "input and insights".

[129] PhD 1980s; interview: Australia, Jan 2017.

Jenny's[130] prioritisation of interpersonal openness was evident from the outset, when she made a distinction between "openness in publication or openness in science". This implicitly placed open access outside of "science", and centred an interpersonal "openness in science". This meant talking to other scientists in situations like meetings, and being completely frank – "no holds barred" – about current research and ideas. In an illuminating narrative, she described how she had developed this approach almost as a matter of necessity, but maintained it as a matter of principle. She characterised her formative career environment, in marsupial genetics, as relatively free from competition, making openness less risky:

> *"I've been on the edge, most of my life, I haven't been sort of in the mainstream human genetics community or the mainstream mouse genetics community, and I'm generally not at risk. So I'm probably much more likely to be open about what I'm doing, what I'm thinking, because I don't have the same sort of competition as human geneticists and mouse geneticists."*

In this setting, Jenny developed a habit of interpersonal openness that not only felt "easy", but was advantageous for her career as it raised awareness of her non-mainstream research organisms and contributions:

> *"...I told my students and post-docs, you know, talk to anybody and everybody you possibly can because we need input and we also want them to know that what we're doing is actually relevant to what they're doing."*

In this way, what may have begun as a personal intuition became a strategy for enriching her science, into which Jenny initiated her research group. She decided to continue this approach as an explicit "policy" under more competitive circumstances:

> *"I actually have made a policy that generosity pays. And for me and I think for our whole group, that has generally been the case, that we get much more by soliciting other peoples' opinion on what we actually care about and what we're actually doing, than we lose [...] It only became a spoken policy much later when I think we observed that the world around us wasn't necessarily as generous with their sharing of ideas and data than, than we were."*

[130] PhD 1970s; interview: Australia, Nov 2017.

Jenny did not underestimate competitive risk – her group had "been burnt, two or three times", but she had learned from formative experience that openness was rewarding, productive, and led to reciprocity: "discussions that are very real".

Ernie[131] also began by distinguishing "openness in publishing" from "openness in how you interact with others". He delighted in the latter:

> *"…one of the things I enjoy most about science is going to meetings and visiting other places and talking to people about my work and having them talk to me about their work. And people who are constantly playing a game, that you know - you show me yours and I'll show you mine, and I'm only going to show you - I'm going to figure out how much what you told me is worth and I'll show you that much - I mean that's a drag. So I cultivate relationships with people who I trust and then we just talk freely about ideas…"*

Once again, it was the exchange of ideas in face-to-face settings, around which interpersonal openness was evoked. Ernie placed emphasis on trust: it was the foundation for a "free" exchange, a condition upon which Adam and Jenny appeared less reliant. This may be linked with the "very competitive and not always particularly friendly" communities in which Ernie was embedded; Adam in contrast occupied a "supportive" community, and Jenny had experience with both "extremely nasty" and "incredibly friendly" communities. In a competitive environment, Ernie's interpersonal openness involved the creation of trusting spaces and relationships in which the risk of sharing was mitigated. His openness was with selected people in order that the interaction itself could be unconditionally open and free: a reciprocity that was, like Jenny's, generous rather than calculated. Palpable in Ernie's account was joy: interpersonal openness was both personally and epistemically fulfilling, underpinning his attachment to science and his identity as a scientist.

Accounts like these were as common as they were compelling. Interpersonal openness was portrayed in positive terms as the most obvious, important, or relevant form of openness in science by an additional nine interviewees, as follows:

[131] PhD 1970s; interview: UK, Oct 2017.

- Andrea's[132] first thought on openness was "talk, talking to people, sharing."
- For David[133], "openness is sharing information before the story is complete" when "meeting people and just talking about what we're doing".
- For Mark[134], "openness is the generous sharing of ideas".
- For Julia[135], openness meant "talk[ing] about your work as it develops".
- For Lena[136], the "most prevalent" form of openness in science was "openness about what I'm doing, to others, at conferences".
- For Miriam[137], "openness is more right at the beginning and the conceptualisation of the work […] So open, in terms of collaborating".
- For Nicole[138], openness was about collaboration, and "the ability to freely talk about what you're doing, and idea share…".
- For Richard[139], openness meant "the extent to which individuals within a community, research groups within a community, share their results informally, with each other".
- And for Steve[140], "what first comes to mind would be feeling free to talk at meetings about work that hasn't been published yet".

In addition to Adam, Jenny, Ernie, and the group above, were interviewees who raised interpersonal openness spontaneously but with less emphasis or in negative terms, and interviewees who did not raise the topic but spoke about it with familiarity and sometimes enthusiasm. The following sections are an exploration of the features of interpersonal openness that arose repeatedly in interviewees' accounts.

[132] Mid-PhD; interview: UK, Oct 2018.
[133] PhD 1970s; interview: UK, Dec 2017.
[134] PhD 1980s; interview: UK, Feb 2017.
[135] PhD 2000s, interview: UK, Jun 2018.
[136] PhD 2000s; interview: UK, Apr 2018.
[137] PhD 2000s; interview: UK, Jan 2018.
[138] Mid-PhD; interview: UK, Oct 2017.
[139] PhD 1980s; interview: US, Mar 2018.
[140] PhD 1970s; interview: UK, Nov 2017.

8.4.1 Pre-publication ideas: valuable and vulnerable

Publication was a highly salient temporal boundary in relation to interpersonal openness, as it was for data openness. Being open interpersonally meant sharing *before publication*: "before the story is complete" (David); as work "develops" (Julia); "right at the beginning" (Miriam); and "work that hasn't been published yet" (Steve). For interviewees embracing this practice, sharing new, unpublished work was seen as the purpose of a conference presentation. This view was so strongly held by some that they viewed presentations of published work – or work holding back unpublished details – as professionally unacceptable, disrespectful to the gathered audience, and epistemically obstructive. Both Adam and Richard characterised their response as deep "irritation", and Mark also expressed frustration. Richard commented:

> *"I just think that is not good behaviour [...] it serves nobody's interest other than the person who is talking about it to try and advance their own personal ends [...] it sucks up time in a public setting. And fails to actually serve the purpose of communication. No real valuable information about biology is communicated. [...] I don't like it, I avoid it, I never do it myself - and I call it out whenever I see it - and I try and encourage all of my students and colleagues to not do that."*

Conference presentations were thus key settings in which interpersonal openness was established as a pre-publication behaviour, as well as moralised as an epistemic virtue.

Beyond this, interviewees tended to associate interpersonal sharing with informal, interactive settings, and the exposure of information that might not be ready to present at a conference. In short, much of interpersonal openness was about sharing science-in-the-making: "what I'm doing, what I'm thinking" (Jenny). This included not only new data, but *ideas*: abstract, malleable epistemic elements that may be a long way from manifesting as a research process or paper. Miriam distinguished these stages of research and made it clear that early, interpersonal openness was particularly meaningful for her in comparison with open publishing:

> *"...I think publishing is the end of the work [...] I'd like to think that openness is more right at the beginning and the conceptualisation of the work [...] So open, in terms of collaborating [...] I would consider openness being that, you know, in science where you can share ideas frankly..."*

Interviewees who embraced interpersonal openness often seemed to do so because they found it valuable and enlivening to discuss science that is not yet settled: to expose their conceptual processes to feedback in in a rewarding, interactive, creative, collaborative, social setting. Some even directly characterised this as a definitive part of doing good science and being a good scientist. As Adam put it:

> *"...to be open and to communicate and to have a free flowing exchange of ideas and criticism is much more healthy, and environments of how science – the best science – is done."*

Here once again, interpersonal openness – of an interactive, intimate kind – is characterised as an epistemic virtue.

Ideas were portrayed as extremely valuable in two interlinked senses that lay at the crux of their role in interpersonal openness. They were materially, professionally valuable because they were foundations for future grants, data, and publications: the currencies of academic science. Ernie told me: "The fact that I can convince people that my ideas are better than other people's ideas, is how I get grant funding." For this reason, Ernie saw it as important to be careful with his ideas: to share them with trusted parties, but not publicly or online. Mark portrayed his wealth of ideas as so materially advantageous to his laboratory that he need not worry about being scooped. He embraced a "generous sharing of ideas" in interpersonal settings:

> *"It's partly because I have the resources, but also a couple of my ideas are sort of, big, in terms of their number, and types of experiments that they open up, and so I can have my ideas, share them, and people will be able to pick up and run with them and it doesn't matter, because my students are doing experiments that nobody else can do, because we're so many, we're quite a few years ahead [...] in some ways ideas are easy, but they're also still valuable."*

In this way, ideas not only held promise of future career security, but could amplify and reinforce existing professional advantage.

As well as their material value, ideas were portrayed as acutely valuable personally. There was an implication in multiple interviews of a deep link between ideas and scientific identity: ideas were expressions of an intimate, imaginative epistemic self, and the generation of ideas gave meaning and purpose to doing science. Ernie spoke directly and eloquently of their personal value:

> *"What I own, and nobody else can have, are my ideas. They come from inside my head, I generate them through the happiness and misery of my life, that's where the ideas come from. When I graduated university, I was going to be an artist I wasn't going to be a scientist. But I found that the process of designing beautiful experiments actually satisfied the same creative urge in me, but I do have sort of an artist's view of ideas - that ideas are really critical, and they're unique, and they come from your soul. And other people don't own them. So even in a totalitarian society - you just can't make people share ideas. I'll share my output. I'm happy to share my output, but my ideas are mine."*

Ernie's impassioned language underlined the value of ideas: he placed them at the core of his internal life, his "soul", as expressions and nourishments of a self that is both artistic and scientific. His "output", in comparison, was far less personally entangled. Ernie also portrayed idea generation as essential to his scientific existence:

> *"If I can't have my own ideas, then I should retire. 'Cause what science for me is about is the creative process of having ideas, and then seeing if you can figure out a way [...] to see if they're true or not. That's what the fun is."*

Thus for Ernie, sharing ideas was not only materially risky, but existentially vulnerable: a baring of his soul. This aligned with the trust-based selectivity of his interpersonal openness, and the strength of his aversion to online, open notebooks. The suggestion that he might be forced to reveal an interior domain of ideas was felt as a violation of his most basic freedoms. Ernie recalled devastation when one of his "best friends" and closest scientific confidantes took advantage of his trust by writing a grant to compete with him: "I felt that I was stabbed in the back". This vulnerability also aligned with his joy in interpersonal openness in a trusting context.

Cedric[141] also aligned his identity and success as a scientist with the production of ideas: "I have a very poor memory, which I argue is why I've been, to some extent, successful in science - because I don't get bogged down by the facts…I can have ideas".

[141] PhD 1960s; interview: UK, Feb 2018.

He drew on an essay by the philosopher Isaiah Berlin, *The Hedgehog and the Fox* (1953), to evoke two intellectual personae defined by their approach to ideas:

> *"…the hedgehog has one big idea, and pursues it relentlessly. And the fox has many ideas, and pursues them as and when, and follows his or her nose, and finds the whole world exciting."*

Cedric identified both himself and influential mentors and colleagues as foxes, and placed great value on prolific idea generation, despite an assertion that hedgehogs are more traditionally successful in science ("I think of my friends who focused more and more on one topic, and have won bigger and bigger prizes"). He linked these personae directly with interpersonal openness and closure, suggesting that hedgehogs, with a limited set of highly developed ideas, "keep it all to themselves", whereas foxes, with their profusion ideas, "want to share the stuff they're doing because they find it all so exciting and interesting". Nonetheless Cedric, like Ernie, portrayed ideas as part of an intimate interior domain that cannot be shared involuntarily:

> *"…in the end, we all keep our ideas to ourselves. Until you can download what I'm thinking, you have no idea what's gone on in my head, what's going on in my head, and what I'm storing there."*

Here the generation of ideas again emerged as an experience and aptitude anchoring a sense of scientific self, and connected with an embrace of interpersonal openness.

Other interviewees also alluded to the generation and sharing of ideas as deeply constitutive of scientific identity. For Melanie[142], sharing ideas interpersonally was so enlivening and essential to a sense of scientific purpose that she accepted a concrete risk of her ideas being used by others:

> *"I mean I tend to be super open. 'Cos, I just feel like […] not expressing what I do, and ideas and sharing - that's my favourite part of science, so if I didn't do it I don't think I would stay in science. But I mean I've been burnt on that several times, where you know, you talk about ideas so freely then someone else does them. But whatever, I'll come up with more ideas. I'm not too worried about that. I always think if I get to the point where I can't share them, then I would leave - just leave science. So I tend to be an over-sharer, probably. But I just get so overexcited, it's so fun to talk about the process, and new*

[142] PhD 2000s; interview: UK, Jul 2018.

results, and interesting ideas and stuff. So I tend to be very open with colleagues. And I encourage my students to be as well."

Again, ideas – and particularly the ability to share them – sustained Melanie's scientific identity. She inhabited the fox persona described by Cedric, and echoed Mark's notion that ideas can be both valuable and easy. Lena[143] similarly felt able to be open because "we have plenty of ideas". She also spoke of ideas as outcomes of interpersonal openness – not just conversational inputs that that are critiqued, but conceptions that emerge "non-linearly" from interaction:

> *"...if you don't talk to people, you don't get as many ideas [...] I get ideas from very, very different kind of people and conversations. I start a conversation with someone, about a scientific project, and end up getting something useful for a completely different project. So it's non-linear, how you get useful information."*

This generative quality of interpersonal openness was implied in other accounts. Julia suggested that it could occur partly within the self, through the act of interpersonal sharing: "I'm quite extroverted, and I realise that I develop a lot of my ideas by talking". The notion of "bouncing ideas off" others in conversation was also part of the generative quality of interpersonal openness.

Collectively, these accounts show that the value and vulnerability of idea-sharing are closely interlinked. As abstract, malleable concepts, they are often far upstream from being realised as academic currency in the form of a grant or publication. As a result their authorship is not established and they are vulnerable to seamless acquisition by others. In this sense ideas were more vulnerable than unpublished data: they are further from being realised, and less traceable to an origin. For the same reason, sharing ideas was seen by some as especially valuable: an opportunity for imaginative, critical input at a formative stage. Ideas – and the process of generating them – were also closely linked to scientific identity, as expressions of a creative interior self. This meant idea-sharing could be akin to baring one's scientific soul – but its rewards could be great for the same reason: an opportunity to share an exciting, purpose-affirming process of idea generation with others. This observation connects with Parker and Hackett's (2012) notion of "hot spots and hot moments": "places and times

[143] PhD 2000s; interview: UK, Apr 2018.

of high emotion, creativity and performance where transformative science is done" (p. 39).

8.4.2 "Talking freely" and "freedom to talk"

Interviewees used similar language in descriptions of interpersonal openness, the clearest example being repeated use of "talk" and "free": "feeling free to talk at meetings about work that hasn't been published yet" (Steve); "freely talk about what you're doing, and idea share..." (Nicole); and "just talk freely about ideas" (Ernie).

8.4.2.1 Talk

"Talk" perhaps coincidentally described both of the main contexts of interpersonal openness: conference presentations and face-to-face conversations. It was also employed by some as a shorthand for open sharing, as if to suggest that there is something honest or revealing about talking in comparison to less embodied or more formalised modes of scientific communication. For instance, Andrea spoke first of "talk, talking to people, sharing" when I asked about openness: she saw "talk" was a means to compensate for the omissions or obfuscations of published science, for example in methods sections. Jenny emphasised the value of meetings at which "...people will talk about what they're doing and what they're thinking, that there'll be discussions that are very real". "Very real" suggests talk without artifice or inhibition: talk that cuts through noise, and perhaps has the intimate, creative, interactive qualities evoked in the previous section. Jenny described an admired colleague's response to competitive environments as "talking more, not less". The implication that talk should be open and honest was also present in the moral conviction that conference talks should be a complete and transparent communication of unpublished work.

8.4.2.2 Free

The use of "free" to characterise interpersonal openness was also recurrent. It implied expansiveness, spontaneity and sincerity in disclosures and interactions, as well as communicating a feeling of agency and unshackling. Jenny used the wrestling metaphor "no holds barred", implying a rule-free and perhaps intensely interactive and exposing approach. "Free" was also associated with "flowing": Adam advocated openness as a "free flowing exchange of ideas and criticism" and Mark spoke of a "free flow of ideas at conferences" as essential. "Flow" again suggests a quality of communication that is unrestricted: not held back, partial or polished, but lively and spontaneous. It also evokes the psychological concept of a "flow state" or being "in the zone" (Csikszentmihalyi 1990) which could describe the enjoyable, absorbing process of idea generation. When ideas are shared and this process becomes a meeting of minds, "flow" suggests not only an inhibited conversation, but one in which social and epistemic boundaries become less distinct, and ideas can mix and re-emerge anew in unpredictable ways, free from preoccupations about ownership or calculations about a degree of openness or closure. In some accounts, "free" was used to emphasise an unshackling from competitive pressures: not just freedom *within* talk, but freedom *to* talk. Ben[144] commented that at early stages of research, "...you're a bit careful about who you share your ideas with. In a free way". The openness that Steve[145] evoked was a liberty to talk about unpublished work, with an implication this is not always the case; Nicole's[146] aspiration to "freely talk" about what she was doing was a response the competitive culture of her field, in which she did not feel that freedom.

8.4.3 Reciprocity: "generosity pays"

Embedded in idea-sharing, free-talking openness is a dynamic of reciprocity. Interviewees spoke about an "exchange" of ideas, and some described this further. A common feature was that reciprocal openness was not expected to be transactional,

---

[144] PhD 1990s; interview: UK, Jan 2018.
[145] PhD 1980s; interview: UK, Nov 2017.
[146] Mid-PhD; interview: UK, Oct 2017.

i.e. repaid in an exact and proportional way, in a single interaction or relationship. For example, Ernie disliked "playing a game" in which you "figure out how much what you told me is worth and I'll show you that much", preferring instead to establish trusting relationships so that he could share freely, without calculation. Interviewees characterised this reciprocity as "generous". Taking an open approach was expected, on average, to prompt openness from others, making it more than worthwhile in the long term. Occasional exploitation of openness was seen as possible or even probable, but tolerable in light of overall benefits. Jenny's "generosity pays" policy was based on this premise: "we get much more by soliciting other peoples' opinion on what we actually care about and what we're actually doing, than we lose". Cedric's equivalent of 'generosity pays' was "enlightened self-interest":

> *"The more open you are with other people, then hopefully the more open they will be with you. And the more you will both be able to move forward, in a meaningful and interesting and exciting and innovative - which is what science is all about in the end - way."*

Both of these concepts mix virtue – "generosity" and "enlightenment" – with ego-centred interests: blending of what is good for others, and for science, with what is good for oneself. This view was not only asserted by established scientists. Olivia[147], a PhD student, told me:

> *"…people have this sort of element of: oh, if you're not going to give me anything then I shouldn't give you anything and keeps this very closed off mentality within science. But actually if you go out there and say: oh you can have this, then quite often you get something back. Even in ways that you don't expect I think. Other doors open to you. That's a big part of openness."*

David[148] asserted these principles with particular directness: "if you give, you get". He saw it as important to "give away" ideas without an expectation of reward, as this upheld collegiality and a "common good" – *and* because it was personally rewarding in the long term. If you are "somebody who shares ideas", you can expect to receive them: the reciprocity of openness was connected with personal character.

---

[147] Mid-PhD; interview: UK, Jul 2018.
[148] PhD 1970s; interview: UK, Dec 2017.

Cedric's choice of words echoed Kohler's (1999) characterisation of the open, co-operative moral economy of early twentieth century *Drosophila* research: "The custom of exchange mixed altruism with enlightened self-interest, bringing substantial benefits…" (p. 254). Although "enlightened self-interest" is an established phrase, Cedric's use of it carries the suggestion that interpersonal openness operates as a moral economy (a trade in "symbolic more than economic values", ibid., p. 249). Interpersonal openness as characterised by interviewees certainly shares more with model organism moral economies (see Section 4.9) than with twenty-first century "open science". Twentieth century model organism communality typically operated within bounded social settings, depended on reciprocity and trust, and eased the sharing of ongoing work (Kelty 2012). "Open science" foregrounds no-strings-attached, universal sharing. It is possible that the prominence of interpersonal openness in my interviews is promoted by the cultural influence of model organism communities in biological science, as generations of biologists have moved through such cultures. But for interviewees, interpersonal openness tended to operate as a diffuse moral economy through networks of like-minded individuals and or small communities; it did not seem to characterise large communities or epistemic cultures.

8.4.4 Trust, collaboration, "circles of trust", and trust technologies

8.4.4.1 Trust as a foundation for interpersonal openness

Trust was a major theme intersecting with interpersonal openness. For many interviewees, trust was a necessary foundation for interpersonal sharing, because of the vulnerability of sharing ideas. This seemed to be why, in a "very competitive and not always particularly friendly" field, Ernie cultivated trusting relationships in which to talk freely. Whilst Ernie implied that he would actively establish trust in order to share, many others portrayed their openness as a response to existing levels of trust that differed across relationships. Michelle[149] told me: "…people who you trust and know well, you might be open to […] if you're just chatting someone at a conference,

[149] PhD 1980s; interview: Australia, Jan 2017.

you might give vague outlines but not specific details. And completely open with collaborators…”. Many interviewees spoke of adjusting their interpersonal openness contextually in this way (see also Section 8.5). Ben[150] summarised this approach:

> *“…I think you make judgement calls on a case-by-case basis, whether that be to do with the state of the development of the work you're on, the actual project you're working on […] but usually actually the person. If you trust the person, and you see them as someone that you'd quite like to collaborate with, then you talk to them about what you're doing […] But there are people out there who I wouldn't tell stuff to, because I don't trust them.”*

Likewise, Miriam[151] emphasised a person-by-person, trust-based approach:

> *“…it's about the people, actually […] within the people who are good, there are going to be people who you want to work with or people you don't want to work with, and talk to things about […] with experience you figure that out […] people you trust, you can be open with…”*

Navigating trustworthiness within communities was portrayed as an important skill by many interviewees, and included circumspection around individuals known – through experience or by reputation – to be untrustworthy.

8.4.4.2 Interpersonal openness as a foundation for trust

Those who most strongly embraced interpersonal openness seemed to practise it without a foundation of trust. In these cases it was clear that openness could establish trust where it did not already exist, through free talk about exciting, valuable, vulnerable ideas. Generosity and vulnerability in sharing ideas seemed to be a social signal that prompted reciprocity, as Cedric noted: “The more open you are with other people, then hopefully the more open they will be with you”. Thus openness indicated trustworthiness, and performed trust before it is established. Vulnerability, in particular, could be expected to establish a social bond, especially if it is shared, and each party depends on the other to treat their ideas with care. Perhaps for this reason, interviewees often depicted interpersonal openness in close association with collaboration. Ben and Miriam, above, both showed how interpersonal openness can

[150] PhD 1990s; interview: UK, Jan 2018.
[151] PhD 2000s; interview: UK, Jan 2018.

pave the way to collaboration. Such openness provided a means to test for, and establish, the trust required for a close, ongoing working relationship. For interviewees like Jude who depended on trust for openness, collaboration was the main setting for interpersonal sharing: “a lot of real openness is only when you're involved in collaboration”. Interviewees mostly spoke of informal collaborations: professional relationships sustained over long periods, perhaps associated with collaborative grant or paper co-authorship. These collaborations were in some cases blended with friendship. Andrea commented that her supervisor’s close collaborators were his “professional friends”. This underlines the degree to which interpersonal openness intertwines the epistemic with the social.

8.4.4.3 Circles of trust

The theme of trust exposed the unevenness with which interpersonal openness operated in scientific communities: it flourished in trusting relationships and networks, but often not outside them. This was one way in which interpersonal openness manifested the complex, relational dynamics described in Chapter 3 (Section 3.4.2.2): it could be both open and closed at the same time. Miriam, for instance, spoke about selecting collaborative partners for big projects based on trust and their past openness with her: “...we'll talk to them, them and them, because they're really open”. Reflecting on this, she commented that this “blocks the people who don’t want to [be open]”, so “I suppose actually it’s the opposite of being open”. Similarly, Ben described it as prudent to only share unfunded ideas within a “circle of trust”. Andrea observed these dynamics amongst senior scientists:

> *“So obviously there is sharing and there is collaboration going, but with selected people. So you find networks of collaboration in topics but they’re normally based on people that know each other through PhDs or postdocs, people that have worked together in the past....”*

Michelle observed this dynamic somewhat from the outside, as someone less “embedded” in such networks:

> *“...I think overall there is a tendency to share less except with those trusted colleagues, and I think that leads to quite a cliquey-ness, that*

if you're in the in group, you get stuff shared with you, and if you're not, you know nothing."

Her account showed that it was possible to be systematically excluded from moral economies of interpersonal openness by being outside cliques, circles of trust, and long-established professional friendships. Michelle had changed fields a couple of times in order to stay in the same location as her scientist husband: as she became less connected, he was "very much embedded within a network of close collaborators and colleagues, where they all know each other and talk to each other". Mutual reinforcement of trust and interpersonal openness promotes this uneven flow of privileged information, ensuring that only some people are aware of the newest ideas and data in a field. Due to its social embeddedness, interpersonal openness could be expected to flow along contours established by existing social inequalities – not only gender, as was evident in Michelle's case, but race, class, socioeconomic status, (dis)ability, sexual orientation, etc., and their intersections – and to reinforce these inequalities. This was not inevitable: where interpersonal openness was practised independently of trust – or most people were assumed to be trustworthy – it had the potential to expand circles of trust or dissolve their boundaries.

8.4.4.4 Trust technologies and "trust-no-one-technologies"

The theme of trust had another key interaction with interpersonal and other categories of openness. Openness was viewed by some as a signal that a scientist's work could be trusted, and closure as a sign of untrustworthy work. This seemed to be true both of interpersonal openness, and other modes of openness that revealed process, such as data openness. Mark commented:

> *"…if somebody's not open, you start to wonder about what other things they might be hiding […] if they're not open, and you're not able to actually really examine, drill into that data in more detail, you start to worry about the integrity of their science."*

This suggests that interpersonal openness may operate as a "trust technology" that "offers a powerful reassurance that things really were done in the way it is claimed that they were" as argued by Grand et al. (2012:1–2) in relation to online forms of open science. These authors were drawing on concepts established by Shapin and Schaffer

(1985:60–69), who showed how seventeenth century experimentalist Robert Boyle established trust through “virtual witnessing”: literary techniques involving “the production in a *reader’s* mind of such an image of an experimental scene as obviates the necessity for either direct witness or replication” [original emphasis] (p. 60). These techniques included text and images rich in circumstantial detail, and inclusion of failed experiments; and a plain, functional style. Such techniques provided testament to the integrity not only of experiments, but also of Boyle himself a modest, objective, and therefore trustworthy observer. Interpersonal openness – not as a form of virtual witnessing, but as a willingness to “talk freely” and be generous with ideas – is a social technique that similarly establishes the sharer as honest and trustworthy.

My observation of a bidirectional connection between interpersonal openness with trust, and Grand et al.’s proposal that open science is a trust technology, both contrast with an implication prevalent in open science discourse: that openness can and should replace trust. This implication is evident in efforts to make the processes of science progressively more open, and especially in the removal of reliance on human actions or testimony. This is exemplified in the words of Jean-Claude Bradley, who coined the term open notebook science and pioneered its practice: “…In the past, trusting people might have been a necessary evil […] Today, it is a choice. Optimally, trust should have no place in science” (Drexel University 2014). In analysing open science discourse Gabrielsen (2020) makes the opposite observation to Grand et al. (2012): “In the Open Science-scenario, there is no real need to trust anyone as everything will be available for checking and validation […] Open Science is therefore rather a ‘trust-no-one-technology’” (p. 499).

As Gabrielsen (2020) concludes, drawing on Leonelli (2016), openness and data-intensiveness of science do not diminish a need for trust, but reconfigure it, placing emphasis on e.g. trust in data curation. Interviewees tended to uphold a role for trust, and several vehemently defended it. Mark said: “…in science, trust is absolutely essential…we rely absolutely, utterly, on trust”; Jacqueline: “I find it really scary if we're going to move to a situation where we don't trust our colleagues”. Some found trust a sufficient alternative to openness. Gavin said: “…if I trust the researchers to do a good job then I really don't mind if they want to keep mining their data, and they're not releasing it". And Michelle found openness an insufficient alternative to trust: "…I can't think of a lot of situations where you would think, well these people are so untrustworthy that I want to see the data, because if they were so untrustworthy you might mistrust the raw data as well." This underlined the limited capacity of online openness to prove trustworthiness, and its impotence where trust or distrust was established by other means, perhaps interpersonally.

The chapter thus far has characterised interpersonal openness according to its distinctive features that arose repeatedly: the sharing of valuable, vulnerable ideas; talking freely; generous reciprocity; and trust. The theme of trust begins to show why not all interviewees felt free to talk. The next sections show how interpersonal open-*ness* is not a single approach, but a spectrum of intertwined attitudes and practices.

8.5 Context and ambivalence: a spectrum of interpersonal openness

8.5.1 Open approaches: from unconditional to selective

Given the many nuances of social context and the bounded nature of interpersonal settings, a universally open or closed interpersonal approach is inconceivable. However, some interviewees portrayed something close to unconditional interpersonal openness: Jenny's "policy of generosity" did not seem to discriminate, even in environments where others were less generous. Several others reported a comparable approach: Elliot[152], for example, said: "I can't think of anything that I've tried to hold back on. Yeah, even to people who are working pretty much on the exact same stuff as us, I've pretty much said what I'm doing." Similarly, Luke[153] told me that he "would just say everything" when networking at a conference. Mat, meanwhile,

[152] PhD 2010s; interview: UK, Jan 2019.
[153] PhD 2010s; interview: UK, Nov 2018.

portrayed his lab's interpersonal approach as completely open, in keeping with his open source methodology, in which ongoing results were made public, online. For Mat, even unconditional interpersonal openness was limiting: he made it a rule to avoid email, due to its exclusive audience: discussions were moved into an online, public forum.

Other interviewees showed *almost* unconditional openness. Adam, for instance, mentioned "one or two occasions" on which he recommended to a student that they hold back in interaction with a scientist who does not "have a good track record". Withholding from particular individuals was common, and likely was implicit in accounts that appeared unconditional. This suggested that even the most open approaches were based in trust that could be selectively withdrawn. Lena[154] exemplified an approach at the open end of a trust-based spectrum:

> *"So my approach is [...] I just talk about everything [...] and if I'm ever scooped, well that's the nature of science, and you prepare your lab for that, and then you lick your wounds, and you get up, and you go."*

Lena assumed trustworthiness, and withdrew her trust and openness if she had adverse experiences with a particular individual: "if there is a problem, then I no longer trust [...] It might change the interaction with that particular person in the future, but that's it, right." She portrayed interpersonal openness as "the right thing to do" and something she would continue on principle regardless of specific negative experiences. In a spectrum of trust-based openness, approaches became gradually less open the more trust was withheld, rather than offered, as a default.

8.5.2 Mid-spectrum: contextual approaches that aspire to openness

Many interviewees' approaches were contingent on multiple overlapping factors and involved nuanced judgements; thus they could be placed somewhere in the middle of a spectrum of interpersonal openness. Typical considerations for these interviewees included not only the trustworthiness of the individuals involved, but also the

[154] PhD 2000s; interview: UK, Apr 2018.

competitiveness of their field and project; the developmental stage of their ideas; the security of their own or colleagues' careers; resourcing advantage compared to competitors; and the proximity of an interlocutor's research aims to their own. Interviewees with mid-spectrum approaches often embraced interpersonal openness attitudinally, or depicted themselves as characteristically open: it was a feature of their identity or sense of scientific self. Gavin[155], for instance, upheld a Mertonian-like ideal of scientific openness and said of his own interpersonal practices: "I think by and large, the default is to be open". In context, this meant "calibrating" against factors like the maturity of his research, and to whom he was speaking – where "some people are more competitive than others". As a result he spoke about his ideas before publication, at conferences – knowing this carried some risk – but only did so after the ideas were fairly secure, having been written up and submitted to a journal. He also described sharing ideas openly if he was unable to pursue them himself owing to a lack of personnel – but holding back on ideas that he planned to develop: "I might give some clues [...] I would not open it up completely". Gavin thus showed how openness could be an important and identity-shaping scientific principle with connections to a messier, nuanced, context-based practice.

Likewise, Oscar[156] told me that he talked about work in progress at conferences, which could be a risky but also beneficial form of exposure. He hesitated if the work was "at a very early stage" though, and shared it only with close collaborators. Greg[157] also characterised his behaviour as tending towards interpersonal openness: "I tend to not be very protective, I tend to talk about our stuff early before we publish it. But at the same time, we live in a practical world...". He told me he might occasionally hold back where a trainee of his was dependent on the work for their next job. Erin[158] too was comfortable being open, partly because she was "far enough along in my career and far enough along in my research database [...] that no one else can do what I'm going to do in [...] any near future time". When emailing potential collaborators with ideas that could be scooped, she was relatively unworried as she knew who was a "good

---

[155] PhD 1990s; interview: Australia, Jan 2017.
[156] PhD 2010s; interview: UK, Dec 2017.
[157] PhD 1990s; interview: US, Feb 2018.
[158] PhD 2000s; interview: UK, Aug 2018.

person” within her small field. She was cautious, however – choosing her words carefully and withholding detail – when the contact was about a new project, or outside her sub-discipline. Similarly, Ben[159] considered himself someone who shared more than others, but identified two “pinch points” before which he considered it normal to withhold: before securing funding he would only share freely within a “circle of trust”; and before getting data he would be careful not to say too much, as it would be possible for others to refocus their work to compete. He portrayed scientists, especially students, as apt to be “bad businessmen” – naïve and trusting – where instead they needed to be socially aware. Arthur[160] told me he “tried to be open”, but would hold back if someone was working on a topic close to his. All of these cases show that an open attitude was often more a guide than a rule: brought into practice through negotiations of social and epistemic context that sought to limit vulnerability whilst seeking some of the rewards of openness.

8.5.3 Mid-spectrum: entirely contextual approaches

Some interviewees did not claim a particular attitude or identity in relation to interpersonal openness, instead portraying their actions as entirely context-dependent. Yvonne[161] told me that “it depends on the project”. One of her projects at the time was confidential: “I have a postdoc who worked on the project, and so I need to not be open about it until we have the publication secured - because it's her future”. In other cases she was happy to be open – when there was less competition, when the discovery was not “big”, and when she had relatively unique resources. Michelle[162] similarly refrained from generalisations about her approach, instead characterising open ideals as out of reach, and in tension with practice:

> *“...I still have an idealistic tendency to think that people should be open and scientists should talk to each other, and share data, but realistically people are competitive and that doesn't happen all the time, and of course there's valid reasons...”*

---

[159] PhD 1990s; interview: UK, Jan 2018.
[160] PhD 2010s; interview: UK, Jul 2018.
[161] PhD 1970s; interview: UK, Nov 2017.
[162] PhD 1980s; interview: Australia, Jan 2017.

Her view was influenced by sensitivity to risk in her research environment: "…there's this constant fear all the time that I see in my colleagues […] that they're going to be scooped". She described being open with collaborators and trusted individuals, but otherwise protective:

> *"…if I'm doing a project then for most of it, it's completely open to some people, but not terribly broadly open […] if you're just chatting to someone at a conference, you might give vague outlines but not specific details".*

She also advised her students to be cautious, "warning them not to give everything away […] just in case", unless they were close to publication. This prioritisation of caution edged towards an interpersonal closure approach. Yvonne and Michelle may in practice have behaved similarly to the interviewees in the previous section, but their hesitation to characterise themselves or their practice as open is notable: they emphasised prevailing conditions, rather than internalised principles, as determinants of openness.

8.5.4 Interpersonal closure: prioritising caution and protectiveness

The accounts above showed that many interviewees were willing to talk about behaviours and attitudes of closure, despite framing openness as good for science. This was to some degree because of the contextual and graduated nature of the interpersonal category: it did not compel scientists to declare themselves open or closed. It was possible to be open and closed at the same time, depending on the exact nature of the information divulged, when, to whom, and with what intention (see Section 3.4.2 for analysis of the same idea in historical context). Some interviewees viewed protection of early-stage scientific information as not only a consideration, but a norm and a priority, due to competitive risks.

It was particularly illuminating to speak with Jacqueline[163] in this regard: she was the only interviewee who questioned the idea of openness at the outset. Her initial

[163] PhD 1980s; interview: Australia, Jan 2017.

response was: “Openness. Actually I don't really know what you mean by ‘openness’ in science, to be honest.” As the interview proceeded, she raised concepts that had more meaning to her, including honesty, trust of colleagues, and humility. However, her follow-up to the quotation above showed that the topic did have meaning for her. It brought to mind protection, rather than openness, of pre-publication ideas:

> *“I guess having worked in the traditional circumstances of competitive labs, where you are...I guess it's not really secretive, but protective of your ideas at least until you get them published [laughs]. To me that's normal. So I don't see that as not being open.”*

It thus became clear that Jacqueline’s initial reaction was in part a critique of the interview premise, which did not resonate with her. She recognised the same category of experience as those embracing interpersonal openness, but approached it in a different way, seeing herself as neither open nor secretive, but normal; and normal meant holding back from sharing unpublished ideas.

Jacqueline’s protective approach did not mean an abandonment of open practices, just as open attitudes did not preclude closures. She described how she might cautiously establish openness from a default position of closure:

> *“I mean there are certain things you're willing to share, and also when you meet certain colleagues or you meet new colleagues, and maybe you strike up a...you realise that you've got common ground, and you may start to share a little bit just to let someone know that you have data that's relevant to what they're thinking about and you may develop a collaboration. Or you might not.”*

Her negotiation of trust and potential collaboration through incremental sharing was comparable with Ernie’s cultivation of trusting relationships. The difference was Jacqueline’s emphasis: she gave weight to protectiveness, and did not associate pre-publication openness with joy, freedom or generosity. Instead, she placed weight on communicating fully and honestly at the end of the process. She saw it as a duty to be protective until the science was ready to be seen and used by others: “if you present your data, and your ideas, and put them out there and somebody runs off with it, that's...you put it out there! You have responsibility for that”. In this view interpersonal openness was a risk to be taken individually: shared information was fair game, and scooping a natural product of competition. This contrasted with the expectation of many who embraced interpersonal openness: that idea-sharing promotes a reciprocal, trusting social contract. Jacqueline's expectation, which conditioned her cautious response, was likely influenced by the norms of her field, which she described as "highly competitive" – in part due to the relevance of her work to climate change. The high stakes of this research topic also contextualised her emphasis on cautious, but full and honest, communication. When the competitive element of her work was put aside – for example when teaching undergraduates, or talking to the public – a joy in interpersonal communication emerged: "many of us biologists like myself love telling stories…we love talking about it, it's hard to stop us!"

Jude[164] had a similar approach that was contextualised by early career experiences:

> *"…my experience is […] okay, that the best scientists don't talk about their research until it's usually at least submitted, very few of the best scientists I know…well, especially if you're in a competitive field […] the one I'm more familiar with probably is* Drosophila *genetics*[165]*. And, you know, nobody was letting on what they were working on until it was at least submitted."*

This observation was from the 1980s, in the US: "that really opened my eyes to big time science, big time grants, and that there is this whole level of science that I didn't know existed". Interpersonal closure was thus portrayed not as a compromise, but as an integral part of performing top-level science, and being a top scientist, that was impressed on Jude as a young scientist. Like others, Jude made decisions based on trustworthiness and competitiveness. He saw his current field as "relatively non-cooperative […] fractured, and fairly fractious": possibly not a safe setting in which to share ideas. In line with this, he observed that "a lot of real openness is only when you're involved in collaboration", and even then, there was a possibility of "feeling ripped off at the end". As someone leading a smaller lab, Jude described how he offset competitive disadvantage by finding a niche in which he could be a "relatively big fish

[164] PhD 1970s; interview: Australia, Jan 2017.
[165] Jude's characterisation of *Drosophila* genetics as interpersonally closed is intriguing given its well-studied association with communality (Kohler 1999). There could be many explanations for this, including changes that occurred in *Drosophila* research between the 1920s and 1980s, differences between the communality described by Kohler and interpersonal openness, and the particular communities and experiences that Jude encountered.

in a small pond". He nonetheless portrayed his position as one that required defending, making a degree of interpersonal closure both necessary and wise.

Jacqueline's and Jude's accounts were distinctive in viewing interpersonal closure as a priority and a settled norm; perhaps part of a scientific self that upholds values in conflict with interpersonal openness. However, several other accounts, from more recently trained scientists, show that practices and attitudes of interpersonal closure are sometimes less settled, or held in tension with open aspirations. Rory's[166] response to the interview topic was influenced by experiences of secrecy and competition:

> *"…one of the things that interests me about science is that you're building on someone else's work, so you need to be open. But I know there's people who are hiding stuff that they don't want other people to build on because they want to get there first. So I don't know if that's all right."*

The group he worked in had, multiple times, had papers rejected because another group – who were both collaborators and competitors in the same field – had published just before them. As a result, he felt uncomfortable about consulting and sharing with this group, in case they used his results to gain advantage. This had begun to stoke a generalised disillusionment and distrust: "…more recently people aren't sharing. Or maybe that's just because I've gone higher up in science that I've realised that people don't really share their results. And aren't telling everybody the truth". During his PhD he had also been concerned about scooping, and had experienced unsupportive supervision. He felt little control over his own interpersonal openness – "I'm told what I can present and what I can't present" – and reported conflicting ideas of how he would behave if he had agency. When he thought about his present situation, he envisioned holding information back from collaborators – staying "a few steps ahead". When he took a more abstracted view, he "would want to be open". Strikingly, his view was influenced by whether he intended to "stay in this career" or not: pursuing a scientific career would mean a more closed approach. Rory's account shows how acutely competitive environments, encountered

[166] PhD 2010s; interview: UK, Jan 2019.

with a lack of support, begin to build attitudes of distrust and closure; or potentially, drive early career scientists out of the profession.

PhD researchers Nicole[167] and Andrea[168] also focused almost entirely on interpersonal openness in interview, and its curtailing by competitiveness in their fields. Nicole was researching dementia, "one of the big problems in the world that just hasn't been solved yet". Competition was inspired by the academic and societal rewards of success in this field, and efforts were focused on a small set of well-defined goals. The stakes were also high due to job insecurity: "a lot of people [...] are on one-year contracts". Nicole was accustomed to being discreet, as modelled and advised by colleagues, "until you're really, really ready to publish, so no one has time to catch up and push it out quicker". She was dissatisfied with this interpersonal closure, and interested in an open, collaborative alternative – especially as the research context was also "emotive" and "caring". However, she described an individually open approach as naïve in the absence of a systematic cultural shift.

Comparably, Andrea worked in an area of veterinary immunology and infection that was both academically and commercially competitive. She spoke about senior colleagues' strategic pursuit of funding ("science is driven by money"), the need for novelty, the danger of scooping ("the worst thing that can happen when you're doing science"), and secrecy around industrial collaborations research. She observed senior scientists limiting their openness to networks and behaving secretively at conferences whilst attempting to expose sensitive information:

> *"For example: I have discovered A, and I haven't told anyone. And I've been told that you have discovered A as well, but we haven't talked directly. So you're doing a talk about things related to A in a conference, and I'm going just to ask you, to push you to see how much you know. And you see that all the time. They challenge each other in questions to see how much they tell, to try to like scoop each other. Instead of just going up and openly talk[ing], like: oh, really interesting talk, we're doing something similar. Shall we sit down*

[167] Mid-PhD; interview: UK, Oct 2017.
[168] Mid-PhD; interview: UK, Oct 2018.

> *with a coffee and discuss about it? No no no, they prefer the bitchiness strategy [laughs], which is horrible. But that happens."*

Like Nicole, Andrea disapproved of these closures but saw them as inevitable, due to the competitive nature of science funding. She did not picture herself becoming a primary investigator: she associated this career path with an off-putting "politics of making money" to which her personal and scientific motivations were opposed. Thus like Rory, Andrea had an ambivalent relationship with her profession in which issues of competition and interpersonal closure were important factors.

One interviewee faced a more extreme case of interpersonal closure in their training. They left academic science after experiencing "complete shutdown of communication" in their PhD lab, to the extent that they never had the opportunity to experience competitive behaviour personally. Their supervisor seemed to have been affected by an earlier incident of scooping that led to "paranoia": "…we weren't allowed to go to conferences. We weren't allowed to talk to people about what we were going to do with our work. He [supervisor] managed all communications even with direct collaborators." This interviewee remarked in particular on an inability to develop networks and collaborations in this setting. Partially in response to what they had experienced, this interviewee subsequently developed a career as an advocate of open science. It was in that capacity that I interviewed them. These instances suggest that an inability to align oneself with a desired scientific self – one that upholds interpersonal openness as an epistemic virtue – can result in disengagement from this notion of scientific self, and associated professional identity, altogether.

8.5.5 Interpersonal closure in a commercial context

Andrea was conflicted about interpersonal closure for both academic and commercial reasons, which was somewhat unusual. There was a tendency for interviewees - who worked in academic settings, with varying degrees of industrial involvement – to view commercial projects and collaborations as uncomplicated exceptions to their usual practices, principles, and negotiations of interpersonal openness. The commercial became a separate domain, in which different rules legitimately applied, and tensions between collaborative and competitive instincts could be dissolved. For example, Mark[169] portrayed himself as “pretty open” and collaborative – and strongly, morally in favour of interpersonal idea-sharing. At the same time, he was pursuing a side-line to his main research in secret, having had “quite a powerful idea”, likely to be lucrative, that he hoped to patent. Mark emphasised how unusual this secrecy was for him, and the detachment of this work from moral obligation, unlike his main research which related to agriculture and was driven by a commitment to “society and planet”.

This dimension of interpersonal openness was an intriguing one. In the light of the wealth of scholarship on academic–industrial integrations (Croissant and Smith-Doerr 2007; e.g. Etzkowitz and Leydesdorff 2000; Gibbons et al. 1994), it would be beyond the scope of my thesis to analyse it fully. However, it was notable that interviewees tended to characterise and treat these domains as separate – and commercial domains as straightforwardly and legitimately secretive – despite the emphasis of STS scholarship on blending between domains.

8.5.6 Making sense of a spectrum of interpersonal openness

The accounts above depict a full range of intertwined attitudes and practices from near-unconditional interpersonal openness, upheld as a personal or group policy; to openness-as-default approaches that are variably sensitive to context; to entirely context-dependent approaches; to caution-first, protective approaches. This underlines Adam’s initial observation that “you see a huge variation”, from holding back completely until publication, to “talk[ing] about everything”. Interpersonal openness was unequivocally framed as an epistemic virtue by those at the open end of this spectrum, who had internalised it in their scientific identities as a means to better science and being a good scientist. These interviewees actively pursued interpersonal openness, initiated their trainees into it, and articulated it with particular care and emphasis: they saw it as a key distinguishing quality between scientists, and one on which they prided themselves. Further along the spectrum were many interviewees

[169] PhD 1980s; interview: UK, Feb 2017.

who framed interpersonal openness as an epistemic virtue, but more of a guide than a rule: they aspired to it and tended to shape their identities around it, but calibrated their practice to context. Nonetheless, they often took pride in working around their circumstances to create openness. Also in the middle of the spectrum were interviewees who saw their approaches as entirely context-dependent: they did not uphold interpersonal openness as a principle, perhaps due to a greater sense of tension with closures in practice. Finally, towards the closed end of the spectrum, were interviewees who saw caution and interpersonal closure as a norm. Some of these individuals had settled in this position, and asserted epistemic principles distinct from interpersonal openness. Others, who were earlier in their careers, upheld interpersonal openness as an epistemic virtue but felt unable to enact it: this state of tension was associated with ambivalence about pursuing a scientific career. Interpersonal secrecy associated with commercial science was treated as an exception that did not define the openness of one's practice or identity.

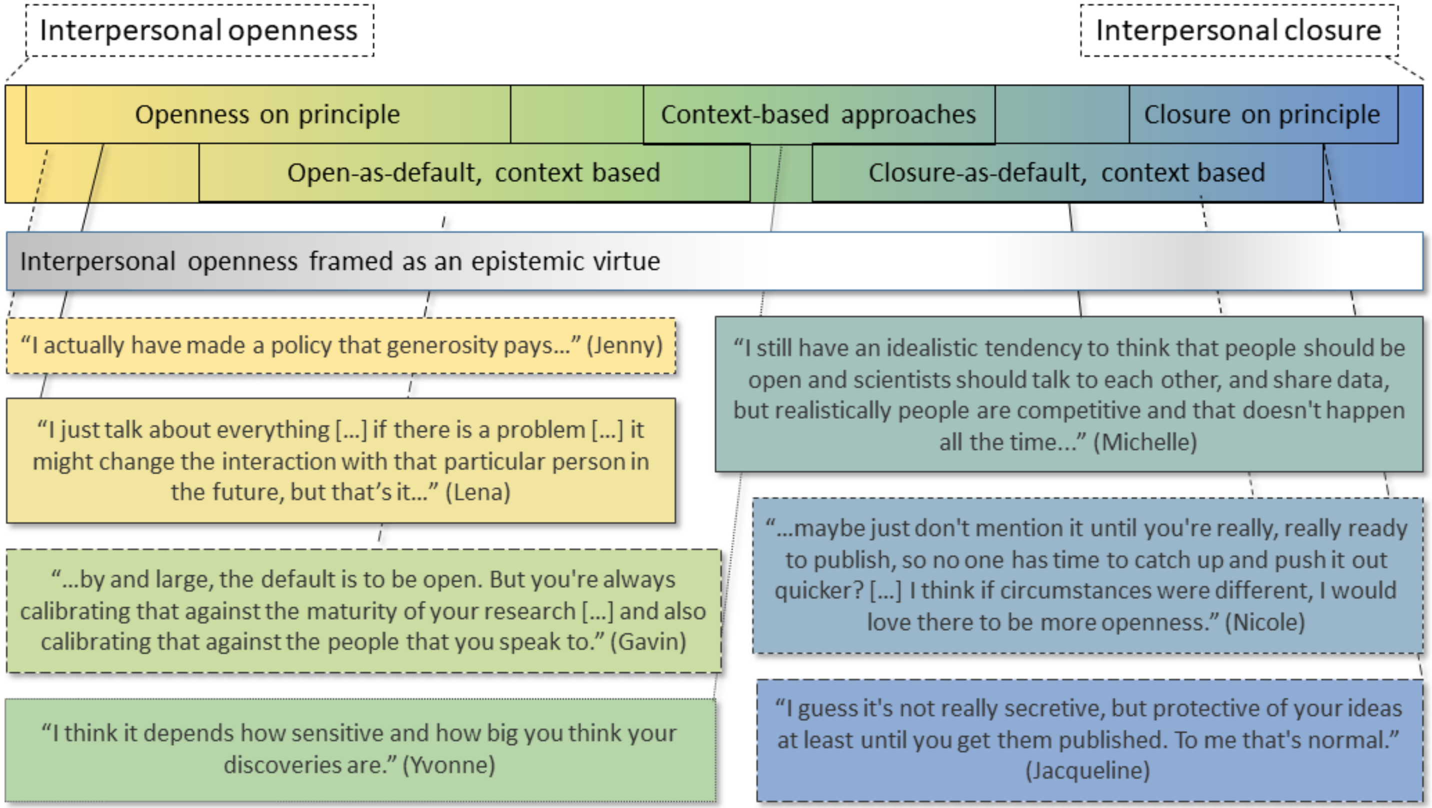


Figure 8 | A spectrum of intertwined attitudes and practices, from interpersonal openness to interpersonal closure.

Figure 8 visualises this spectrum of intertwined attitudes and practices of interpersonal openness. Different approaches are shown as overlapping to indicate that they are not absolute or clear-cut, but overall they fall on a spectrum. Interview quotes illustrate different positions on the spectrum. Categories are larger on the openness side to indicate that more interviewees saw themselves as embracing or aspiring towards openness than those who tended towards closure. A box with a grey gradient shows where interpersonal openness was constructed as an epistemic virtue. Because interpersonal openness is subjectively defined and experienced, the spectrum does not aim to present an objective view of interpersonal practice. It is possible that to an observing party, the behaviours of interviewees towards the middle of the spectrum may seem more similar than their accounts suggest.

8.6 Enabling interpersonal openness: security, resources, support

As is evident from interviewees' testimony thus far, their approaches to interpersonal openness were not simply a matter of personal preference. Attitudes and practices of interpersonal openness were enabled and constrained by a complex array of intersecting factors, both circumstantial and systemic. Here I refer the reader to the previous chapter on data openness, and in particular **Table 6**, where I compiled a list of factors that either build or diminish a sense of control and security, and respectively enable or constrain openness. Most of the medium- and long-term factors in this list also apply to interpersonal openness, because both forms of openness – especially when they reveal unpublished information – create vulnerability. Interpersonal openness exposes science at an earlier stage, but in a more limited and controlled setting. However, both are primarily focused on the risk of scooping, and are generalised enough that insights can be shared across the data and interpersonal categories: Table 6 was informed by both contexts. In this chapter I will limit my exploration of enabling and constraining factors to those most saliently and explicitly connected with interpersonal openness by interviewees. Factors like these interact with one another, influencing but not determining interviewees' approaches.

8.6.1 Career security

Several interviewees made comments linking their capacity to be open with their senior position and established career. Steve articulated this as follows:

> *"...well I'm [age in sixties], and my career has gone pretty well so far, and I'm not so much, not at this stage in my career, likely to miss out on promotion or anything like that by being second to publish anything. So it doesn't stop me from talking openly about things..."*

Adam was similarly conscious that his position afforded certain freedoms:

> *"...I'm aware that I'm in a luxurious position now [...] you can say well Adam, you can afford to be totally open [...] if you get scooped lots of times now and you never publish much again that doesn't really matter to you, you've had most of your career and you've got jobs now where it doesn't matter so much."*

Career security meant that Steve and Adam could enjoy benefits of interpersonal openness – such as feedback, intellectual stimulation, and networking – whilst being able to weather any negative consequences. In contrast Rory, an early-career researcher, felt that interpersonal closure would be necessary if he continued to pursue a scientific career (see Section 8.5.4).

8.6.2 Resourcing

The ability to be open was also connected with medium-term factors like the resourcing of a research group or project. This included team size, budget, and access to equipment and facilities: factors that allow work to be done faster or in unique or sophisticated ways. It also included access to privileged social and epistemic resources: collaborative links, contextual knowledge, materials, research settings, or data. Jude described both of these resourcing issues:

> *"...if you have the only material in the world, well you can be much freer about sharing ideas than if...a guy with a big lab over there could do what you're talking, what you're thinking about in two days when it's going to take you two months."*

Jude described unique access to material as "holding all the cards", a contextual advantage that could allow smaller and less wealthy research groups to thrive and share their ideas openly, where they might otherwise struggle. Nicole similarly had "a particular element that we got from one of our collaborators that's particularly valuable": this was a source of reassurance when her conference poster was photographed by a competitor.

8.6.3 Competitiveness, collegiality, and social support

Interviewees' characterisations of their research communities differed widely, from competitive, even nasty, to friendly and supportive. Competitiveness tended to be linked with cultures of interpersonal closure, whereas friendly communities seemed to support openness. Nicole's account was an example of a highly competitive, high-stakes field being linked with interpersonal closure. Jacqueline, too, connected the high stakes of her field – related to climate change – with its competitiveness, whilst upholding interpersonal closure as a norm. Ernie, who showed a strong commitment to interpersonal openness, nonetheless limited this to certain trusting circumstances, in his very competitive field. In contrast, Adam's interpersonal openness was set within a "lively, supportive" field, a character he helped to promote: "we, the senior people know each other and like each other, and try to be supportive of each other and the people coming through the labs as well". David[170] also strongly upheld interpersonal openness, and did so partially in a collegial context of his own creation. As an institute director, he had sought to promote interpersonal openness at a community level through architecture:

> *"This building is actually designed deliberately to encourage openness. So you can't get to the laboratory floors without going through the cafeteria. That's deliberate. There's only one block of loos in the building, and they're in the middle of the building. Which means you've got to get up, and walk the length of the building, and walk past people. Open-plan, shared laboratories...is all intended avoid any kind of territorialism."*

[170] PhD 1970s; interview: UK, Dec 2017.

These accounts showed that interpersonal openness or closure sometimes characterised a community or field. This made it easier or more difficult to trust one's colleagues, and practise openness as an individual or research group. Adam and David had reached points of security and influence in their own careers from which they could shape openness culture at a community level. Of course, each interviewee characterised their community through a partial lens. It would be possible to see and experience one's colleagues as untrustworthy in a field perceived by others as supportive, and vice versa, depending on factors like social embeddedness. Types of social support described in Chapter 7 are also relevant (see Section 7.10).

8.6.4 Early experiences, role-modelling, and generational dynamics

Interviewees commonly mentioned interpersonal openness of their supervisors or mentors, and those in supervisory positions readily spoke of advising their students and post-docs about interpersonal openness. Adam linked his open identity with the influence of a mentor who "talked about everything to everyone" and espoused this as a philosophy, reckoning that the risk of scooping was worth it. "That has stayed with me", Adam told me – and he passed it on to his students. This connected with a depiction, throughout the interview, of his enjoyment and care in training students. Interviewees pursuing interpersonal openness as an epistemic virtue were particularly inclined to speak of cultivating this approach in their students: encouraging them to share liberally, and not to feel constrained. Adam actively introduced his students to "big names" in the field and left them together to converse "as equals" without "any sort of barrier".

Others also highlighted the people and experiences that had encouraged them, early in their careers – and often decades ago – to be open. This suggested that role-modelling, advice, and supervisory care practices were influential in passing interpersonal open approaches down through generations of researchers. There was also an indication that these approaches might persist via a virtuous cycle, as those who practise interpersonal openness from early in their careers benefit from reciprocal rewards. Jenny demonstrated this possibility: although she did not have a mentor, a non-competitive early environment and a need for visibility and feedback encouraged her to adopt an open approach. By the time she encountered riskier environments, she had already established her practice and seen its benefits. Thus, it seemed that freedoms or constraints of training environments, including but not limited to supervisory influence, were formative, and could develop a degree of resilience to vulnerability in subsequent environments.

It is harder to say whether more cautious approaches travelled in the same way, but it is likely, as interpersonal openness largely operated at a research group level. When asked, senior scientists with more contextual or protective approaches readily reported passing this circumspection on to their students as a matter of common sense (e.g. Michelle warned "not to give everything away [...] just in case"). This influence also seemed to have the capacity to last: Jude's cautious approach was contextualised by his early observations that "the best scientists don't talk about their research until it's usually at least submitted". Importantly, both open and protective role-modelling and mentorship were practised as acts of care for students.

Generational influences were not always strong, determining, or unidirectional. Olivia, a PhD student, displayed a commitment to interpersonal openness that was actively discouraged by her primary supervisor:

> *"I was at a conference and I was speaking to one of, well, probably our biggest competitor [...] we just ended up having quite a lot of interesting scientific conversations together, and by the end of the conference he offered to collaborate with me [...] So I went quite excitedly back to my supervisors and said: oh well this has happened and it all sounds great. And then, he sort of had quite a negative reaction to it."*

Luke similarly reported defying advice from his Master's supervisor not to talk to a rival group. These cases showed that some scientists, even early in their careers, have acquired sufficient agency, confidence and security to pursue their own approach, and may even be motivated by their unease with their seniors' practices.

Complexities of generational influence also arise from the fact that early career researchers, if they stay in academia beyond their PhD, typically experience more than one training environment and approach to openness. Julia[171] described her open approach as influentially shaped by her PhD supervisor, "one of the most open people in terms of his science that I've ever come across". She only realised later on that "many people don't work like that", including the PI of a lab she subsequently worked in: "that was a time when I had to contain myself". At the time of interview, Julia had recently become a PI herself, and planned an open approach both interpersonally and online. Her cautious former PI had recently supported eight trainees in starting their own labs, and these new labs "are now networking and sharing information with each other in a way that she [cautious PI] probably wouldn't have done". Julia implied that role-modelling may begin to work in reverse, as her former PI seemed happy with the new sharing behaviour of her scientific offspring. Olivia also reported reversing the direction of influence: "I think I generally try and push for him [supervisor] to be more open all the time. And he's responded well." These cases are probably exceptional, but they show that interpersonal openness is not embraced or perpetuated by senior or junior scientists exclusively.

8.7 Conclusions

In this chapter I explored the third and final category of openness that was highly salient in scientist interviewees' responses. Interpersonal openness was distinct in that it did not have formal links with openness as constructed by twenty-first century open science movements. Accounts of interpersonal openness often had a captivating quality, and several interviewees underlined its value to them in contrast with online openness, in particular open access. It emerged as a category of openness that gave meaning to scientific lives, and to which many had a moralised attachment, or even an explicit "policy" commitment. Interpersonal openness could thus readily be conceptualised as an epistemic virtue with an important place in many scientists' sense of self or professional identity. The emergence of this category was partly

[171] PhD 2000s, interview: UK, Jun 2018.

contingent on my specific choice of words: “openness in science”, not “open science”. “Openness in science” allowed interviewees to articulate something outside normative discourses: outside open–closed binaries; outside outputs; and outside policy prescriptions.

It might be argued that interpersonal openness is not comparable with other categories under the expanding open science umbrella. Open science advocates and policymakers tended not to bring it up as a form of openness[172]. It was not revolutionary, but mundane and low-tech: simply a mode of human interaction. Moreover, it was constitutive of a science that open movements sought to transform: conditionally, relationally open; messy, ephemeral, replete with human influence; unaccountable to outsiders; and coextensive with networks of privilege. However, because of its separation from contemporary “open science” – and its high level of recognition by scientists – interpersonal openness offers novel insights on the meaning and practice of openness in science. Here, I bring it into conversation with the themes of my research sub-questions.

In historical context (sub-question a), interpersonal openness offers a perspective different from both traditional, Mertonian narratives of essential openness (see Chapter 3), and contemporary “open science” narratives. In contrast with the former, interpersonal openness is defined by unpublished, not published work. A history of interpersonal openness – as a history of unpublished communication – may even mirror a history of scientific publishing in interesting ways: perhaps interpersonal domains have gained new value as published communications have, over time, become more formal, commodified, and competitive (see Section 4.6).

[172] Open science advocates/policymakers sometimes identified the similar kinds of distinctions between “open science” and “openness in science” as scientist interviewees, discussed in Section 8.3, e.g. the latter as more process-, behaviour-, or person-oriented. Peter Murray-Rust described open-*ness* as a “philosophy […] a way you feel”, and Mark Patterson associated it with scientists who embody open practices. However, advocate/policymaker interviewees tended not to describe or recognise interpersonal openness as part of open science. Jon talked about variable disclosure levels at conferences, but commented, “I don't think it is open science, but I think open science might factor into it somehow”.

The emergence of interpersonal openness in interview again underlined the expansive potential of “open” terminology (the theme of sub-question b). It also demonstrated how contemporary open science movements, despite their breadth are mainly expanding in particular directions defined by technology and visions of universality. Interpersonal openness was most salient category outside this direction of movement, but interviewees also raised the following as forms of openness: support of junior colleagues; openness with research participants; openness about animal work; receptivity to new ideas and questions; inclusion of diverse individuals and perspectives; and freedom to be open about identity in science (in this case, LGBT identity). These dimensions of openness, which are more about social challenges and contextual opening, illuminate the partial focus of movements that appear all-encompassing.

The dynamics of interpersonal openness offer particular insight on questions of epistemic virtue (c) and scientists' apparent disengagement from open science (d). I explore these connections more fully in Chapter 9, where I also bring themes from Chapters 6 (open access) and 7 (data openness) into conversation with interpersonal openness. While interpersonal openness was constructed by many interviewees as an epistemic virtue with a relationship to the self, it differs in interesting ways to the main epistemic virtue characterised by Daston and Galison (2007), mechanical objectivity (see Section 2.2.1). The relationship of interpersonal openness with the self seems to be one of trust, enrichment, and excitement rather than denial or restraint as in the case of mechanical objectivity. I also note that understandings from this chapter can expand ways of thinking about scientists' relationships with openness beyond current policy and advocacy discourses. In interpersonal contexts, openness is not new: scientists have career-long experience of negotiating it around situated vulnerabilities and opportunities. Calls to be publicly open, online – particularly when they apply to unpublished science – therefore meet and contrast with established landscapes of openness that are nuanced and relational. An understanding of where, why, and how scientists and their communities come to embrace interpersonal openness – or interpersonal closure – is valuable its own right, to promote security and wellbeing in scientific lives and careers. It is also valuable as a rich, layered, empathetic way to understand (dis)engagement with online openness practices, many of which prompt similar vulnerabilities, but in a setting of altered or diminished trust, talk, and reciprocity: features that make interpersonal openness feel safe and worthwhile.

I have now identified, characterised and analysed the three most salient categories of openness raised in interviews with scientists. The final chapter is a discussion that brings these categories, and the different ways of being open that they represent, into conversation with one another and with a broader context.

Chapter 9 | "Actively engaged and willing to be open" **Discussion and conclusions**

Lena spoke to me with candour and directness. The group leader of an interdisciplinary lab who had completed her PhD in the 2000s, she was in search of nothing less than the "equations that describe life" – or a reason why they do not exist. She responded to my questions with precision and feeling, extending her attention as a critical researcher to the topic at hand. She mentioned "scientific integrity rules" which guided her towards interpersonal openness. When I asked about the origin of these rules, she gave a striking account of identity, ethics, and epistemology – bringing epistemic virtue to life:

"...I do science because it's how I am, right. I wake up in the morning, and the first thing I want to do is go to my lab, right, and that's since I started [...] That is the motivation to do good science. Of course we want to be acknowledged [...] of course we want to publish the best papers that we can – but with that good science, right. If that's gone, then there is no point in everything else [...] It's just a very important thing to preserve and I'm just not sure where it comes from [...]

So it's hard to define. Let me try. So, if you are motivated by the wish to find the answers to certain questions, right - the real answers – not what seems to be a real answer if I present it to a reviewer [...] But if you really want - you, yourself - want to know what is the real answer. *Then I think this science integrity comes naturally. Because it's kind of like there, to enable those real answers to be found."*

Lena, biophysicist

9.1 Introduction: the story so far

I began this thesis by observing six curious aspects of open science that made it a timely and rich topic for study: its rapid rise to salience; future promise; historical resonance; elusive expansiveness; moral overtones; and apparent distance from the lives of (many) scientists. In this chapter, I look back upon these overarching themes – guided by my research questions – and consider what I have learned. First, I will briefly revisit each empirical chapter to summarise its discrete findings. Research sub-questions (a) and (b) – set out in Section 1.7 – are primarily addressed in the earlier chapters. I will then bring these findings into conversation with one another, focusing on the interview study, to consider the differences and similarities between different categories of openness from scientists' perspectives. This allows me to address research sub-question (d) directly and holistically: *why does it appear that many academic scientists are disengaged from "open" discourses and practices?*

I then introduce epistemic virtue, to consider whether it adds depth this explanation, and to address sub-question (c): *to what extent is "open" (or "openness") being constructed as an epistemic virtue?* I then suggest another layer of interpretation: that the breadth of openness might contain several distinct epistemic virtues – which I identify – and that these play different roles in the present open science "revolution". I then contextualise open epistemic virtue(s), re-asserting their contingent place and enabling and constraining factors that make them unequally available to scientists. I consider, in this context, one of my early curiosities: in what sense we might be experiencing a shift in the agreed meaning of "good science", defined according to openness.

9.1.1 Chapter 3: Historical framings of openness in science

My overarching interest, through this research, has been in how meanings of "open" in science are being constructed. This first literature chapter resulted from a curiosity about historical framings: both contemporary open science texts, and older accounts of the workings of science, implied that scientific openness had a long history. Moreover, the narratives in these texts built up a picture of science as essentially, inevitably, enduringly open. This was an intriguing juxtaposition with contemporary narratives, which are of crisis and opportunity: today's science falls short in its openness, but can be restored and elevated through Internet technology. Because of this, I asked (a): *How do contemporary "open science" discourses relate to older, historical discourses about the essential value of openness in science?*

The answer was more complex than I expected. I began by reading histories of scientific openness, and learned about its contingent origins. These are commonly traced to seventeenth-century Western Europe and co-located with the first scholarly journals and the birth of modern science (David 2008; e.g. Eamon 1985). However, an examination of openness and secrecy as social practices (in historical works, e.g. Balmer 2012; Iliffe 1992; Vermeir and Margócsy 2012) showed that the two tend to be closely intermingled in cultures, individuals, and even actions. I also observed that flexible semantic qualities of both "open" and "science" enable a circular kind of boundary work, through which it is possible to find openness in almost anything considered scientific. I therefore sought to describe the openness of seventeenth-century Western science in specific and situated ways: public, communal, and cumulative; defined by journal publishing; and within elite, gendered, not-yet-professionalised community structures of the time. It was more difficult to say – owing to its subjective, relational qualities – whether a distinctive cultural idealisation of openness arose at that time. I was able to conclude, however, that historical and contemporary narratives of scientific openness are closely entangled. Such historical narratives are *part of* a contemporary phenomenon that re-imagines past science as "open" in contemporary terms, and adds depth and authenticity to today's open science narratives. Even historians have subtly participated in boundary work – seemingly from the 1980s – that attaches "openness" to Western science (e.g. Eamon 1985; McMullin 1985). In the 2000s, this became "open science" (David 2008).

9.1.2 Chapter 4: Twentieth century preludes to "open science"

In this second literature chapter, I viewed the history of "open science" differently: instead of looking for a long-ago origin for scientific openness, I thought about salience and breadth of contemporary "open science", and its apparently sudden appearance in the twenty-first century. The Internet is often framed as the driving force behind open science, and there is a tendency towards technological determinism. I sought a more holistic, nuanced backstory that would help my understanding of how open science came about. Many parts of this story have been told, but my aim was capture breadth, and explain salience. In doing so, I sought to address (b): *How has "open" in science recently gained such salience and discursive power, despite its multiple meanings and lack of specificity?*

In seeking the roots of open science, I brought together diverse twentieth century histories and trends, including:

- an increase in scale, and archetypal "big science";
- the birth of preprints in physics;
- the influence of World Wars;
- a growth in commercial publishing;
- an intensification of academic–industry relations in fields including biology;
- the rise of the early Internet;
- the development of sharing cultures in model organism communities;
- and the emergence of free and open source software movements.

This showed that in the century before the appearance of "open science", science itself shifted in profound ways: its openness was not simply present or corrupted; instead the relationship between science and openness changed qualitatively in multiple dimensions. Some of these historical developments suggested an encroaching secrecy, and others depict new or reinforced cultures of scientific openness. Distinct elements of today's "open science" – e.g. open access – can be traced in one or several of these histories. They seem to have converged under the heading "open" in the early 2000s. Addressing (b), I proposed that the salience and discursive power arises *because* of its ambiguity. "Open" has become a rallying point for meaning, standing for diverse values that bringing together many constituencies, with myriad histories – magnifying the salience of "open". In this way, "open" comes to represent great deal of all that is considered good – and wanting – in contemporary science.

9.1.3 Chapter 5: Advocacy and policy framings of open science

This final empirical literature chapter moved from a wide-angle perspective history and meaning of scientific openness towards a narrower focus on the contemporary period. My goal was to concretely situate the interview study by providing detail on two "open" movements – open access and open data – as well as the holistic movement of open science. I also analysed the construction of openness in these advocacy and policy contexts through nine document case studies from the period 2000-2019. This was an engagement with my overarching research question – *How is the meaning of "open" (or "openness") being constructed in the context of science?* – that addressed the themes of all four sub-questions.

Through this document analysis, I confirmed a tendency for advocacy and policy narratives to draw on a long history of openness in science. However, this framing was not universal, and seemed to gain momentum only after "open science" was envisioned holistically. It also seems to lose relevance as "open research" and "open scholarship" become preferred framings in Anglophone contexts. I also illustrated an accumulation of meanings associated with "open", from a broad base. Values associated with advancing knowledge and public benefit were present early on; accountability and transparency/reproducibility framings appear later; and critical reckonings with inequality are recent. Both grassroots and policy documents tended towards technological, business-friendly framings, obscuring a complex relationship between openness and capitalist values. The documents depicted a gradual shift towards the construction of openness as an epistemic virtue, and towards the portrayal of researchers and their cultures as by-products of incentive systems.

9.1.4 Chapter 6: Open access to research articles

Open access was the first of three main categories of openness constructed by scientists in interview, consistent with its salience in advocacy and policy. Most interviewees supported open access, at least in principle, but it was characterised by prominent problem framings. The clearest of these was financial: open access seen as an expense. This made it seem impracticable to some, easy to others, and associated it with conditionality, scarcity, and injustice. This financial framing was linked with the construction of open access as a "gold" journal-mediated, and fee-paying, practice.

The cost-free repository route ("green") was mentioned less often and framed as a less valuable or actualised form of open access, despite its prominent role in local policy mandates. There was a hint that more recently trained scientists see the benefits of both "gold" and "green" in preprints. A minority of interviewees associated open access with a loss of quality, rigour, and community in scientific publishing. Many more associated it with lower "impact", which they expected would be judged negatively by *others*, even though they tended to profess disagreement with that judgment. Some more recently trained interviewees framed open access separately from quality and impact, or saw open access as a reputational asset.

Experiences of open access were shaped by a policy compliance framing through which open access was distanced and bureaucratic. This was particularly the case in the UK context, where the open access was often administrated by professional services staff. This shaped framings of open access as easy, on one hand – but lacking in personal meaning – or frustrating on the other, where bureaucratic friction arose. Open advocates at UK universities gave congruent responses, and seemed caught in a bind between a desire to win "hearts and minds", and need to compel compliance.

9.1.5 Chapter 7: Opening up research data

Data openness was the second, comparably salient category raised by interviewees, and again was associated with advocacy and policy discourses. Reactions to this category were highly varied, so I grouped them by attitudinal and emotional qualities. For a sizeable group, data openness was familiar and accepted, but not highly salient or emotive. I concluded that for this group, who tended to work in genetic or crystallographic fields, data openness was probably shaped by digital data sharing norms established 1980s and 1990s for specific data types.

A second, smaller group found the topic salient, and reacted with unease and defensiveness. Here, data openness carried a blended epistemic and personal threat:

worries about competition, and losing the context attached to data, were prominent. However, these interviewees were happy to share data post-publication. The assumption that pre-publication data sharing would be required seemed to emerge from the expansive qualities of open science, and distrust in those implementing it.

A third, large group also found data openness salient, and were passionate about it. They were driven by a variety of principles that reflected the breadth of contemporary open science movements: including links with software culture, reproducibility, public duty, and epistemic self-improvement. Nearly all these interviewees had trained since the 2000s, suggesting a profound influence of open science movements. Principled attachments to data openness could still be marked by vulnerability. I proposed that a feeling of control – arising from a combination factors – is a pivotal in enabling scientists to embrace data openness. I suggested that the vulnerability of data openness may be more tolerable to scientists who internalise it as an epistemic virtue, because openness can become an act of care.

9.1.6 Chapter 8: Interpersonal openness

The final category I described as "interpersonal openness". Its salience alone was remarkable, since it had no foundation in established open science discourses: but interviewees raised it consistently and with conviction. They spoke of face-to-face interactions with other scientists in which unpublished information – especially ideas – were exchanged. They spoke of talking freely, sharing generously, with indirect reciprocal benefits; and a dependence on, or a creation of, trust. Where interpersonal openness was dependent on trust, it was associated with cliques through which information flowed selectively. The emergence interpersonal openness was likely contingent on my phrase "openness in science" (not "open science"). This contingency allowed interviewees to express something important to them outside normative discourses. Attitudes and practices of interpersonal openness fell on a spectrum from near-unconditional openness, through to principled protectiveness. Those who embraced it tended to construct it as an epistemic virtue, core to their professional identity and an approach to instil in trainees. Mid-spectrum, interpersonal openness was constructed as context-dependent: an aspiration that seemed achievable or naïve to varying extents. For some, interpersonal closure was the norm: either as a settled position, or – often for junior researchers – an uncomfortable compromise.

Interpersonal openness was both epistemically and personally vulnerable, so was enabled by a sense of control, like data openness. Role modelling and social support at training stages seemed to be an important basis for long-term interpersonal openness, further underlining its connection to development of the (scientific) self. As a category of openness defined by scientists, it offered particular insight into scientists' apparent disengagement from "open" discourses and practices. I will elaborate these insights in the following discussion.

9.2 Synthesis: comparing interview-based openness categories

The three most salient categories of openness that emerged in interview were contrasting: conceptually, in practice, and in their relationship with scientists. This led to an opportunity for comparison and direct reflection on sub-question (d): *why does it appear that many academic scientists are disengaged from "open" discourses and practices?* Here, I present a comparison broken down into thematic sections, and summarised in **Table 7** (Section 9.2.5).

9.2.1 Stage and interactivity of openness

One of the clearest differences between these categories in scientists' accounts was the stage or state of knowledge-making that is "opened". Open access applies to the publication stage: a relatively settled knowledge claim in a bounded, permanent physical or virtual form. Data openness exposes science-in-process through bounded information objects that contain multiple interpretative possibilities, and may structurally enable re-use. Interpersonal openness reveals ideas upstream, non-linear components of knowledge-making in ephemeral, fluid, and appropriable form. In this way, the three main categories identified by interviewees happen to fall along a spectrum.

Another lens through which to view this spectrum is interactivity. Open access publications are primarily one-way communications. Online data openness also takes the form of one-way dissemination, but may prompt more interaction than open access, e.g. if users ask for contextual information or contact the producer as a matter of etiquette. Moreover, the re-use of data creates a form of epistemic intimacy between producer and user, as the latter becomes familiar with minute details. At a pre-publication stage this may reshape findings. Interpersonal openness is inherently interactive: ideas are produced and altered *through* interaction. These categories of openness thus differ profoundly in their receptivity to outside perspectives, and thus their capacity to shape knowledge. An inverse relationship emerges here between *access* and *receptivity* as connotations of openness: public online access (open access and data) seems to entail a lower potential for interactive, creative, upstream re-shaping of knowledge; and interpersonal openness creates the latter through the exclusivity of intimacy and trust. This raises the question of whether online, public openness *necessarily* limits other types of openness.

9.2.2 Scale, infrastructure and standardisation of openness

Interviewees' three categories of openness are also standardised to different degrees, in association with the scales at which they are organised and governed. Open access, especially in the UK, encompasses a limited range of practices that are definable, administrable, and measurable by the top-down policies of funders and universities (limited but not simple: as advocate/policymaker interviewees noted, there are several main funder policies in the UK, leading to administrative complexity). This large-scale standardisation and governance seems to occur partly because "opening" a journal article means the same thing across epistemic contexts: its openness is not entangled with the specific form of knowledge it conveys. Open publishing and repository infrastructures can therefore operate in the same way in different academic fields, and researchers across fields encounter open access in similar ways – with the notable exception of humanities and social sciences fields in which monographs hold particular value. In short, there are a limited number of ways to be “open” in the current context of open access.

Data openness features a complex layering of scales, infrastructures and standards. There are now funder and government-level open data requirements and generic infrastructures for data sharing (e.g. Zenodo n.d., “a catch-all repository”). Before this, some disciplinary communities had created infrastructures for specific data types, associated with local norms and journal-governed mandates (Strasser 2019). Any one scientist may now encounter multiple standards and norms from different sources and scales; or none at all. Data openness is less amenable to top-down governance, because of variety in what constitutes “data” and sufficient openness[173]. In other words, there are a large number of ways to be “open” with data. This relates to the entanglement of data – compared to articles – with science-in-the-making.

Standardisation does not apply in the case of interpersonal openness, which falls outside a policy framing, and would be difficult – and deeply problematic – to measure and control. Through this social, relational lens there are myriad qualitative shades of openness, though none are public. Some uniformity seems to occur at the research group level. These comparisons again depict an inverse relationship, which may hold more widely: the more amenable a form of openness is to large-scale standardisation and governance, the more constrained it is, and the less contextual depth it reveals about processes and relations of knowledge-making.

[173] Graeme, an open science policymaker/advocate at a major UK research funder, commented: “With data, we don't say: you have to put your data in [specific repositories], we recognise that [...] there's so many different types of it, and there are so many different norms about how people share...”. Interview: UK, Feb 2018.

9.2.3 Enforcement and agency of openness

Differing scales of standardisation are thus associated with different approaches to enforcement, and different experiences of agency for scientists. The relative standardisation, infrastructuring and measurability of open access render it amenable to enforcement, which is strongly pursued in the UK context. This shapes its construction through the binary lens of compliance, and reduces the need for any in-depth form of engagement by scientists to a minimum.

Some open data policies are similarly top-down, but with less extensive enforcement, partly because data openness is less amenable to standardisation and measurement (although Data Management Plans are a top-down mechanism for accountability, Neylon 2017). Not many interviewees framed data openness as a matter of compliance, perhaps because open data policies had not made a wide impact at the time. Alternatively, the more contextually sensitive range of openness practices encompassed by such policies may allow a greater sense of agency. A substantial group of interviewees had a principled attachment to data openness that seemed independent from policy requirements, but shaped by a wider advocacy and policy landscape. This may be because of the epistemic entanglement of data openness.

Interpersonal openness is not governed or governable by formal structures, and was characterised in interviewees' accounts by freedom and agency: a free quality of interaction; a freedom *to* be open; and agency to control precise levels of closure. For many it seemed to represent an escape from structure and pressure and a space for nurturing one's scientific identity.

9.2.4 Attitudes and emotions of openness

The dimensions above were connected with patterns in attitude and emotion. Positive responses to open access were approving, acquiescent, and occasionally enthusiastic.

Many positions were neutral or passive, perhaps connected with a compliance framing or the perception that open access is a "done deal" (Thomas) that no longer needs to be discussed (Henry and Kate). On the negative side, scepticism and frustration were characteristic orientations, linked with the financial framing of open access, and distrust towards publishing and institutional systems.

Data openness was associated with a wider and qualitatively distinct attitudinal range. On the positive side, it was embraced with passion by a significant group – often more recently trained scientists – who seemed to have internalised openness as an epistemic virtue across a range of context and practices. Another significant group were approving and acquiescent, having adjusted to discipline-specific data sharing norms. A third group of reactions were defensive and vulnerable, and linked with the contextual and competitive sensitivity of opening up data.

Interpersonal openness was associated with strong positive emotional reactions as well as ambivalence. Distinctive to this form of openness were expressions of joy and excitement linked with the experience of freely exchanging ideas in a social context. Principled and passionate expressions of commitment were also characteristic. Neutral or negative reactions seemed less demonstrative, perhaps because of the agency of interpersonal openness: risky or vulnerable elements of the practice could be avoided through contextual forms of closure.

9.2.5 Bringing it together: a multi-layered spectrum of openness experiences

These comparisons build up a picture of the distinctive ways that open access, data openness, and interpersonal openness are experienced by scientists. Many of these experiences are specific to the present context: biological science, at relatively privileged institutions, in the UK and Australia in 2017-2019. However, there is potential to make more general observations about the reasons for scientists' relationships with these forms of openness, and the factors that promote engagement or disengagement. This addresses research sub-question (d): *why does it appear that many academic scientists are disengaged from "open" discourses and practices?*

First, I propose that forms of openness that are more epistemically involved – that is, entangled with knowledge-making *processes* – have a greater capacity to engage scientists. Being open can feel like part of doing science in these cases. This happens, for example, when interpersonal openness leads to new ideas or collaborations, or critiques. It also happens when an embrace of data openness causes a scientist to organise and examine their data in new ways, or when data sharing leads to critical input that catches a mistake or shapes an interpretation. This may partially explain why it is difficult to engage scientists in open access.

Secondly, I propose that it is easier to engage scientists in forms of openness in which social enrichment – and feelings of joy and excitement – are part of epistemic enrichment; and in which trust mitigates vulnerability. This is a lesson drawn from interpersonal openness, which blends knowledge-making with social interaction, trust-building, and collaborative friendship. This presents a challenge to open science practices, which in part rely upon a dissolution of necessary social ties between creators and recipients of information in order to enable universal online access, independent of social networks. Data-sharing as a technical, decontextualized, human–machine interaction may feel both less rewarding and more vulnerable than an equivalent interpersonal exchange.

This ties in with a third proposal: that top-down, standardised, enforceable forms of openness present a problem for engagement, and for nurturing openness in research communities. This is for a variety of interlinked reasons. Forms of openness that are amenable – or have been made amenable – to top-down administration are, firstly, likely to connect less closely with local epistemic processes and contexts, and thus to lack personal meaning for scientists. As a result, large-scale governance of open science may both disproportionately measure, and incentivise, forms of openness from which scientists are disengaged. This underlines the importance of local, contextual approaches; and the potential harm of both metrics-based and other standardised approaches to incentivising openness in research communities.

Moreover, top-down governance associates openness with structures and institutions that scientists may distrust. There were signs in interview that disengagement from open access and other centrally regulated forms of openness are due to suspicion of financial arrangements and a sense of injustice rather than objections to openness practices themselves. This is analogous to findings by STS scholars that distrust in scientific institutions is a prominent reason for publics to lack trust or acceptance of scientific knowledge (Irwin and Wynne 1996).

Finally, enforcement of openness reduces agency, resulting in unease and defensiveness when openness requires labour or vulnerability. Moreover, bureaucracies for large-scale management of openness reinforce framings of open practices as administrative rather than scientific. In **Table 7** I present a summary of comparisons between openness categories, and these observations about engagement. However, the factors above are not the only ones that mediate engagement. One crucial factor can be understood through the lens of epistemic virtue.

Table 7 | Comparison of the three main openness categories from scientist interviews, leading to the identification of factors influencing engagement.

	Open access	**Data openness**	**Interpersonal openness**
Stage and interactivity	Downstream: "opens" settled, published findings; low interactivity	Mid-process: "opens" science-in-the-making; variable interactivity	Upstream: "opens" ideas and process; high interactivity
Scale, infrastructure and standardisation	Large-scale, standardised governance and infrastructure	Mixture of large- and mid-scale standards and infrastructure: interaction of top-down and community governance	Research group and individual-scale governance without standardisation or infrastructure
Enforcement and agency	Enforced top-down; bureaucratic; limited agential involvement (especially UK)	Expected/encouraged top-down and in some communities; some bureaucracy; agential involvement required	Unenforced (bottom-up); some lab-level expectations; agential and anti-bureaucratic
Emotional and attitudinal range	Enthusiasm, approval, acquiescence, passivity, scepticism, frustration	Passion, principle, approval, acquiescence, scepticism, unease, defensiveness, vulnerability	Passion, principle, joy, excitement, circumspection, protectiveness, tension
Overall engagement (+) factors that may promote engagement; (-) factors that may discourage engagement	*Tends towards passive acceptance*; (+) visibility of work; (-) low epistemic/social entanglement, low contextual relevance, institutional/structural distrust, compliance framings	*Full spectrum*; (+) local, trusted governance & infrastructure, high epistemic/social entanglement, high agency; (-) standardised governance, distrust, low epistemic/social entanglement, low agency	*Outside policy/engagement discourse*, but experience is more involving with (+) interactivity, agency, epistemic/social entanglement, trust

9.3 Sensitising to epistemic virtue

The comparisons above show that interviewees' levels of engagement with scientific openness was highly varied, and that disengagement might be productively reframed through varied lenses such as lack of epistemic/social interactivity; lack of contextual relevance; distrust of institutions or systems; and lack of agency. In this section, I will explore whether my sensitising concept, epistemic virtue, can bring any extra depth and explanatory power. This involves directly addressing sub-question (c): *to what extent is "open" (or "openness") being constructed as an epistemic virtue?*

9.3.1 Finding virtue in open access?

Open access was a distinctive case. It was widely recognised, accepted and practised by scientists in the study context, especially in the UK. Thus in many ways it represented a runaway success story all those who wish to see a wider dissemination and reuse of research. Cameron[174], a high profile open science advocate and researcher, noted this achievement with nuance:

> *"...there's been various criticisms [...] But honestly I think, massive increase in accessibility of content fundamentally in a decade, 15 years, is actually not a bad achievement [...] Access was always an important part of open access* [laughs]. *I'm pretty comfortable with the notion that that's on the way to being delivered."*

The analyses above, however, emphasise strain and conditionality in scientists' relationships with open access. This was underlined by Danny Kingsley's[175] reckoning that "voluntary engagement" with open access would drop dramatically if policies were removed. This conditionality can be viewed through the lens of epistemic virtue – or rather, its absence. For many interviewees, open access was not good science because it was not science at all: it had value, but it was separate; it was publishing; it was money and paperwork; it was "a little thing" (Richard) compared to the science.

[174] Interview: UK, Feb 2018.
[175] Head of Scholarly Communication, University of Cambridge; interview: UK, Feb 2018.

Moreover, open access was not typically a concern of the scientific self. Indeed, the process could be largely handled without scientists. Erin was one of several interviewees who made it clear that one can practise open access without being an "open researcher" who is "active engaged and willing to be open" (see Section 8.3). This disconnection from epistemic virtue is another way to understand the conditionality in scientists' relationships with open access: why acquiescence was more common than passionate commitment; and why sacrifices of cost or journal choice were not easily made for this practice.

However, an absence of epistemic virtue was not the whole story. For some scientists open access is closely entangled with knowledge-making, because articles take on the role of primary data. Accessing and computationally processing articles – at scale – is pivotal for this kind of research. Peter Murray-Rust[176], a veteran open science advocate and computer scientist, described making knowledge in this way: "content mining, which is the use of the whole scientific literature - the whole scholarly literature - but particularly science, to create new knowledge artefacts". Additionally, Sarah[177] - a journal editor who was involved in early open access movements – indicated that the biomedical scientists who founded PLOS were initially motivated by these kinds of epistemic goals rather than accessibility for the public:

> *"…they were doing genome-scale experiments, and doing microarrays: they've got hundreds of thousands of things they want to find the references for, and doing a sort of pay-per-view option is just not reasonable - or even working out how you get round the paywall in each case is not reasonable."*

Open access is entangled with many other values and causes including, vitally, resistance to big publishers. In terms of engagement, however, it is illuminating to consider that some differences in scientists' orientations to open access may depend on project- or field-specific differences in how "good science" is made.

[176] Interview: UK, Feb 2018.
[177] Interview: UK, Jan 2018.

This specific epistemic virtue framing is not end of the story, either. Some interviewees were proactively committed to open access without it being a direct and intimate part of their knowledge-making process. There were moral and epistemic elements to these commitments: open access was sometimes framed as the right thing to do for scientists with less access, or for publics, or because it aligns with other "open" values and principles. This is an interesting challenge to Daston and Galison's (2007) concept, which engages at the level of the scientific self. When a virtue is internalised not for its truth-telling value in one's own scientific endeavours, but for its assumed benefit at a systemic level, is it an epistemic virtue? Or a moral or social virtue, without epistemic entanglement? This question speaks not only to contemporary open science, but to the much longer-held idea that communal visibility of findings is essential to science (see Chapter 3).

Elsewhere – but absent in *Objectivity* (2007) – Daston (1998) identifies "communitarian objectivity" that "urges scientists to standardize their instruments, clarify their concepts, and depersonalise their writing styles to achieve communicability and commensurability across continents and centuries..." (p. 88). This certainly resonates with aspects of contemporary open science, but it does not capture commitment to open access on equality grounds, which might be considered more moral than epistemic: the right thing to do, whether or not it is good for science. This is different from the idea of better science through inclusion of diversity of voices and perspectives, which interviewees rarely alluded to (but see Section 9.4.4), and which open access, in current forms centred in the global north, may not be well placed to accomplish (Chan et al. 2020; Nabyonga-Orem et al. 2020).

Whether open access is considered a moral, social or epistemic virtue in these cases – and perhaps the distinction begins to break down at systemic level – the virtue- and identity-associated framing introduced by epistemic virtue is useful, as it captures an emerging, generationally-specific internalisation of "open" values that extends to open access, but is centred upon open data and methodology.

9.3.2 Finding virtue in data openness?

Because data openness takes so many forms, theorising its relationship with scientists is not a single task. However, it was clear – as noted in Chapter 7 (see Section 7.8) – that some scientists internalise and pursue openness of data as an epistemic virtue. I could see this most clearly when opening data was constructed not only as a route to good science, but an act of reflective self-improvement for the scientist. It was also instructive to see data openness experienced as a form of vulnerability that was worth enduring for the improvement of both the scientist and their work. This level of "engagement" was distinct from the case of open access. In other cases, linked to pre-2000s genomic data-sharing norms, interviewees seemed "engaged", but in limited and specific ways. They seemed to experience openness as a managed process of exposure that limited vulnerability through standardised formats, infrastructures, and community expectations. Defensive rejection of data openness suggested a conflict that went beyond frustration: vulnerability of the (scientific) self, and perhaps a challenge to the epistemic virtues anchoring that sense of self, which may be in tension with openness. For example, some expressed a view that pre-publication data sharing threatens the integrity of knowledge-making: Cedric[178] drew on the Scottish saying, "fools and bairns should never see a job half-done".

The internalisation of data openness as an epistemic virtue had a particularly strong tie to rationales of transparency and reproducibility, and seemed to incorporate methodological openness including open code, preregistration, and occasionally open notebook science, as well as extending to assume open publication, and often an open interpersonal approach. As I have argued, this indicates that contemporary open science is gaining traction amongst more recently trained scientists. I suggest that an epistemic virtue framing helps to articulate what this "traction" might consist of: a strong, moralised, identity-associated conviction likely to be long-lasting in individuals, and instilled in their trainees and other close colleagues. This depth of engagement may be necessary for data openness to be practised with a purposeful

[178] PhD 1960s; interview: UK, Feb 2018.

integrity that makes it worthwhile. Enrico[179] articulated the practical difference this can make:

> *"I can supply a code of what I did, it's open, it's accessible. I can supply a code that's a lot of comments for the support document to help documentation, that's also open. But clearly the latter is helping out [...] And that really generally has to come from the researcher thinking this can help and not because I'm forced to put my data because the reviewer, or the publisher, or the grant body have asked."*

As Enrico and others point out, there is potential for openness of this kind to be enacted instrumentally. In some cases interviewees were aware of strategies employed to limit the openness of "open data", such as deliberate withholding of parts of a dataset, or a lack of annotation make the data understandable to others. Even without such strategies, some scientists approached data sharing as a tick-box task, without expectation that the opening of data would improve their own work or others'; or they did not share data at all for the same reasons. This is important to understand from a policy perspective, since top-down governance is limited in its ability to motivate – or perhaps even distinguish – the kinds of openness that are most epistemically caring and enriching. Moreover, the depth and specificity of engagement required cannot be displaced from data creators. This certainly makes open data an ambitious policy goal, and one that is distinct from open access. It is helped, however, by the fact that data openness is a task close to the epistemic self.

9.3.3 Finding virtue in interpersonal openness?

Because interpersonal openness is so different from the categories above, it is an illuminating point of comparison. Many interviewees' relationships with interpersonal openness suggested its internalisation as an epistemic virtue: they connected it to formative elements of knowledge-making and the nourishment of their scientific identity and purpose (see Sections 8.4.1 and 8.5.6). These experiences were characterised by excitement, joy, moral conviction, and a desire to pass this approach on. Adherents to this approach often saw it as a fundamental quality distinguishing scientists. Some interviewees who felt unable to practise interpersonal openness still

[179] PhD 2010s; interview: UK, Nov 2018.

framed it as an epistemic virtue. This tension was tied, in younger scientists, to ambivalence about continuing in a scientific career, which emphasises a link to identity and purpose. However, others expressed no systematic commitment to the practice. Interpersonal openness seemed to operate as an epistemic virtue, and a moral economy, in certain research groups, communities and networks.

Distinctive to this epistemic virtue was its embodied quality: it was an opening – a vulnerability, receptivity, exposure – of the self. This gave it with a particularly close tie to identity: scientists who internalised this virtue were not just practising openness – they were *being* open. This is significant in contrast with online open science, which is characterised by the opening of information objects, and their abstraction and separation from their human creators: in Nielsen's (2009) terms, moving information "out of people's heads and labs, onto the network" (p. 32), and in Bartling and Friesike's (2014b) terms, enabling knowledge to "flow quickly, regardless of institutional and personal networks" (p. 8). These narratives imply that there is a technological solution to making science open; and that scientists' connections with their knowledge hamper rather than enable openness.

Interpersonal openness presents an illuminating counterpoint because it is a thoroughly human practice of scientific openness, in which social entanglement is both problematic and a source of profound epistemic enrichment. It is also an example of an openness practice that many scientists pursue energetically in the absence of external incentives, and despite the risk it entails. Moreover, accounts of interpersonal openness show that many scientists care deeply about sharing ideas and data well upstream of publication, but may only do so when they feel secure. Interviewees' accounts also show how a sense of security might come about in scientific careers and communities, and thus how cultures of openness might be nurtured. None of this makes interpersonal openness a model "open" epistemic virtue. In many ways, it represents the opposite of what open science movements seek: universally accessible, transparent, permanently recorded knowledge. However, in the context of difficulty "engaging" scientists with online open science I suggest that lessons might be learned from interpersonal openness. Communities in which it has been widely internalised as an epistemic virtue may even have laid the cultural groundwork required for online forms of openness to thrive (see examples in Section 7.10).

9.4 What "openness" hides: distinct virtues within?

Through the discussion above, I have considered the question of scientists' (dis)engagement with open science (d); the construction of scientific openness as an epistemic virtue (c); and an illuminating connection between the two. Here, I extend my focus on epistemic virtue in one further direction. I consider what happens if the "openness" label is removed from epistemic virtues (above) to which scientists are committed, allowing these to be understood in different ways. I have analysed the breadth of meaning encompassed by scientific openness throughout this thesis, and others have proposed categorisations of the breadth of open science (e.g. Fecher and Friesike 2014). Here I propose four categories specifically grounded in scientist interviewees' constructions of openness as an epistemic virtue, and informed by historical and policy/advocacy analyses. In other words, these are epistemic virtues that an "openness" framing hides, which cut across practice-based categories like open access and open data. They are summarised in **Table 8**.

9.4.1 Commoning

One distinct purpose and value of scientific openness is to make knowledge resources communally available, for use in future knowledge-making. This purpose characterises centuries-old scientific openness in Western Europe associated with the tradition of journal publication (see Chapter 3). It is also associated with the idea of "standing on the shoulders of giants" (Newton 1675); and Merton's norm of communism (1973 [1942]). In the twentieth century, this type of openness is also manifest in the moral economies of model organism communities (Kohler 1999), their newsletters (Kelty 2012) and later online databases (Leonelli and Ankeny 2012); and in the public Human Genome Project (see Chapter 4). It is also, of course, a purpose that drives contemporary open access and data movements. I have called it "commoning" in reference to prevalent discourses that characterise open online resources as knowledge or information commons (Lawson 2019; Moore 2018). "Commoning" places critical emphasis on "the practice of cultivating and caring for the relationships that exist *around* the production of shared resources" (Moore 2018:17), as distinct from the treatment of commons as decontextualised, free, universally appropriable resources, as they are often configured in open movements (ibid). Interview accounts showed that the latter arrangement – which does not entail community ownership, reciprocity, and trust – tends not to engage scientists. More epistemically, socially involved forms of resource sharing governed at smaller scales engage scientists more (see Section 9.2.5), and are more likely to be constructed as epistemically virtuous. Thus, "commoning" captures one distinct way in which openness is pursued by scientists for the improvement of knowledge-making their communities. Daston and Galison (2007) define epistemic virtues in relation to problematic tendencies of the self that they counteract, in order to seek truth. I propose that commoning is a virtue pursued to counteract tendencies of isolation, idiosyncrasy, and individualism.

9.4.2 Transparency

Distinct from commoning is transparency: exposing processes of knowledge-making so that their validity can be examined and improved. This purpose is salient in open data movements, alongside commoning (or at least, the creation of "commons"), and is a sensibility distinguishing today's pursuit of open data[180] from pre-2000s digital data sharing traditions in genomics. Of the four categories I propose, transparency is most identifiable as an epistemic virtue in Daston and Galison's terms, perhaps because it could be considered a type, or extension of, objectivity. The transparency component of open data, for example, is focused towards detecting and kerbing error, bias, fraud and other problematic entanglements of the scientist with their work. Where mechanical objectivity (Daston and Galison 2007:20) encourages the scientist to self-examine and suppress biasing tendencies, transparency no longer trusts the

[180] See Section 5.5.2: I documented rise transparency/accountability connotations of openness as open data joins open access as a policy issue).

scientist alone with this task: the eyes of a community (perhaps including online "data thugs", see Section 7.10), and the panoptical eye of public visibility, are turned in this direction.

Transparency is closely linked to reproducibility in scientists' accounts (Section 7.8.2), and the need to prove the integrity of knowledge-making in response to "reproducibility crises" (Section 5.3.2). There is an implication here that good knowledge-making must be amenable to exposure and extraction from context. An extreme manifestation of transparency as an open epistemic virtue is the automation of data sharing to eliminate human imperfection from this process (Kekecs et al. 2019), and the removal of a need for trust (observed by Gabrielsen 2020), although these are not inevitable directions. Transparency and reproducibility were strongly tied to scientist interviewees' epistemic virtue framings, and seemed to drive principled and passionate orientations towards open science in general (see Section 7.8). After my analysis, I found that Freese and Peterson (2018) also link reproducibility concerns with the emergence of a new epistemic virtue in Daston and Galison's sense, which they describe "statistical objectivity". This is a "meta objectivity" focused upon populations of studies, rather than individual studies, that configures scientists as "economic actors led into bed practices by a poorly aligned system of incentives" (Freese and Peterson 2018:293). This is an interesting synthesis overlapping with the "transparency" category I propose as well as my observation that researchers are portrayed as incentive-driven in open discourses (see Section 5.5.4). However, my analysis of epistemic virtue focuses more on openness than reproducibility, and is based on scientists' interview accounts, where Freese and Peterson draw primarily on literature and academic activism associated with the "crisis of false positives" in social psychology.

9.4.3 Collegiality

A further distinct open epistemic virtue is collegiality, which involves offering oneself and one's embodied knowledge generously to a scientific community. It is a defining feature of interpersonal openness, but incorporates other community-oriented practices identified by interviewees. It involves building and nourishing communities by fostering trust and collective identity; establishing networks; supporting colleagues and particularly students; and even designing physical environments to promote interaction (as one interviewee, David, had done: I borrowed his use of "collegiality"). Collegiality improves knowledge through free-flowing idea-generation; creativity; excitement; and the creation of settings in which this can occur. As an epistemic virtue it is dissimilar to objectivity as identified by Daston and Galison: although it is identity-defining and tied to truth-seeking, it builds upon rather than suppressing human influence. It exposes knowledge-making not through surveillance, but through relational vulnerability, and is strongly trust-mediated. Where objectivity, and its "open" extension – transparency – seek to counteract error-prone tendencies of the self, collegial openness is trusting of the self and its insight, and of other community members. It thus has self-possessed qualities that contrast with some stereotypic characterisations of science. If it guards against any tendencies of the self, they might be distrust, competitiveness, narrowness, isolation or stagnation.

9.4.4 Cognitive justice

This fourth open epistemic virtue is only thinly represented in my interview data, but signposts an important, newer direction in open science discourse that has been centred in the global north (see Chan et al. 2020; and Section 5.5.2). It has important distinguishing features that variously challenge or complement the categories above. Cognitive justice is "open to all knowledge and all epistemologies, and not an abstract universalism based on Western standards that exclude what is different from themselves" (Piron, 2018, cited in translated form in Chan et al. 2020). This virtue acknowledges the inequalities that are embedded in knowledge-making systems, and has feminist, decolonial sensibilities. It might be embodied by a reflexive, situated, empathetic self that seeks truth(s) by counteracting tendencies towards narrowness, assumed omniscience, supremacy, and perhaps – objectivity. It is self-examining but not self-denying, and with parallels to Haraway's (1988) feminist objectivity . This virtue is in conflict with some others, especially transparency as outlined above: an uncomfortable implication for some open science movements.

Table 8 | Proposed categories of epistemic virtue that underlie or cut across "openness" categories identified by scientists in interview. Cognitive justice framings were uncommon in interview, but represent an important emerging, contrasting direction in open science.

Types of virtue	**Commoning**	**Transparency**	**Collegiality**	**Cognitive justice**
Description	Communal resource sharing, especially in community-managed, trusting, reciprocal settings	Making detail and processes of knowledge production transparent for inspection and reproduction	Making oneself and one's knowledge available as a colleague; cultivating trust and community; supporting colleagues	Acknowledging contextual embeddedness of oneself and one's knowledge; addressing effects of systemic inequality
Linked practices	Open access; open preprints; open data; data stewardship; open materials (some of these are more "commons" than "commoning)	Open data; open code; open methods; pre-registration; open notebook science	Interpersonal openness; collaborative practices; fair authorship; proactively supporting students	Representing and including diversity; reflexivity; receptivity to other views and epistemologies
Tendency of scientific self that it counteracts	Isolation, idiosyncrasy, individualism	Bias, imperfection, error, fraud (extension of objectivity)	Distrust, competitiveness, narrowness, isolation, stagnation (nourishes: excitement, joy, care)	Narrowness, omniscience, supremacy, objectivity (nourishes: reflexivity, empathy)
Example quote	"...the more scientists especially that have access to research data, resources, publications, the better your science will be..." (Jason)	"...if they're not open, and you're not able to actually really examine, drill into that data in more detail, you start to worry about the...yeah, the integrity of their science." (Mark)	"I think if people were less worried about somebody, I don't know, maybe...using their ideas, or wanting to come on board, I think maybe there'd be a lot more idea sharing." (Nicole)	"...it's way better at the end of the day because you have multiple people with multiple perspectives and expertise. And that's going to produce better science than what's been done previously." (Erin)

9.5 Contextualising virtue: insecurity, inequality, and culture

There is a vital counterpoint to the construction of openness as one or several epistemic virtues: virtuous orientations are not equally accessible. I have explored this topic in depth in Chapters 7 and 8 in relation to data and interpersonal openness. The scientists I spoke to – even those in positions of significant power – tended to feel beholden to a "system" through which jobs, promotions, grants and prestige are distributed. This system was associated with feelings of anxiety, scarcity, precarity, and pressure. An environment of this kind in academia is increasingly well-documented, and related to trends such as the rise of the "entrepreneurial university"; short term job contracts and the projectification of research; metrics and rankings including the journal impact factor and its use in performance assessments; and a large population of PhD graduates compared with available academic jobs (Felt 2017; Müller and De Rijcke 2017; Schönbauer 2019, 2020; Slaughter and Leslie 1997).

Interview accounts showed how such an environment impacts on openness. Opening up some stages and types of knowledge-making was associated with acute risks and anxieties about scooping and transparent exposure. As I have discussed, made openness easier for scientists who are more secure than others in a pressured system. In Chapter 7 I identified short- to long-term factors that seemed to shape this feeling or experience of security (see **Table 6**). Importantly, some forms of security not only made openness *feel* easier, but buffered its risks. Scientists in insecure positions may not be able to afford such risk-taking. Unequal access to openness is likely to be reinforced by the fact that openness is often rewarding – both personally and professionally – as well as risky. Open access was easier for scientists in wealthy labs or with access to wealthy institutional systems: thus, the benefit of greater visibility through openness also accrues to those scientists. The effect was more notable in relation to data and interpersonal openness, which were epistemically and socially enriching. Scientists who embraced interpersonal openness in particular spoke of its reciprocal rewards and long-term benefits. Current proposals to promote open scientific cultures top-down by creating rewards for open practices (e.g. LERU 2018; Wilsdon et al. 2017) risk feeding into this feedback loop that reinforces the security of scientists who already had the privilege to be open. Inevitably, woven through these privilege dynamics are systematic societal inequalities, which likely make openness more accessible to scientists at intersections of power in gender, race, ethnicity, disability, sexual orientation, class, and so on.

One final form of privilege is worthy of particular mention because it may be a means to rebalance – although not overcome – others. There was evidence that supportive role modelling, mentoring, and community dynamics had a formative effect on scientists' feelings of security in their early careers. It may be the case that risks associated with openness can be buffered in a community that builds confidence and generosity in individuals, and provides a secure foundation from which to experiment with the risks and rewards of openness. Conversely, experiences of scarcity, competition, and isolation seemed to promote long-term protective approaches.

9.6 Comparison with similar studies

The findings of my interview study in Chapters 6-8, and their synthesis in this chapter, can be productively compared with the findings of similar studies mentioned in the Introduction (see Section 1.9.1). The most relevant comparison is to Levin et al. (2016), because their interview study with 22 UK biomedical researchers employed a similar broad framing and in-depth qualitative methodology in which interviewees were asked about their understanding "openness" in science (p. 3). The thematic analysis conducted by these authors yielded seven themes characterising understandings of openness, and nine factors shaping openness practices. These themes and factors were broadly consistent with my observations: for example, interviewees across both studies highlighted the importance of "timely donation of and access to research components" and "access to research components in non-Western and/or Nonacademic contexts" (themes identified by Levin et al.). Interviewees from my study also concurred that factors such as "the existence of repositories and databases for data, materials, software, and models", "the competitiveness of academic fields", and "credit systems in academic research" shape experiences and practices of openness. My study also aligned with overall conclusions of Levin et al., including that researchers' decisions about openness are highly contextual, especially for data, and that open policies which are too stringent or standardised may be counterproductive.

There are also key differences and complementarities between my findings and those of Levin et al. (2016), which highlight the contribution of my research. These seem to arise from differences in analytical approach, theoretical lens and sampling strategy. Most obviously, my analysis took a different form to that of Levin and co-authors. My use of multiple analytical techniques, including but going beyond *NVivo* coding, resulted in a decision to focus on three highly salient themes (open access, data openness and interpersonal openness) that incorporate many sub-themes and capture a cross-section of experience. This focus enabled me to study scientists' *relationships* with each of these categories in depth, sensitised by epistemic virtue and experiences of scientific self. This in turn enabled the synthesis in the current chapter, which goes beyond these three categories to identify underlying virtues of "openness" that scientists uphold in their practices and identities. The analytical process described by Levin et al. seemed instead to rely more directly upon thematic coding ("...the authors coded the interviews with the software program NVivo. The selected themes were included in the study if they were referred to by at least three different interviewees", p. 3). As a result their findings are best summarised as a collection of themes and factors associated with openness that are difficult to cluster due to their variety. In contrast, my study features more thick description evoking openness *as constructed by scientists*; continuities and contrasts in their relationships with openness, including affective dimensions; and systematic conceptualisation of why experiences and identities in relation to openness are diverse, including consideration of factors that may be implicit (such as vulnerability).

The different process and form of my analysis compared to that of Levin et al. (2016) is also associated with at least one clear novel finding in my study: the identification of interpersonal openness as a category that is highly salient and meaningful for scientists. There is a hint that interviewees in Levin et al.'s study raised similar ideas, as those authors identify "collaboration and cooperation with peers and communities" as a theme. This theme touches on a link between openness, informal sharing, and collaboration, but does not capture the distinctiveness and salience of the category in my research, which includes its relative absence from established "open science" discourse, and its coherent links to the sharing of unpublished ideas, trust, reciprocity, generosity, creativity, vulnerability, and a fulfilment of scientific self.

There are likely several reasons why this category did not arise with the same distinctive salience in the study by Levin et al. (2016), despite the similar breadth and framing of their study. Firstly, my analytical focus on scientists' relationships with openness promoted a focus on personally meaningful and identity-defining categories such as interpersonal openness. Secondly, like other similar exploratory, qualitative studies on this topic – such as Scheliga and Friesike (2014) and Ali-Khan et al. (2017) – Levin et al.'s analytical lens may focus their attention on established categories of "open science" over alternative interpretations by scientists that do not fit coherently with current advocacy and policy discourses. Despite this, Levin et al. were more likely than others to encounter interpersonal openness due their use of "openness" rather than "open science" as an interview prompt (I reflect on this in Section 8.3). Finally, my inclusion of interviewees without pre-existing links to open science movements may have been a pivotal difference: both Levin et al. and Scheliga and Friesike (2014) recruited interviewees through their connections with open science. This would have made it less likely that interviewees would raise alternative categories of openness, even within a broad, exploratory study. This difference in sampling may also have shaped a greater focus by Levin et al. on themes reflecting "expert" knowledge of open policy and advocacy, such as metadata, standard formats, and open-source approaches to intellectual property. In sum, breadth in both participant input to my study and my analytical gaze, with a particular emphasis on how "openness" is constructed outside advocacy and policy perspectives, likely contributed to my identification of interpersonal openness. This approach and finding proved valuable in making sense of many scientists' lack of "engagement" with open science – partly by showing that openness *is* of deep interest to most scientists, but perhaps not in the ways that are expected.

Levin and Leonelli (2016) analysed selected aspects of the interviews by Levin et al. (2016) in more depth and concluded that openness is a mode of valuing research, rather than a value or virtue in its own right. This conclusion is complementary with aspects of my own findings: throughout the thesis I have observed that openness connotes a diverse, expanding and internally contradictory range of values and practices; it is not coherent or self-explanatory as a value, but rather plays a social and sematic role as a rallying point for meaning, through its flexibility and the history of its use in contexts connected to science. Levin and Leonelli (2016) also make nuanced observations of scientists' negotiations of openness and closure; how different aspects of the research process are valued; and the risk that open science policies will disrupt the "close and ever-changing relationships" between scientists and the aspects of research they value and have laboured over (p. 18). These observations strike a chord with my own, especially those in Chapters 7 and 8, which touch on partial and finely calibrated forms of openness, and the value that is gained and lost through opening up or holding back personally meaningful data or ideas.

Levin and Leonelli (2016)'s conclusion about value in relation to openness is thus productive in light of my own research as well as theirs. My conceptualisation of openness in relation to epistemic virtue does not challenge Levin and Leonelli's conclusion. Instead, it offers a new layer or depth of understanding of scientists' *relationships* with openness. While openness itself may not have essential qualities as a value or virtue in science, my conclusion is that many scientists assign significance to the idea and/or terminology of openness; that it plays a role in their understanding of "good science" and being a "good scientist"; and that these understandings can be collective, albeit situated, shaping consensus views of epistemic virtue in scientific communities. This is not to say that "openness" has one meaning when it is upheld or internalised as a virtue in these settings. It means many things, but is loosely bounded by recurring features in scientists' experiences, such as an experience of epistemic exposure that is both vulnerable and potentially rewarding. It also connects and perhaps unites several distinct principles that scientists may already consider important to their practice and identity (such as commoning, transparency, collegiality, and cognitive justice; see Section 9.4). All of the nuances of openness and its situated value, as identified by Levin and Leonelli, remain relevant. But in this study, I showed that – alongside processes and outputs of research – scientists' identities and notions of what is "right" when making knowledge are also subject to openness as a "mode of valuation" (ibid., p. 17). Thus, these identities and moralities are shifting in complex ways with the rising salience of open science movements.

9.7 Conclusion: an open science "revolution"?

This study is set within the context of the so-called open science "revolution" (Bartling and Friesike 2014b; Nielsen 2009; The Royal Society 2012) in which scientific practice is expected, by some advocates and policymakers, to transform on a scale not seen for centuries. I have critiqued and contextualised this narrative, but it remains the case that some kind of transformation is happening as open science becomes a priority in research governance, and advocate voices embrace more and more varieties of openness. My research enables me to comment on the "cultural" aspect of this transformation. The internalisation – or otherwise – of openness as an epistemic virtue suggests the extent to which open science is transforming ways of *being* as a scientist, and deeply-held conceptions of good science. In sum, based on the analyses above, there is at present no consensus on the virtue of contemporary open science, at least amongst a disciplinarily broad population of biologists in the study context. Open access movements by themselves appear not to have had a widely, deeply felt impact on internalised notions of "good science". However, the idea of an "open" epistemology seems to have gained traction as open data movements rise to join open access, with accompanying movements seeking transparency and reproducibility. The main indication of this change - a particular principled, passionate stance of a group of more recently trained interviewees – suggests the beginning of a transition marked by a re-making of epistemic subjects: a generation of "open" scientists. This does not suggest a natural ascendency, but an ongoing, multi-sited, cross-generational, value-based negotiation or contest in which prevailing conceptions of good science are challenged and may be displaced.

The shaping of epistemic subjects is a robust indicator of deep cultural change. However, this cultural shift – if it is indeed taking place – is fragile, and in its early stages. It depends on the trajectory that younger scientists take through their careers, the effects of these journeys on their relationship with "open", and the nature of the ongoing influence from open science advocacy and policy movements. Moreover, several implications of this shift require ongoing attention and critique. Firstly, as I have argued, current pressures, diverse epistemic contexts, and systematic inequalities in academic research make openness unequally available, and may amplify privilege. Top-down incentives to be open may magnify this effect. Scientists who are less secure in their careers; who are isolated from support; who feel unable to trust their communities; or who work with less standardisable data and methods, are at risk of being disadvantaged in a shift towards "open". This, in turn, may entrench a need for self-protective closure. This applies to all career generations: not all recently trained scientists were in a position to internalise open values, and some continued to receive formative training in highly competitive, unsupportive, or precarious settings.

Secondly, the nature of the "open" epistemology rising to prominence requires ongoing attention. My analysis suggests that it is predominantly influenced by a vision of transparency and reproducibility that extends a traditional, objective "view from nowhere" (Daston and Galison 2007; Haraway 1988). This view may narrow or lock out alternative ways of knowing. Leonelli (2018) argues that a focus on reproducibility reduces researchers' attention to local, idiosyncratic, and variable aspects of their research, as well as "devaluing the role of expertise and embodied knowledge in data production, processes and assessment" (p. 13). This type of openness is thus in conflict with a decolonising view of open science framed by cognitive justice (see Chan et al. 2020 and Section 9.4.4). Moreover, the mechanism by which this type openness governs integrity – and produces "good science" – is analogous to panoptical surveillance (Foucault 1991): the creation of a system in which scientists avoid imperfections, errors and malpractice because they face the ongoing potential of exposure. This mechanism may build trust in outputs of science, but it operates by institutionalising distrust amongst scientists, and contains a logic by which the elimination of trust – and of human influence – is ideal. This conflicts with a collegial, generous, trusting view openness that nurtures community. There is also an emphasis in open science discourse and policy on the creation of universal information "commons" rather than on "commoning" practices that acknowledge the social entanglements of knowledge production and sharing (Moore 2018).

These are not inevitable directions of "open" epistemologies. My interview study indicated that scientists deeply inspired by "open" sometimes had broad and imaginative views that went beyond transparency and reproducibility, and sometimes included a deep sense of moral responsibility towards publics; a receptivity to collaboration and outside ideas; and a community-oriented, supportive mind-set. These scientists also tended to combine an open approach to interpersonal interactions with their online openness. Interpersonal openness is imperfect, but its partial diffusion as an epistemic virtue through scientific communities suggests – almost maps – how current scientific cultures are locally enabling or constraining collegiality and generosity. A healthy, equitable embrace of online open science surely depends on the capacity for scientific communities – and the individuals that make them up – to be collegial, generous, and supportive of diversity. For an open science revolution that raises up – and arises from – a diverse community of scientists in all their humanity, we must think deeply about the kinds of openness we hope future generations of scientists will embody as they practice "good science".

See Appendix A for recommendations to open science advocates and policymakers.

References

Abir-Am, P. G. 1997. "The Molecular Transformation of Twentieth Century Biology." Pp. 495–524 in *Science in the Twentieth Century*, edited by J. Krige and D. Pestre. Amsterdam: Harwood.

Adema, Janneke and Samuel A. Moore. 2018. "Collectivity and Collaboration: Imagining New Forms of Communality to Create Resilience in Scholar-Led Publishing." *Insights* 31(3):1–11.

Agar, Jon. 2012. *Science in the Twentieth Century and Beyond*. Cambridge UK and Malden USA: Polity Press.

Ali-Khan, Sarah E., Liam W. Harris, and E. Richard Gold. 2017. "Motivating Participation in Open Science by Examining Researcher Incentives." *ELife* 6(e29319):1–12.

Altman, Lawrence K. 1996. "The Ingelfinger Rule, Embargoes, and Journal Peer Review-Part 1." *The Lancet* 347:1382–86.

Ankeny, Rachel A. 2001. "The Natural History of Caenorhabditis Elegans Research." *Nature Reviews Genetics* 2:474–79.

Ankeny, Rachel A. 2010. "Historiographic Reflections on Model Organisms: Or How the Mureaucracy May Be Limiting Our Understanding of Contemporary Genetics and Genomics." *History and Philosophy of the Life Sciences* 32(1):91–104.

Ankeny, Rachel A. and Sabina Leonelli. 2011. "What's so Special about Model Organisms?" *Studies in History and Philosophy of Science Part A* 42(2):313–23.

Anon. n.d. "Open Data: Revision History." Retrieved June 15, 2020 (https://en.wikipedia.org/w/index.php?title=Open_data&dir=prev&action=history).

ARC. n.d. "Discovery Projects: Funding Rules for Funding Commencing in 2008." Retrieved November 1, 2020 (https://web.archive.org/web/20070829003846/http://www.arc.gov.au/pdf/DP08_FundingRules.pdf).

Archer, Margaret. 2003. *Structure, Agency and the Internal Conversation*. Cambridge: Cambridge University Press.

ASAPbio. 2019. "Preprints Are Valid Research Outputs for REF2021." Retrieved October 27, 2020 (https://asapbio.org/preprints-valid-for-ref2021).

Attenborough, Rosalind. 2015. "Ethos-in-Making: A Pilot Study of Openness in the Professional Practices and Identities of Biologists." University College London.

Aymar, Robert. 2009. "Scholarly Communication in High-Energy Physics: Past, Present and Future Innovations." *European Review* 17(1):33–51.

Baker, Monya. 2016. "Is There a Reproducibility Crisis?" *Nature* 533:452–54.

Balmer, Brian. 2012. *Secrecy and Science: A Historical Sociology of Biological and Chemical Warfare*. England: Ashgate Publishing Company.

Barbour, Virginia. 2016. "AOASG Response to Productivity Commission Issues Paper on Data Availability and Use." *AOASG*. Retrieved June 9, 2020 (https://aoasg.org.au/2016/08/05/aoasg-response-to-productivity-commission-issues-paper-on-data-availability-and-use/).

Barnes, S. B. and R. G. A. Dolby. 1970. "The Scientific Ethos: A Deviant Viewpoint." *European Journal of Sociology* 11(1):3–25.

Bartling, Sönke and Sascha Friesike. 2014a. *Opening Science: The Evolving Guide on How the Internet Is Changing Research, Collaboration and Scholarly Publishing*. Springer Open.

Bartling, Sönke and Sascha Friesike. 2014b. "Towards Another Scientific Revolution." Pp. 3–15 in *Opening Science*, edited by S. Bartling and S. Friesike. Springer International Publishing.

Bastian, Hilda. 2016. "Reproducibility Crisis Timeline: Milestones in Tackling Research Reliability." *PLOS Blogs: Absolutely Maybe*. Retrieved June 5, 2020 (https://blogs.plos.org/absolutely-maybe/2016/12/05/reproducibility-crisis-timeline-milestones-in-tackling-research-reliability/).

BBC. 2010. "Tim Berners-Lee Unveils Government Data Project." *BBC News*. Retrieved June 8, 2020 (http://news.bbc.co.uk/1/hi/technology/8470797.stm).

Beall, Jeffrey. 2012. "Predatory Publishers Are Corrupting Open Access." *Nature* 489(7415):179.

Begley, C. Glenn and Lee M. Ellis. 2012. "Raise Standards for Preclinical Cancer Research." *Nature* 483:531–33.

Berlin Declaration. 2003. "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities." Retrieved June 23, 2020 (https://openaccess.mpg.de/Berlin-Declaration).

Berlin, Isaiah. 1953. *The Hedgehog and the Fox: An Essay on Tolstoy's View of History*. London: Weidenfeld & Nicolson.

Bernal, J. D. 1975. "Appendix A: Lessons of the War for Science." *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences.* 342(1631):555–74.

Berners-Lee, Tim, Robert Cailliau, Ari Luotonen, Henrik Frystyk Nielsen, and Arthur Secret. 1994. "The World-Wide Web." *Communications of the ACM* 37(8):76–82.

Bethesda Statement. 2003. "Bethesda Statement on Open Access Publishing." Retrieved June 23, 2020 (http://legacy.earlham.edu/~peters/fos/bethesda.htm).

Björk, Bo-Christer. 2012. "The Hybrid Model for Open Access Publication of Scholarly Articles: A Failed Experiment?" *Journal of the American Society for Information Science and Technology* 63(8):1496–1504.

Björk, Bo-Christer. 2017a. "Gold, Green, and Black Open Access." *Learned Publishing* 30:173–75.

Björk, Bo-Christer. 2017b. "Growth of Hybrid Open Access, 2009-2016." *PeerJ* 5:e3878.

Björk, Bo-Christer, Mikael Laakso, Patrik Welling, and Patrik Paetau. 2013. "Anatomy of Green Open Access." *Journal of the Association for Information Science and Technology* 65(2):237–50.

Blumer, Herbert. 1954. "What Is Wrong with Social Theory?" *American Sociological Review* 19(1):3–10.

BOAI. 2002. "Read the Budapest Open Access Initiative." Retrieved June 3, 2020 (https://www.budapestopenaccessinitiative.org/read).

Boulton, Geoffrey, Michael Rawlins, Patrick Vallance, and Mark Walport. 2011. "Science as a Public Enterprise: The Case for Open Data." *The Lancet* 377:1633–35.

Bowman, Nicholas David and Justin Robert Keene. 2018. "A Layered Framework for Considering Open Science Practices." *Communication Research Reports* 35(4):363–72.

Bradley, Jean-Claude, Kevin Owens, and Antony Williams. 2008. "CDI-Type I : Chemistry Crowdsourcing Using Open Notebook Science." *Nature Precedings* 1–8.

Brown, Brené. 2012. *Daring Greatly: How the Courage to Be Vulnerable Transforms the Way We Live, Love, Parent and Lead*. New York: Gotham Books.

Bryman, Alan. 2012. *Social Research Methods*. 4th ed. New York: Oxford University Press.

Bulock, Chris. 2019. "Open Dialog: SciELO's Approach to Open Access Publishing." *Serials Review* 45(4):245–47.

Bulpin, Kate and Susan Molyneux-Hodgson. 2013. "The Disciplining of Scientific Communities." *Interdisciplinary Science Reviews* 38(2):91–105.

Callaway, Ewen. 2011. "Report Finds Massive Fraud at Dutch Universities." *Nature News*. Retrieved November 1, 2020 (https://www.nature.com/news/2011/111101/full/479015a.html#update2).

Calvert, Jane. 2001. "Goodbye Blue Skies? The Concept of 'Basic Reserach' and Its Role in a Changing Funding Environment." University of Sussex.

Calvert, Jane. 2006. "What's Special about Basic Research?" *Science, Technology, & Human Values* 31(2):199–220.

Capshew, James H. and Karen A. Rader. 1992. "Big Science: Price to the Present." *Osiris* 7:2–25.

Carr, Dave. 2017. "Our New Policy on Sharing Research Data: What It Means for You." Retrieved June 8, 2020 (https://wellcome.ac.uk/news/our-new-policy-sharing-research-data-what-it-means-you).

Castells, Manuel. 2002. "Lessons from the History of the Internet." Pp. 1–28 in *The*

Internet Galaxy: Reflections on the Internet, Business, and Society, edited by M. Castells. Oxford: Oxford University Press.

CAUL and AOASG. 2019. "Joint CAUL-AOASG Election Statement: Developing a Strategic Approach to Open Scholarship in Australia." Retrieved (https://aoasg.files.wordpress.com/2019/05/caul-aoasg2019election-statement.pdf).

de Chadarevian, Soraya. 1998. "Of Worms and Programmes: Caenorhabditis Elegans and the Study of Development." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 29(1):81–105.

Chan, Leslie, Budd Hall, Florence Piron, Rajesh Tandon, and Lorna Williams. 2020. "Open Science beyond Open Access: For and with Communities, A Step towards the Decolonization of Knowledge (Version 1)." *Zenodo*. Retrieved November 1, 2020 (http://doi.org/10.5281/zenodo.3946773).

Chavan, Vishwas and Lyubomir Penev. 2011. "The Data Paper: A Mechanism to Incentivize Data Publishing in Biodiversity Science." *BMC Bioinformatics* 12(Suppl 15):52.

Chow-White, P. A. and M. Garcia-Sancho. 2011. "Bidirectional Shaping and Spaces of Convergence: Interactions between Biology and Computing from the First DNA Sequencers to Global Genome Databases." *Science, Technology & Human Values* 37(1):124–64.

Christensen-Dalsgaard, Birte, Marc van den Berg, Rob Grim, Wolfram Horstmann, Dafne Jansen, Tom Pollard, and Annikki Roos. 2012. "Ten Recommendations for Libraries to Get Started with Research Data Management."

Cobey, Kelly D., Manoj M. Lalu, Becky Skidmore, Nadera Ahmadzai, Agnes Grudniewicz, and David Moher. 2018. "What Is a Predatory Journal? A Scoping Review." *F1000Research* 7:1001.

Comprehensive Knowledge Archive Network. n.d. "Data-Gov-Au." Retrieved June 8, 2020 (https://ckan.org/portfolio/data-gov-au/).

Confrey, Eugene A. 1966. "Information Exchange Groups To Be Discontinued." *Science* 154(3751):843.

Cook-Deegan, Robert. 1994. *The Gene Wars: Science, Politics, And The Human Genome*. New York: W. W. Norton & Company.

Cragin, M. H., C. L. Palmer, J. R. Carlson, and M. Witt. 2010. "Data Sharing, Small Science and Institutional Repositories." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368:4023–38.

Creative Commons Wiki. 2011. "History." Retrieved May 21, 2020 (https://wiki.creativecommons.org/wiki/History).

Croissant, Jennifer L. and Laurel Smith-Doerr. 2007. "Organizational Contexts of Science: Boundaries and Relationships between University and Industry." Pp. 691–718 in *The Handbook of Science and Technology Studies*, edited by E. J. Hackett and M. E. Lynch. MIT Press.

Csikszentmihalyi, Mihaly. 1990. *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

Daston, Lorraine. 1998. "Fear and Loathing of the Imagination in Science." *Daedalus* 127(1):73–95.

Daston, Lorraine and Peter Galison. 2007. *Objectivity*. Brooklyn: Zone Books.

David, Paul A. 2008. "The Historical Origins of 'Open Science': An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution." *Capitalism and Society* 3(2):Article 5.

David, Paul A. 2014. "The Republic of Open Science: The Institution's Historical Origins and Prospects for Continued Vitality." Retrieved June 14, 2016 (http://econpapers.repec.org/paper/sipdpaper/13-037.htm).

Debat, Humberto and Dominique Babini. 2019. "Plan S in Latin America: A Precautionary Note [Preprint]." *PeerJ Preprints*. Retrieved November 2, 2020 (https://doi.org/10.7287/peerj.preprints.27834v1).

Delfanti, Alessandro. 2010. "Open Science, a Complex Movement." *SISSA – International School for Advanced Studies Journal of Science Communication ISSN* 9(3):1824–2049.

Delfanti, Alessandro. 2013. *Biohackers: The Politics of Open Science*. London: Pluto Press.

Delfanti, Alessandro. 2016. "Beams of Particles and Papers: How Digital Preprint Archives Shape Authorship and Credit." *Social Studies of Science* 46(4):629–45.

Delfanti, Alessandro and Nico Pitrelli. 2015. "Open Science: Revolution or Continuity." Pp. 59–68 in *Open science, open issues*, edited by S. Albagli, M. L. Maciel, and A. H. Abdo. Brasília: IBICT; Rio de Janeiro: Unirio.

Department of the Prime Minister and Cabinet. n.d. "Open Data." Retrieved June 8, 2020 (https://www.pmc.gov.au/public-data/open-data).

Desjardins-Proulx, Philippe, Ethan P. White, Joel J. Adamson, Karthik Ram, Timothee Poisot, and Dominique Gravel. 2013. "The Case for Open Preprints in Biology." *PLoS Biology* 11(5):1–5.

Dingwall, Robert. 1997. "Accounts, Interviews and Observations." Pp. 52–65 in *Context and Method in Qualitative Research*, edited by G. Miller and R. Dingwall. London: SAGE Publications.

Dixon, Paul. 2009. "'Hearts and Minds'? British Counter-Insurgency from Malaya to Iraq." *Journal of Strategic Studies* 32(3):353–81.

DOAJ. n.d. "DOAJ (Directory of Open Access Journals)." Retrieved June 1, 2020 (https://doaj.org/).

DORA. 2012. "San Francisco Declaration on Research Assessment." Retrieved April 20, 2020 (https://sfdora.org/read/).

Drexel University. 2014. "Mourning Jean-Claude Bradley, PhD, Department of Chemistry." Retrieved August 21, 2020 (https://drexel.edu/coas/academics/departments-centers/chemistry/news/2014/May/mourning-jean-claude-bradley-department-of-chemistry/).

Eamon, William. 1985. “From the Secrets of Nature to Public Knowledge: The Origins of the Concept of Openness in Science.” *Minerva* 23(3):321–47.

Easterbrook, Steve M. 2014. “Open Code for Open Science?” *Nature Geoscience* 7:779–81.

EC. 2015a. “Goals of Research and Innovation Policy.” Retrieved June 9, 2020 (https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy_en).

EC. 2015b. “‘Science 2.0’: Science in Transition.” Retrieved August 20, 2015 (http://scienceintransition.eu/).

EC. 2016. “Open Science Policy Platform | Open Science - Research & Innovation - European Commission.” Retrieved March 27, 2017 (http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-policy-platform).

EC. 2017. *Providing Researchers with the Skills and Competencies They Need to Practise Open Science: Open Science Skills Working Group Report*.

EC. 2018. *OSPP-REC: Open Science Policy Platform Recommendations*. Brussels.

EC. n.d. “Data Management.” Retrieved June 16, 2020a (https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/data-management_en.htm).

EC. n.d. “Data on Open Collaboration.” Retrieved June 9, 2020b (https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/data-open-collaboration_en).

Edge, David. 1995. “Reinventing the Wheel.” Pp. 3–24 in *Handbook of Science and Technology Studies, Revised Edition*. Thousand Oaks: SAGE Publications.

ESRC. n.d. “Research Data Policy.” Retrieved September 13, 2020 (https://esrc.ukri.org/funding/guidance-for-grant-holders/research-data-policy/).

Etzkowitz, Henry, Carol Kemelgor, and Brian Uzzi. 2000. *Athena Unbound: The Advancement of Women in Science and Technology*. Cambridge: Cambridge University Press.

Etzkowitz, Henry and Loet Leydesdorff. 2000. “The Dynamics of Innovation: From National Systems and ‘Mode 2’ to a Triple Helix of University–Industry–Government Relations.” *Research Policy* 29(2):109–23.

European Science Foundation. n.d. “Guidance on the Implementation of Plan S.” Retrieved October 28, 2020a (https://www.coalition-s.org/guidance-on-the-implementation-of-plan-s/).

European Science Foundation. n.d. “Plan S: Making Full and Immediate Open Access a Reality." Retrieved June 2, 2020b (https://www.coalition-s.org/).

Eve, Martin Paul. 2014. *Open Access and the Humanities: Contexts, Controversies and the Future*. Cambridge: Cambridge University Press.

Van den Eynden, Veerle, Gareth Knight, Anca Vlad, Barry Randler, Carol Tenopir, David Leon, Frank Manista, Jimmy Whitworth, and Louise Corti. 2016. *Towards Open Reserach: Practices, Experiences, Barriers and Opportunities*.

Fanelli, Daniele. 2018. "Is Science Really Facing a Reproducibility Crisis, and Do We Need It To?" *Proceedings of the National Academy of Sciences of the United States of America* 115(11):2628–31.

Fecher, Benedikt and Sascha Friesike. 2014. "Open Science: One Term, Five Schools of Thought." Pp. 17–47 in *Opening Science*, edited by S. Bartling and S. Friesike. Springer International Publishing.

Federer, Lisa M., Christopher W. Belter, Douglas J. Joubert, Alicia Livinski, Ya Ling Lu, Lissa N. Snyders, and Holly Thompson. 2018. "Data Sharing in PLOS ONE: An Analysis of Data Availability Statements." *PLoS ONE* 13(5):1–12.

Felt, Ulrike. 2009. *Knowing and Living in Academic Research: Convergence and Heterogeneity in Research Cultures in the European Context*. Prague: Institute of Sociology of the Academy of Sciences of the Czech Republic.

Felt, Ulrike. 2017. "Under the Shadow of Time: Where Indicators and Academic Values Meet." *Engaging Science, Technology, and Society* 3:53–63.

Finch, Janet. 2012. "Accessibility, Sustainability, Excellence: How to Expand Access to Research Publications." Retrieved July 16, 2016 (https://www.sconul.ac.uk/sites/default/files/documents/finch-report-executive-summary.pdf).

Ford, Emily. 2013. "Defining and Characterizing Open Peer Review: A Review of the Literature." *Journal of Scholarly Publishing* 44:311–26.

FOSTER. n.d. "Open Science." Retrieved October 5, 2020 (https://www.fosteropenscience.eu/taxonomy/term/7).

Foucault, Michel. 1991. "Panopticism." Pp. 195–228, 316–17 in *Discipline and Punish: the birth of the prison, translated by A. Sheridan*. London: Penguin.

Free Software Foundation. n.d. "Overview of the GNU System." Retrieved July 23, 2016a (https://www.gnu.org/gnu/gnu-history.html).

Free Software Foundation. n.d. "What Is Free Software?" Retrieved May 21, 2020b (https://www.gnu.org/philosophy/free-sw.html).

Freese, Jeremy and David Peterson. 2018. "The Emergence of Statistical Objectivity: Changing Ideas of Epistemic Vice and Virtue in Science." *Sociological Theory* 36(3):289–313.

Fyfe, Aileen, Kelly Coate, Stephen Curry, Stuart Lawson, Noah Moxham, and Camilla Mørk Røstvik. 2017. *Untangling Academic Publishing: A History Of The Relationship Between Commercial Interests, Academic Prestige And The Circulation Of Research*. St Andrews.

G7. 2017. "Annex 4: G7 Expert Group on Open Science." Retrieved November 2, 2020 (http://www.g7italy.it/sites/default/files/documents/ANNEX 4_WG Open Science/index.pdf).

G8. 2013. "G8 Open Data Charter." Retrieved November 2, 2020 (https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex).

Gabrielsen, Ane Møller. 2020. "Openness and Trust in Data-Intensive Science: The Case of Biocuration." *Medicine, Health Care and Philosophy* 23:497–504.

Garfield, Eugene. 1970. "Citation Indexing for Studying Science." *Nature* 227:669–71.

Garfield, Eugene. 2006. "The History and Meaning of the Journal Impact Factor." *The Journal of the American Medical Association* 295(1):90–93.

Gaudet, Joanne. 2014. "All That Glitters Is Not Gold: The Shaping of Contemporary Journal Peer Review at Scientific and Medical Journals." *UO Research* 1–23.

Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Trow. 1994. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*. London: Sage.

Gieryn, Thomas 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):781–95.

Ginsparg, Paul. 2011. "It Was Twenty Years Ago Today… [Preprint]." Retrieved November 2, 2020 (https://arxiv.org/abs/1108.2700).

Github. n.d. "Openingscience / Book." Retrieved June 17, 2020 (https://github.com/openingscience/book/commits/master/_posts).

Glaser, B. G. and A. L. Strauss. 1967. *The Discovery of Grounded Theory*. Chicago, IL: Aldine.

Gowers, Timothy. 2012. "Elsevier — My Part in Its Downfall." *Gowers's Weblog*. Retrieved October 4, 2020 (https://gowers.wordpress.com/2012/01/21/elsevier-my-part-in-its-downfall/).

Grand, Anne, Clare Wilkinson, Karen Bultitude, and Alan Winfield. 2012. "Open Science: A New 'Trust Technology'?" *Science Communication* 34(5):679–89.

Greenwald, Glenn. 2013. "The Inspiring Heroism of Aaron Swartz." *The Guardian*. Retrieved October 4, 2020 (https://www.theguardian.com/commentisfree/2013/jan/12/aaron-swartz-heroism-suicide1).

Guédon, Jean-Claude. 2001. "In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing." *138th Membership Meeting of the Association of Research Libraries*. Retrieved November 2, 2020 (http://www.arl.org/storage/documents/publications/in-oldenburgs-long-shadow.pdf).

Hagerlid, Jan. 2012. "Open Access - Looking Ten Years Back and a Few Years Ahead." *ScieCom Info* 8(1).

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14(3):575–99.

Harnad, Stevan. 2012. "Why the UK Should Not Heed the Finch Report." Retrieved June 13, 2020 (https://blogs.lse.ac.uk/impactofsocialsciences/2012/07/04/why-the-uk-should-not-heed-the-finch-report/).

Heberle, H., G. V. Meirelles, F. R. da Silva, G. P. Telles, and R. Minghim. 2015. "InteractiVenn: A Web-Based Tool for the Analysis of Sets through Venn Diagrams." *BMC Bioinformatics* 16:169.

HEFCE. 2014. "Open Access in the Post-2014 Research Excellence Framework." Retrieved July 25, 2016 (http://www.hefce.ac.uk/media/hefce/content/pubs/2014/CL,072014/Print-friendly version.pdf).

HEFCE, RCUK, Universities UK, and Wellcome Trust. 2016. "Concordat On Open Research Data." Retrieved November 2, 2020 (https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf).

Hessels, Laurens K. and Harro van Lente. 2008. "Re-Thinking New Knowledge Production: A Literature Review and a Research Agenda." *Research Policy* 37(4):740–60.

HGP Information Archive. 1997. "Summary of the Report of the Second International Strategy Meeting on Human Genome Sequencing (Bermuda, 27th February - 2nd March, 1997) as Reported by HUGO." Retrieved June 5, 2020 (https://web.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml#1).

Hicks, Diana and Paul Wouters. 2015. "The Leiden Manifesto for Research Metrics." *Nature* 520:429–31.

Higher Education Statistics Agency. n.d. "Who's Working in HE?: Personal Characteristics." Retrieved August 31, 2020 (https://www.hesa.ac.uk/data-and-analysis/staff/working-in-he/characteristics).

Higman, Rosie, Daniel Bangert, and Sarah Jones. 2019. "Three Camps, One Destination: The Intersections of Research Data Management, FAIR and Open." *Insights: The UKSG Journal* 32(18):1–9.

Himmelstein, Daniel S., Ariel Rodriguez Romero, Jacob G. Levernier, Thomas Anthony Munro, Stephen Reid McLaughlin, Bastian Greshake Tzovaras, and Casey S. Greene. 2018. "Sci-Hub Provides Access to Nearly All Scholarly Literature." *ELife* 7:e32822.

HM Government Cabinet Office. 2012. "Open Data White Paper: Unleashing the Potential." 51. Retrieved November 2, 2020 (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/78946/CM8353_acc.pdf).

Holden, Kerry. 2014. "Lamenting the Golden Age: Love , Labour and Loss in the Collective Memory of Scientists." *Science as Culture* 1–22.

Iliffe, Rob. 1992. "'In the Warehouse': Privacy, Property and Priority in the Early Royal

Society." *History of Science* 30(1):29–68.

Ioannidis, John P. A. 2005. "Why Most Published Research Findings Are False." *PLoS Medi* 2(8):e124.

Irwin, Alan and Brian Wynne. 1996. "Conclusions." Pp. 213–21 in *Misunderstanding Science? The Public Reconstruction of Science and Technology*, edited by A. Irwin and B. Wynne. Cambridge: Cambridge University Press.

Jeffers, Susan. 1987. *Feel The Fear And Do It Anyway: How to Turn Your Fear and Indecision into Confidence and Action*. London: Century Hutchinson.

Jisc. n.d. "Sherpa Services." Retrieved October 28, 2020 (https://www.jisc.ac.uk/sherpa).

Kekecs, Zoltan, Balazs Aczel, Bence Palfi, Barnabas Szaszi, Peter Szecsi, Mark Zrubka, Marton Kovacs, Bence Bakos, Denis Cousineau, Patrizio Tressoldi, Massimo Grassi, Dana Arnold, Thomas Evans, Yuki Yamada, Jeremy Miller, Huanxu Liu, Fumiya Yonemitsu, and Dmitrii Dubrov. 2019. "Raising the Value of Research Studies in Psychological Science by Increasing the Credibility of Research Reports: The Transparent Psi Project [Preprint]." *PsyArXiv*. Retrieved November 2, 2020 (10.31234/osf.io/uwk7y).

Kelty, Christopher M. 2008. *Two Bits: The Cultural Significance of Free Software*. Vol. 17. USA: Duke University Press.

Kelty, Christopher M. 2012. "This Is Not an Article: Model Organism Newsletters and the Question of 'Open Science.'" *BioSocieties* 7(2):140–68.

Kerr, Anne and Lisa Garforth. 2016. "Affective Practices, Care and Bioscience: A Study of Two Laboratories." *Sociological Review* 64:3–20.

Kling, Rob. 2001. "The Internet and Unrefereed Scholarly Publishing." *Annual Reivew of the Information Science and Technology* 38(1):591–631.

Kling, Rob and Geoffrey McKim. 2000. "Not Just a Matter of Time: Field Differences and the Shaping of Electronic Media in Supporting Scientific Communication." *Journal of the American Society for Information Science and Technology* 51(14):1306–20.

Knorr Cetina, Karin. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. London: Harvard University Press.

Kohler, Robert E. 1999. "Moral Economy, Material Culture, and Community in Drosophila Genetics." Pp. 243–57 in *The Science Studies Reader*, edited by M. Biagioli. New York and London: Routledge.

Kozak, Marcin and James Hartley. 2013. "Publication Fees for Open Access Journals: Different Disciplines—Different Methods." *Journal of the American Society for Information Science and Technology* 64(12):2591–94.

Kramer, Bianca and Jeroen Bosman. 2018. "Rainbow of Open Science Practices." *Zenodo*. Retrieved November 2, 2020 (https://doi.org/10.5281/zenodo.1147025).

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. UK: Oxford University Press.

Lawson, Stuart Andrew. 2019. "Open Access Policy in the UK: From Neoliberalism to the Commons." Birkbeck, University of London.

Lawson, Stuart, Kevin Sanders, and Lauren Smith. 2015. "Commodification of the Information Profession: A Critique of Higher Education Under Neoliberalism." *Journal of Librarianship and Scholarly Communication* 3(1):eP1182.

Leonelli, Sabina. 2007. "Growing Weed, Producing Knowledge An Epistemic History of Arabidopsis Thaliana." *History and Philosophy of the Life Sciences* 29(2):193–223.

Leonelli, Sabina. 2013. "Why the Current Insistence on Open Access to Scientific Data? Big Data, Knowledge Production, and the Political Economy of Contemporary Biology." *Bulletin of Science, Technology & Society Society* 33(1–2):6–11.

Leonelli, Sabina. 2016. *Data-Centric Biology*. Chicago: Chicago University Press.

Leonelli, Sabina. 2018. "Rethinking Reproducibility as a Criterion for Research Quality." Pp. 129–46 in *Including a Symposium on Mary Morgan: Curiosity, Imagination, and Surprise (Research in the History of Economic Thought and Methodology, Vol. 36B)*. Emerald Publishing Limited.

Leonelli, Sabina and Rachel A. Ankeny. 2012. "Re-Thinking Organisms: The Impact of Databases on Model Organism Biology." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 43(1):29–36.

Leonelli, Sabina, Daniel Spichtinger, and Barbara Prainsack. 2015. "Sticks and Carrots: Encouraging Open Science at Its Source." *Geography and Environment* 2:12–16.

LERU. 2018. "Open Science and Its Role in Universities: A Roadmap for Cultural Change." Retrieved November 2, 2020 (https://www.leru.org/publications/open-science-and-its-role-in-universities-a-roadmap-for-cultural-change).

Levin, N. and S. Leonelli. 2016. "How Does One 'Open' Science? Questions of Value in Biological Research." *Science, Technology & Human Values* 1–26.

Levin, N., S. Leonelli, D. Weckowska, D. Castle, and J. Dupre. 2016. "How Do Scientists Define Openness? Exploring the Relationship Between Open Science Policies and Research Practice." *Bulletin of Science, Technology & Society* 1–13.

Long, Pamela O. 2001. *Openness, Secrecy, Authorship : Technical Arts and the Culture of Knowledge from Antiquity to the Renaissance*. Baltimore: Johns Hopkins University Press.

Malterud, Kirsti. 2001. "Qualitative Research: Standards, Challenges, and Guidelines." *Lancet* 358(9280):483–88.

Mann, Steve. 2016. "Surveillance (Oversight), Sousveillance (Undersight), and Metaveillance (Seeing Sight Itself)." Pp. 1408–17 in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. Las Vegas.

Mann, Steve, Ryan Janzen, Valmiki Rampersad, Jason Huang, and Lei Jimmy Ba. 2016. "'sQUEAKeys': A Friction Idiophone, for Physical Interaction with Mobile Devices." in *2015 IEEE Games Entertainment Media Conference*. Toronto, Canada.

Marginson, Simon. 1996. "Power and Modernisation in Higher Education: Australia after the Reforms of 1987–1989." *Critical Studies in Education* 37(2):77–99.

Maxson Jones, Kathryn, Rachel A. Ankeny, and Robert Cook-Deegan. 2018. "The Bermuda Triangle: The Pragmatics, Policies, and Principles for Data Sharing in the History of the Human Genome Project." *Journal of the History of Biology* 51(4):693–805.

McMullin, Ernan. 1979. "Review: Laudan's Progress and Its Problems." *Philosophy of Science* 46(4):623–44.

McMullin, Ernan. 1985. "Openness and Secrecy in Science: Some Notes on Early History." *Science, Technology & Human Values* 10(2):14–22.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago and London: University of Chicago Press.

Mirowski, Philip. 2018. "The Future(s) of Open Science." *Social Studies of Science* 48(2):171–203.

Mitroff, Ian 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):579–95.

Mittermaier, Bernhard. 2015. "Double Dipping in Hybrid Open Access – Chimera or Reality?" *ScienceOpen Research*. Retrieved November 2, 2020 (10.14293/S2199-1006.1.SOR-SOCSCI.AOWNTU.v1).

Mittra, James. 2016. *The New Health Bioeconomy: R&D Policy and Innovation for the Twenty-First Century*. Palgrave Macmillan.

Moore, Samuel A. 2017. "A Genealogy of Open Access: Negotiations between Openness and Access to Research." *Revue Française Des Sciences de l'information et de La Communication* (11):1–16.

Moore, Samuel A. 2018. "The Commons and Care." Pp. 16–25 in *The Commons and Care*, edited by Samuel Moore and Mattering Press. Coventry, UK: Post Office Press and Rope Press.

Moore, Samuel A. 2019a. "Common Struggles: Policy-Based vs. Scholar-Led Approaches to Open Access in the Humanities." King's College London.

Moore, Samuel A. 2019b. "Revisiting 'the 1990s Debutante': Scholar-Led Publishing and the Prehistory of the Open Access Movement." *Journal of the Association for Information Science and Technology* 1–11.

MRC. 2007. "Data Access." Retrieved June 8, 2020 (https://web.archive.org/web/20070612032911/http://www.mrc.ac.uk/PolicyGuidance/EthicsAndGovernance/DataAccess/index.htm).

Mulkay, M. 1976. "Norms and Ideology in Science." *Social Science Information* 15:637–56.

Müller, Ruth and Sarah De Rijcke. 2017. "Thinking with Indicators. Exploring the Epistemic Impacts of Academic Performance Indicators in the Life Sciences." *Research Evaluation* 26(3):157–68.

Murray-Rust, Peter. 2008. "Open Data in Science." *Serials Review* 34(1):52–64.

Murray-Rust, Peter, Rufus Pollock, Bryan Bishop, Mike Chelen, Cameron Neylon, Daniel Mietchen, and Andy Powell. 2010. "Panton Principles: Principles for Open Data in Science." Retrieved June 5, 2020 (https://pantonprinciples.org/faq/index.html).

Nabyonga-Orem, Juliet, James Avoka Asamani, Thomas Nyirenda, and Seye Abimbola. 2020. "Article Processing Charges Are Stalling the Progress of African Researchers: A Call for Urgent Reforms." *BMJ Global Health* 5(9):1–4.

NCBI. n.d. "PMC FAQs: What Is PubMed Central (PMC)?" Retrieved June 2, 2020 (https://www.ncbi.nlm.nih.gov/pmc/about/faq/#q1).

Newton, Isaac. 1675. "Isaac Newton Letter to Robert Hooke, 1675." *Historical Society of Pennsylvania*. Retrieved November 1, 2020 (https://digitallibrary.hsp.org/index.php/Detail/objects/9792).

Neylon, Cameron. 2017. "Compliance Culture or Culture Change? The Role of Funders in Improving Data Management and Sharing Practice amongst Researchers." *Research Ideas and Outcomes* 3:e21705.

NHMRC. 2018. "Open Access Policy." Retrieved November 2, 2020 (https://www.nhmrc.gov.au/about-us/resources/open-access-policy).

NHMRC, ARC, and Universities Australia. 2007. "Australian Code for the Responsible Conduct of Research." *Book Australian Code for the Responsible Conduct Of*. Retrieved November 2, 2020 (https://web.archive.org/web/20170119145725/https://www.nhmrc.gov.au/_files_nhmrc/file/research/research-integrity/r39_australian_code_responsible_conduct_research_150811.pdf).

Nielsen, Michael. 2009. "Doing Science in the Open." *Physics World* 22(5):30–35.

Nielsen, Michael. 2011. *Reinventing Discovery: The New Era of Networked Science*. United States of America: Princeton University Press.

Niles, Meredith T., Lesley A. Schimanski, Erin C. McKiernan, and Juan P. Alperin. 2020. "Why We Publish Where We Do: Faculty Publishing Values and Their Relationship to Review, Promotion and Tenure Expectations." *PLoS ONE* 15(3):e0228914.

Van Noorden, Richard. 2020. "Open-Access Plan S to Allow Publishing in Any Journal." *Nature News*. Retrieved October 4, 2020 (https://www.nature.com/articles/d41586-020-02134-6).

Nosek, Brian A. 2014. "Open Science." Retrieved October 18, 2020 (https://www.nih.gov/news-events/videos/open-science).

Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of*

Sciences of the United States of America 115(11):2600–2606.

November, Joseph A. 2012. *Biomedical Computing: Digitizing Life in the United States*. Baltimore: Johns Hopkins University Press.

OCSDNet. n.d. "Open Science Manifesto." Retrieved June 19, 2020 (https://ocsdnet.org/manifesto/open-science-manifesto/).

ODI. 2013. "What Makes Data Open?" Retrieved June 15, 2020 (https://theodi.org/article/what-makes-data-open/).

ODI. n.d. "About the ODI." Retrieved June 8, 2020a (https://theodi.org/about-the-odi/).

ODI. n.d. "ODI Queensland." Retrieved June 8, 2020b (https://theodi.org/organisation/odi-queensland/).

OECD. 2004. "Declaration on Access to Research Data from Public Funding." Retrieved November 2, 2020 (https://legalinstruments.oecd.org/public/doc/157/157.en.pdf).

OECD. 2014. "OECD Science, Technology and Industry 2014." Retrieved November 2, 2020 (http://dx.doi.org/10.1787/sti_outlook-2014-en).

OKF. 2006. "Open Knowledge Definition 1.0." Retrieved June 10, 2020 (https://opendefinition.org/od/1.0/en/).

OKF. 2015. "Open Definition 2.1." Retrieved June 5, 2020 (https://opendefinition.org/od/2.1/en/).

OKF. n.d. "History of the Open Definition." Retrieved June 5, 2020a (https://opendefinition.org/history/).

OKF. n.d. "The Open-Science Archives." Retrieved June 9, 2020b (https://lists-archive.okfn.org/pipermail/open-science/).

Open Science Collaboration. 2015. "Estimating the Reproducibility of Psychological Science." *Science* 349(6251):aac4716.

Pandora, Katherine and Karen A. Rader. 2008. "Science in the Everyday World: Why Perspectives from the History of Science." *Isis* 99(2):350–64.

Parker, John N. and Edward J. Hackett. 2012. "Hot Spots and Hot Moments in Scientific Collaborations and Social Movements." *American Sociological Review* 77(1):21–44.

Pinfield, Stephen, Andrew M. Cox, and Jen Smith. 2014. "Research Data Management and Libraries: Relationships, Activities, Drivers and Influences." *PLoS ONE* 9(12):e114734.

Pinfield, Stephen, Jennifer Salter, and Peter A. Bath. 2015. "The 'Total Cost of Publication' in a Hybrid Open-Access Environment: Institutional Approaches to Funding Journal Article-Processing Charges in Combination With Subscriptions." *Journal of the Association for Information Science and Technology* 67(7):1751–66.

Piwowar, Heather a. 2011. "Who Shares? Who Doesn't? Factors Associated with Openly Archiving Raw Research Data." *PloS One* 6(7):e18657.

Piwowar, Heather, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. 2018. "The State of OA: A Large-Scale Analysis of the Prevalence and Impact of Open Access Articles." *PeerJ* 6:e4375.

PLOS. 2000. "Open Letter." Retrieved June 2, 2020 (https://plos.org/open-letter/).

PLOS. 2017. "Who We Are." Retrieved October 5, 2020 (https://web.archive.org/web/20170829215357if_/https://www.plos.org/who-we-are).

PLOS. n.d. "Early History." Retrieved June 2, 2020a (https://web.archive.org/web/20120209174104/http://www.plos.org/about/what-is-plos/early-history/).

PLOS. n.d. "Publication Fees." Retrieved June 1, 2020b (https://plos.org/publish/fees/).

Polanyi, Michael. 1962. "The Republic of Science: Its Political and Economic Theory." *Minerva* 1(1):54–73.

Pollock, Rufus. 2014. "Open Definition v2.0 Released – Major Update of Essential Standard for Open Data and Open Content." Retrieved June 15, 2020 (https://blog.okfn.org/2014/10/07/open-definition-v2-0-released-major-update-of-essential-standard-for-open-data-and-open-content/).

Popper, Karl R. 1945. *The Open Society and Its Enemies*. London: Routledge.

Poynder, Richard. 2006. "Interview with Vitek Tracz." *Open and Shut?* Retrieved July 29, 2019 (https://poynder.blogspot.com/2006/05/interview-with-vitek-tracz.htm).

Prainsack, Barbara and Sabina Leonelli. 2018. "Openness and Responsibility." Pp. 97–106 in *Science and the politics of openness: here be monsters*, edited by B. Nerlich, S. Hartley, S. Raman, and A. Smith. Manchester: Manchester University Press.

Price, Derek J. de Solla. 1963. *Little Science, Big Science*. New York: Columbia University Press.

Procter, R., R. Williams, J. Stewart, M. Poschen, H. Snee, A. Voss, and M. Asgari-Targhi. 2010. "Adoption and Use of Web 2.0 in Scholarly Communications." *Philosophical Transactions of the Royal Society A* 368(1926):4039–56.

Proctor, Rob, Robin Williams, and James Stewart. 2010. "If You Build It, Will They Come? How Researchers Perceive and Use Web 2.0." Retrieved June 7, 2016 (www.rin.ac.uk/web-20-researchers).

Puig de la Bellacasa, María. 2011. "Matters of Care in Technoscience: Assembling Neglected Things." *Social Studies of Science* 41(1):85–106.

Quintana, Dan and James Heathers. 2020. "Everything Hertz: A Podcast by Scientists, for Scientists. Methodology, Scientific Life, and Bad Language." Retrieved October 27, 2020 (https://everythinghertz.com/).

QUT. n.d. “Open Scholarship and Open Access.” Retrieved June 2, 2020 (https://www.library.qut.edu.au/about/whatwedo/support-researchers/open/).

Rader, Karen A. 1998. “‘The Mouse People’: Murine Genetics Work at the Bussey Institution, 1909-1936.” *Journal of the History of Biology* 31(3):327–54.

RCUK. 2011. “RCUK Common Principles on Data Policy.” Retrieved June 8, 2020 (https://web.archive.org/web/20111129024454/http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx).

RCUK. 2013. “RCUK Policy on Open Access and Supporting Guidance.” (i):1–13. Retrieved July 27, 2016 (http://www.rcuk.ac.uk/documents/documents/rcukopenaccesspolicy-pdf/).

RCUK. n.d. “RCUK Policy on Open Access.” Retrieved June 2, 2020 (https://web.archive.org/web/20141221062054/http://www.rcuk.ac.uk/research/openaccess/policy/).

RDA. 2016. “About RDA.” Retrieved June 8, 2020 (https://www.rd-alliance.org/about-rda).

Robinson-Garcia, Nicolas, Rodrigo Costas, and Thed van Leeuwen. 2020. “Open Access Uptake by Universities Worldwide.” *PeerJ* 8:e9410.

Ross, John. 2019. “Repository Requirements a Deal-Breaker for Plan S Down Under.” *Times Higher Education*, February 22.

Rothman, Robert A. 1972. “A Dissenting View on the Scientific Ethos.” *The British Journal of Sociology* 23(1):102–8.

Rowhani-Farid, A. 2018. “Towards a Culture of Open Science and Data Sharing in Health and Medical Research.” Queensland University of Technology.

Savransky, Martin. 2014. “Of Recalcitrant Subjects.” *Culture, Theory and Critique* 55(1):96–113.

Scheliga, Kaja and Sascha Friesike. 2014. “Putting Open Science into Practice: A Social Dilemma?” *First Monday* 19(9).

Schönbauer, Sarah. 2019. “At Least a Little Bit of Stability. How Scientists Do Identity Work in Contemporary Life Sciences.” University of Vienna.

Schönbauer, Sarah Maria. 2020. “‘From Bench to Stage’ How Life Scientists’ Leisure Groups Build Collective Self-Care.” *Science as Culture* 1–22.

Shapin, S. and S. Schaffer. 1985. *Leviathan and the Air-Pump: Hobbes, Boyle and the Experimental Life*. Princeton, N.J.: Princeton University Press.

Shapin, Steven. 1984. “Pump and Circumstance : Robert Boyle’s Literary Technology.” *Social Studies of Science* 14(4):481–520.

Shapin, Steven. 1988. “The House of Experiment in Seventeenth-Century England.” *Isis* 79(3):373–404.

Shapin, Steven. 1994. *A Social History of Truth*. Chicago and London: The University of Chicago Press.

Shapin, Steven. 1996. *The Scientific Revolution*. United States of America: The University of Chicago.

Shaw, Allison K. and Daniel E. Stanton. 2012. "Leaks in the Pipeline: Separating Demographic Inertia from Ongoing Gender Differences in Academia." *Proceedings of the Royal Society B: Biological Sciences* 279:3736–41.

Shearer, Kathleen. 2015. *Comprehensive Brief on Research Data Management Policies*.

Shotton, David. 2018. "Funders Should Mandate Open Citations." *Nature* 553:129.

Silva, Liz. 2014. "PLOS' New Data Policy: Public Access to Data." *PLOS Blogs: EveryONE*. Retrieved June 8, 2020 (https://everyone.plos.org/2014/02/24/plos-new-data-policy-public-access-data-2/).

Šimukovič, Elena. 2020. "Plan S, Open Access and the Potential Roles for STS Research." *Nordic Journal of Science and Technology Studies* 8(1):27–30.

Skinner, Quentin. 1969. "Thomas Hobbes and the Nature of the Early Royal Society." *The Historical Journal* 12(2):217–39.

Slaughter, Sheila A. and Larry L. Leslie. 1997. *Academic Capitalism: Politics, Policies and the Entrepreneurial University*. Baltimore: Johns Hopkins University Press.

Song, Miri. 2020. "Rethinking Minority Status and 'Visibility.'" *Comparative Migration Studies* 8(5):1–17.

SSAC. 2019. "Open Research: How Data Sharing Can Advance Scientific Impact in Scotland." Retrieved November 2, 2020 (https://www.scottishscience.org.uk/sites/default/files/article-attachments/SSAC Briefing Note - Open Research.pdf).

Stallman, Richard. 2016. "Why Open Source Misses the Point of Free Software." Retrieved July 23, 2016 (https://www.gnu.org/philosophy/open-source-misses-the-point.html).

Stam, Henderikus J. 2018. "Once More with Feeling: The Eternal Recurrence of the Reproducibility Crisis in Psychology." in *ESHHS GRONINGEN: European Society for the History of the Human Sciences*. Groningen: Open University of the Netherlands.

Star, Susan Leigh and James R. Griesemer. 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.

Steele, Colin. 2013. "Open Access in Australia: An Odyssey of Sorts?" *Insights: The UKSG Journal* 26(3):282–89.

Strasser, Bruno J. 2019. "Open Science." Pp. 228–404 in *Collecting Experiments : Making Big Data Biology*. University of Chicago Press.

Suber, Peter. 2006. "Open Access Mandates Coming to the RCUK." *SPARC Open Access Newsletter*. Retrieved June 2, 2020 (https://legacy.earlham.edu/~peters/fos/newsletter/07-02-06.htm#rcuk).

Suber, Peter. 2009. "Timeline of the Open Access Movement." Retrieved June 8, 2020 (https://legacy.earlham.edu/~peters/fos/timeline.htm).

Suber, Peter. 2012. *Open Access (The MIT Press Essential Knowledge Series)*. Cambridge, Massachusetts: MIT Press.

Talkhani, Zeba. 2019. *My Past Is a Foreign Country: A Muslim Feminist Finds Herself*. London: Hodder & Stoughton.

Taubes, Gary. 1993. "Publication by Electronic Mail Takes Physics by Storm." *Science* 259:1246–48.

Tennant, Jonathan. 2018. "Open Science Is Just Good Science." Retrieved August 28, 2020 (https://figshare.com/articles/Open_Science_is_just_good_science_pptx/5783004/1).

Tennant, Jonathan, Jennifer Elizabeth Beamer, Jeroen Bosman, Björn Brembs, Neo Christopher Chung, Gail Clement, Tom Crick, Jonathan Dugan, Alastair Dunning, David Eccles, Asura Enkhbayar, Daniel Graziotin, Rachel Harding, Johanna Havemann, Daniel S. Katz, Kshitiz Khanal, Jesper Norgaard Kjaer, Tim Koder, Paul Macklin, Christopher R. Madan, Paola Masuzzo, Lisa Matthias, Katja Mayer, David Nichols, Elli Papadopoulou, Thomas Pasquier, Tony Ross-Hellauer, Michael Schulte-Mecklenbeck, Dan Sholler, Tobias Steiner, Pawel Szczesny, and Andy Turner. 2019. "Foundations for Open Scholarship Strategy Development [Preprint]." *MetaArXiv Preprints*. Retrieved November 2, 2020 (https://doi.org/10.31222/osf.io/b4v8p).

Teytelman, Leonid, Alexei Stoliartchouk, Lori Kindler, and Bonnie L. Hurwitz. 2016. "Protocols.Io: Virtual Communities for Protocol Development and Discussion." *PLoS Biology* 14(8):e1002538.

Thackray, Arnold. 1998. *Private Science: Biotechnology and the Rise of the Molecular Sciences*. Philadelphia: University of Pennsylvania Press.

The PLoS Medicine Editors. 2006. "The Impact Factor Game." *PLOS Medicine* 3(6):0707–8.

The Royal Society. 2012. *Science as an Open Enterprise*. London.

Thomas, Gary and David James. 2006. "Reinventing Grounded Theory: Some Questions about Theory, Ground and Discovery." *British Educational Research Journal* 32(6):767–95.

Till, James E. 2001. "Predecessors of Preprint Servers." *Learned Publishing* 14:7–13.

Tkacz, Nathaniel. 2012. "From Open Source to Open Government: A Critique of Open Politics." *Ephemera: Theory & Politics in Organization* 12(4):386–405.

Trust, Wellcome. 2020. "Wellcome and COAF Open Access Spend 2018/19." Retrieved June 2, 2020 (https://wellcome.ac.uk/funding/wellcome-and-coaf-open-access-spend-201819).

UK Government. 2010. "Government Removes Barrier to Public Data Use." Retrieved June 8, 2020 (https://www.gov.uk/government/news/government-removes-barrier-to-public-data-use).

UK Government. 2012. "'Accessibility, Sustainability, Excellence: How to Expand Access to Research Publications' Finch Group Report: Government Response on Open Access." Retrieved June 2, 2020 (https://www.gov.uk/government/publications/letter-to-dame-janet-finch-on-the-government-response-to-the-finch-group-report-accessibility-sustainability-excellence-how-to-expand-access-to-research-publications).

Vanderstraeten, Raf and Joshua Eykens. 2018. "Communalism and Internationalism: Publication Norms and Structures in International Social Science." *Serendipities* 3(1):14–28.

Vermeir, Koen. 2012. "Openness versus Secrecy? Historical and Historiographical Remarks." *British Journal for the History of Science* 45(2):165–88.

Vermeir, Koen and Dániel Margócsy. 2012. "States of Secrecy: An Introduction." *British Journal for the History of Science* 45(2):153–64.

Vermeulen, Niki. 2016. "Big Biology: Supersizing Science During the Emergence of the 21st Century." *NTM Zeitschrift Für Geschichte Der Wissenschaften, Technik Und Medizin* 24(2):195–223.

Vermeulen, Niki, John N. Parker, and Bart Penders. 2013. "Understanding Life Together: A Brief History of Collaboration in Biology." *Endeavour* 37(3):162–71.

Vicente-Saez, Ruben and Clara Martinez-Fuentes. 2018. "Open Science Now: A Systematic Literature Review for an Integrated Definition." *Journal of Business Research* 88:428–36.

Wadman, Meredith. 2012. "A Workforce out of Balance." *Nature* 486:304.

Wallace, Nicholas. 2019. "Open Science and Research Data – a View from down Under." *Science | Business*, June 18.

Wallis, Jillian C., Elizabeth Rolando, and Christine L. Borgman. 2013. "If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology." *PloS One* 8(7):e67332.

Walsh, John P. and Todd Bayma. 1996. "Computer Networks and Scientific Work." *Social Studies of Science* 26(3):661–703.

Watson, Mick. 2015. "When Will 'Open Science' Become Simply 'Science'?" *Genome Biology* 16:101.

Weckowska, Dagmara M., Nadine Levin, Sabina Leonelli, John Dupré, and David Castle. 2017. "Managing the Transition to Open Access Publishing: A Psychological Perspective." *Prometheus* 35(2):111–35.

Weinberg, Alvin M. 1961. "Impact of Large-Scale Science on the United States." *Science* 134(3473):161–64.

Weinberg, Alvin M. 1967. *Reflections on Big Science*. Cambridge: MIT Press.

Wellcome Trust. 2003. "Sharing Data from Large-Scale Biological Research Projects: A System of Tripartite Responsibility." Retrieved November 2, 2020 (http://www.genome.gov/Pages/Research/WellcomeReport0303.pdf).

Wellcome Trust. 2004. “Costs and Business Models in Scientific Research Publishing.” Retrieved November 2, 2020 (https://wellcome.ac.uk/sites/default/files/wtd003184_0.pdf).

Wellcome Trust. 2007. “Policy on Data Management and Sharing.” Retrieved June 8, 2020 (https://web.archive.org/web/20090726101426/http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTX035043.htm).

Wellcome Trust. 2008. “Authors’ Guide and FAQ: Open and Unrestricted Access to Published Research.” Retrieved June 2, 2020 (https://web.archive.org/web/20081013110247/http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD018855.htm).

Wellcome Trust. 2012. “Wellcome Trust Strengthens Its Open Access Policy.” Retrieved January 29, 2016 (http://www.wellcome.ac.uk/News/Media-office/Press-releases/2012/WTVM055745.htm).

Wellcome Trust. n.d. “Open Access Policy – before 1 January 2021.” Retrieved October 25, 2020a (https://wellcome.org/grant-funding/guidance/open-access-guidance/open-access-policy).

Wellcome Trust. n.d. “Open Research.” Retrieved June 9, 2020b (https://wellcome.ac.uk/what-we-do/our-work/open-research).

Wikipedia contributors. n.d. “Open Science.” *Wikipedia*.

Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J. G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A. C. t Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan Van Der Lei, Erik Van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao, and Barend Mons. 2016. “Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data* 3:1–9.

Willinsky, John. 2005. “The Unacknowledged Convergence of Open Source, Open Access, and Open Science.” *First Monday* 10(8):1–18.

Wilsdon, James, Judit Bar-Ilan, Robert Frodeman, Elisabeth Lex, Isabella Peters, and Paul Wouters. 2017. “Next-Generation Metrics: Responsible Metrics and Evaluation of Open Science.” *European Commision. Directorate-General for Research and Innovation*. Retrieved November 2, 2020 (https://ec.europa.eu/research/openscience/pdf/report.pdf).

Wright, Susan. 1986. “Recombinant DNA Technology and Its Social Transformation, 1972-1982.” *Osiris* 2:303–60.

Wykle, Stacy S. 2014. “Enclaves of Anarchy: Preprint Sharing, 1940-1990.” *Proc. Am. Soc. Info. Sci. Tech.* 51:1–10.

Yearley, Steven. 2005. *Making Sense of Science*. London: Sage Publications.

Yu, Harlan and David G. Robinson. 2011. “The New Ambiguity of ‘Open Government.’” *UCLA Law Review Discourse* 59:178–209.

Zenodo. n.d. “About Zenodo.” Retrieved November 1, 2020 (https://about.zenodo.org/).

Ziman, John. 2000. *Real Science: What It Is and What It Means*. Cambridge: Cambridge University Press.

Zuckerman, Harriet and Robert K. Merton. 1971. “Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System.” *Minerva* 9(1):66–100.

지우 황. 2013. “Open Science 우산.” *Flickr.Com*. Retrieved November 2, 2020 (https://flic.kr/p/gxKYeW).

Appendix A: Recommendations for open science advocates and policymakers

I acknowledge that individuals and organisations advocating and making policy to promote open science are diverse, and face complex challenges with which I am not practically familiar. Individuals in these roles likely have contextual knowledge that aligns with or challenges my findings. I present these recommendations with the hope that they may become conversations, and enable mutual learning.

- Where scientists seem to lack awareness, understanding, engagement, or incentives to engage with open practices, consider alternative framings. For example, their position may arise from a combination of:
    - distrust of institutions that represent or mediate open practices, including commercial publishing systems;
    - vulnerability, especially if they are experiencing other forms of insecurity (or have been influenced by such experiences in the past);
    - implementations of “open” that lack sensitivity or relevance to particular local contexts of knowledge-making;
- Open practices that reveal *processes* of knowledge-making (e.g. open data) have a greater capacity to engage scientists deeply than practices that make finalised science open (e.g. open access).
    - But for similar reasons, opening up unpublished and/or process-related information can be vulnerable and risky experience, especially for scientists with existing vulnerabilities.
- Avoiding one-size-fits-all approaches is important. Also important: scientists’ relationships with openness vary not just according to face-value factors like discipline and ethical or commercial considerations, but layered and nuanced factors like the uniqueness of project resources; trust dynamics in their community; and long-term career security (see **Table 6**).
- It may be easier to engage scientists in open practices that allow trust-building and social, collaborative interaction between sharer and recipient.

- Top-down, enforceable forms of openness present a problem for engagement for interlinked reasons related to trust, agency, and sensitivity to local context.
    - There may be an inverse relationship between the measurability and enforceability of an open practice, and its value and meaning in context.
- Open practices tend to be easier for scientists who are systemically advantaged. Consider that reward systems incentivising openness may reinforce existing forms of advantage and disadvantage.
- Supportive supervision and mentoring is an important factor promoting a long-term sense of security in early career scientists, which in turn is a foundation for openness.
- Most importantly: consider nurturing a sense of wellbeing, security and trust in research communities as a first step to supporting openness, especially if openness requires vulnerability. This may require imaginative, empathy-based policymaking. I include a thought experiment below.

Thought experiment: empathy-based policymaking for openness

The kinds of openness that scientists and advocates of open science may value the most entail in-depth knowledge work, caring labour, and vulnerability. Such practices cannot be straightforwardly measured, incentivised, and implemented. They connect to questions like: what makes people secure pursuing this profession? What makes them care about it, deeply? Are they at liberty to pursue that? If they are generous with their community and with society, will they be valued or punished? If they are unable to be generous, will they be supported – not isolated, not punished? Making sense of research culture is hard. It's even harder to admit that we probably know some of the ways to improve it already, but they are not convenient, politically, financially, or procedurally. They are not amenable to standard policy measures.

Think about the best contexts for openness that you have witnessed. Not the most open individuals. Think about the student who feels safe, inspired, and ambitious because their department holds a coffee morning every day (see Section 7.10).

They don’t just hold it, people attend. People don’t just attend, they talk. They don’t just talk, they talk openly. They don’t just talk to people they know. They don’t just talk to people of their own rank. Senior scientists go out of their way. Senior scientists are at liberty to go out of their way to help. Because they don’t feel precarious themselves. And because they didn’t feel precarious when they were students. Think about why these things might come about. How they are nourished. How you can empower and entrust scientists to create these kinds of communities.

Appendix B: Interview schedule

Sample interview schedule for scientists

Thank you.
Introduce - sci training in bio, work in sci publishing, interest in social/cultural aspects.
Topic - deliberately broad, you can frame the discussion, no right or wrong answers.
Biographical component at the beginning for context.
Time to read/explain consent forms, emphasise main points (anonymity, withdrawal)
Any Qs?

--

Could you tell me briefly about the research you're currently involved in?

How would you define the discipline/disciplines that you are part of?
Your primary identity as a scientist? What is comfortable?
When in your career did you start thinking of yourself as this?
Previous identities in career?

Which scientific community or communities would you say you were part of?

How would you describe those communities? (character words)

--

Could you tell me what first comes to mind when you think about openness in science? Most salient issues/thoughts

Does [that kind of] openness feel like something relevant to you?
Does it make you feel any particular way - interested, passionate, bored, worried?

Do you recall when you first came across the term "open" or "openness" used in relation to science - or has it always been used, in your experience?

How would you describe science that is open? What does it look like for you, in practice?
Do you have a vision for what open science should look like?

Is there a difference between "openness in science" and "open science" as far as you are concerned? Open source?

In your everyday work (e.g. talking to colleagues, in the lab/field, at conferences, writing publications…) is openness something you have tended to think about?
Are you making decisions about what to say or not say to others?
When advising students?

Has this changed over your career?

Have you ever faced a situation in which you wanted to share more scientific information than you did in the end? Or when you kept information back but regretted doing so?

Have you ever regretted sharing as much information as you did?

Did (or do) you tend to:
- Publish open access?
- Share raw/primary data?
- Share information about a study before publication? (preprints)
- Open peer review
- Open notebook science

Has this changed?

Have you been aware of any expectations or rules from your university of funder about how "open" your research should be?

--

Did/do you feel that there was a consensus in your discipline about how "open" science should be?
Should there have been? What? Why?

Think about a scientist who you know and really admire - someone who you think is a really good scientist. Could you describe their qualities as a scientist - what makes them a good scientist? Is this anything to do with openness, or is openness unrelated?
(or a bad scientist)

--

Is there such a thing as an "open scientist" for you? If so, what are they like?
Are they defined by what they do? Their attitudes? Their communities? Their character?

Anything else we haven't covered?

Sample interview schedule for advocate/policymaker

- Tailored to the individual

To start off with - for context -
I'm wondering if you could talk a bit about your background in science and how you've got to where you are now.

Could you tell me a bit about your background in research, as [e.g. a biologist]?
What type of research did you do in your PhD? How did you find the experience? Did you continuing to work as a biologist?

Which scientific community or communities would you say you were/are part of in biology?

How would you describe those communities? (character words)

--

Could you describe how you have got involved in advocating for openness in science?
What were your reasons for becoming an advocate?

Do you recall when you first came across the term "open" or "openness" used in relation to science - or has it always been used, in your experience?

Could you tell me what first comes to mind when you think about openness in science now? What does it mean to you?

Is there a difference between "openness in science" and "open science" as far as you are concerned? Open source?

How would you describe science that is open? What does it look like for you? Do you have a vision for what open science should look like?
Open access, open data, open peer review, open notebooks, preprints…

Is there such a thing as an "open scientist" for you? If so, what are they like?

Are they defined by what they do? Their attitudes? Their communities? Their character?

Have your views on openness in science changed or developed while during your years of advocacy? If so, how?

What do you hope to achieve in [role?]?

--

How do you think openness in science is understood by academic scientists?
How would you describe their attitudes? What types of openness are they keen (or not) to engage with? Why do you think some scientists are ambivalent or negative about openness? What are their priorities?

What is your opinion on the current state of openness in science (broader / UK policy)?
Are things heading in the right direction or not?
What, if anything, still needs to change?
What do you think the priorities should be?

How do you understand the policy landscape in the UK and Europe currently?
Who are the major players who you deal with in trying to create change?

--

From your point of view, is there anything particular that you hope my research could shed light on?

Are there any particular articles, reports or projects which you think I should pay attention to, that I might not be aware of?

Appendix C: Ethics approval

ATTENBOROUGH Ros

From:	Graduate School Office <gradschool.sps@ed.ac.uk>
Sent:	24 November 2016 16:09
To:	ATTENBOROUGH Ros
Subject:	Ethics form submission

Thank you for making a submission using the Research Ethics form. The information you submitted is detailed below.

Your submission will be reviewed by the relevant reviewer, who will get back to you in due course.

Section 1: Project details

Your name	Rosalind Attenborough
Your email address	s1562082@sms.ed.ac.uk
Project name	Scientific openness as a new epistemic virtue
Are you a student or staff?	Student
Are you an Undergraduate or Postgraduate student?	Postgraduate
Is your course Taught or Research?	Research
Please select your subject area	Science, Technology and Innovation Studies
Supervisor (if known)	Niki.Vermeulen@ed.ac.uk
Name of supervisor	Niki Vermeulen
Please paste a description of the research problem in the box on the right	Since the turn of the twenty-first century, an embrace of "openness" has become increasingly salient in scientific discourse, practice and policy. The dominant narrative tells us that science is open at heart, but in its present form lacking; and that now, the technology of the Internet can restore its essential nature. In the scholarly literature, few studies have examined scientific openness as an ideological and epistemic trend rather than a set of related, technologically-enabled practices; even fewer have considered the relationship between this trend and the experiences of scientists, who as epistemic subjects are creators and products of such trends. To address this knowledge gap, I propose "epistemic virtue" as a novel theoretical frame through which to understand present-day scientific openness. The concept is used by Daston and Galison (2007) to illuminate the construction of objectivity as a simultaneously moral and epistemic imperative for scientists in the nineteenth century. In the proposed research, I will conduct qualitative interviews - with a diverse sample of biological scientists - in order to characterise the twenty-first century vogue for scientific openness and investigate the possibility that it represents a moral-epistemic shift comparable to the rise of objectivity. I will also conduct interviews with policy makers and open science advocates, and analyse policy and advocacy documents, in order to build contextual understanding. My research questions can be summarised as follows: 1. What meaning does "open" have in the context of science for: (a) biological scientists; (b) policy makers; and (c)

1

	open sciences advocates? [Where meaning is created and expressed both in discourse and practice. The emphasis of the investigation is on (a).] 2. To what extent, and in what ways, are biological scientists constructing openness as an epistemic virtue? The focus on biology is due to particularly close links between parts of this field and the emergence of new "open science" (Delfanti, 2013); the field as a whole provides a bounded yet internally diverse context in which to explore the issue. Most interviewees will be recruited in the UK and thus to a significant the study will be specific to a UK context; however as scientific communities are international - as are discourses of openness that spread in scientific, policy and advocacy communities - this study will have some relevance for a wider context, and will include some non-UK interviewees.

Section 2: Potential risks to participants and researchers

Will any human or animal participants be studied either directly or indirectly?	Yes
State the nature of the risk and what measures will be taken to deal with such problems	Human participants (biological scientists, policy makers, and open science advocates) will be invited to take part in qualitative interviews as part of this research. The interviews themselves will be low risk as they will occur in a safe environment, and they will not involve any invasive or traumatic procedures or questions. Questions will largely relate to the participant's professional career and interests and are thus unlikely to provoke discomfort; however because I am analysing issues of identity and virtue in connection with professional life, I must approach the interview with appropriate care and respect. Deception will not be employed. All the participants will be adults who occupy relatively empowered positions in society and are able to give informed consent. Even after informed consent has been given, participants will also be allowed to withdraw at any time, and I will maintain their confidentiality and anonymity by default. Their data will be stored securely.
Are you recruiting participants that are or may be vulnerable in other ways?	No
Do the researchers need to be cleared through the Disclosure (Protecting Vulnerable Groups) Scheme?	No
Will it be difficult to ascertain whether participants are vulnerable in any of the ways listed above (e.g. where participants are recruited via the internet)?	No
Does the research involve sensitive topics (such as participants'' sexual behaviour, illegal activities, their experience of violence, their abuse or exploitation, their mental health or other)?	No
Is it likely that this research will lead to the disclosure of information (such as participants''' sexual behaviour, illegal	No

activities, their experience of violence, their abuse or exploitation, their mental health or other) that would require the researchers to breach confidentiality conditions agreed with participants?	
Is it likely that participation in this research or the dissemination of research findings could adversely affect participants in any other way?	No
Are you planning to reimburse participants for their involvement?	No
Will the true purpose of the research be concealed from the participants?	No

Section 3: Confidentiality and handling of data

Will the research require the collection of personal information from e.g. universities, schools, employers, or other agencies about individuals without their direct consent?	No
Does the research involve the collection of sensitive data (including visual images of respondents)?	No
Will any part of the research involving participants be audio/film/video taped or recorded using any other electronic medium?	Yes
What medium is to be used and how will the recordings be used?	The interviews will be audio recorded for the purpose of transcription, if the participant has given their consent.
Who will have access to the raw data?	During analysis, only the PhD researcher (Rosalind) will have access to the raw data. At this stage I do not plan to employ others to perform interview transcriptions, however should this become necessary they will sign a confidentiality agreement, and the information sheet will be adjusted to inform participants. Participants will provided with copies of their own raw data (in the form of an anonymised interview transcript). Participants may choose whether their raw data is deleted after completion of the project or archived for future researchers by an organisation like the UK Data Service. If they consent to archiving, the raw data will be accessible by registered users of the UK Data Service who have agreed not to attempt identification of individuals. Some participants may request that their raw data be made publicly available, and I will allow this as long as they are aware of the risks and as long as data do not breach the confidentiality or reputation of others.
Many funders encourage making datasets available for use by other researchers. Will the data collected in this research be made available for secondary use?	Yes
What arrangements are in place to ensure the consent of participants to secondary use?	Participants will be given a choice about whether or not their data is made accessible to future researchers via the UK data service. Although my funder strongly encourages such

X

	archiving, I wish to make this is a free choice for my participants as some may feel that the interview contained personal information that should remain confidential. Participants will be informed about the options regarding data archiving on the consent form, but they will not be asked to confirm their choice in writing until after they have had the opportunity to view the interview transcript. Even where the participant wants their raw data be made available post-study, I will review the sensitivity of that data before archiving, and will reserve the right not to archive the data it contains comments that, according to my best judgement, jeopardise the confidentiality or reputation of the participant or others that they are associated with (if the participate has chosen to waive anonymity I will still make this judgement re. reputation, but not re. their own confidentiality).
Will participants be identifiable, including through internet searches?	Yes
How will their consent to quotations/identifications be sought?	Participants will be anonymous by default: in my stored data, audio files and transcripts will only refer to participants by a number/pseudonym, and quotes/opinions of participants will only be reported or discussed in relation to those numbers/pseudonyms. A single file, stored separately, will link numbers/pseudonyms with minimal identifying information (names, contact details) for the researcher's purposes only. All files will be stored in digital form on the researcher's password-protected, encrypted laptop, or in hard copy in locked storage. Sometimes contextual information in the interview may give clues as to the participant's identity even if it is not explicitly stated. Therefore I will redact not only names, but also other potentially identifying information (e.g. job title) from the transcript as part of the anonymisation process. Where this information must be retained for the meaning of the quote/data, or where the participant may be identifiable based on context despite anonymisation of the transcript, I will seek specific consent from the participant for publishing/archiving those quotes/data. Because my research concerns "openness" and those with different ways of valuing and enacting openness, I will give my participants the option to waive anonymity and have their quotes/views reported with their name. This choice will be offered after they have had the opportunity to view their interview transcript; before that anonymity and confidentiality will be assumed. I will make it clear that anonymity remains the default choice and identification of the individual will only take place with their consent *and provided their identification does not compromise others' anonymity or their own/others' reputation*.
Will the datafiles/audio/video recordings, etc. be retained or disposed of after the study?	Retained
If retained, how will they be kept secure?	Audio recordings will be deleted after the project is complete as they cannot be fully anonymised, but I will retain anonymised interview transcripts - subject to consent by the participant (this option is given on the consent form) - for post-

	PhD research that I may do in this area. This data will be kept on an encrypted, password-protected computer (accessible only by the researcher, Rosalind). When they are no longer needed for relevant research being conducted by the researcher, they will be deleted for data protection reasons.
If disposed of, how will they be securely disposed of?	I will not keep copies of the raw data in my personal possession after the study is complete: digital files will be deleted and paper documents shredded. However, some data may be "retained" in the sense that some participants may have opted for their data to be archived by the UK Data Service (or similar).
Will feedback of findings be given to participants?	Yes
How and when will this feedback be provided?	Participants will be given the option of viewing their interview transcript after the interview is complete, and before any information is archived. Additionally, once I have completed an initial analysis of the interviews I will give participants the option of viewing a written summary of my analysis, and making comments. This will occur before I publish or submit my fundings, so that I can respond to or incorporate feedback in the final version.

Section 4: Participant information and consent

Does your research project require administrative and/or participant consent?	Yes
Does your research project require administrative consent?	No
Will administrative consent be obtained in lieu of participants consent?	No
Will written consent be obtained from participants?	Yes
Please attach a copy of the information sheet and consent forms	Interviewinformationsheet_24Nov2016.pdf, type application/pdf, 55.0 KB
[Consent forms]	Consentform_24Nov2016.pdf, type application/pdf, 45.4 KB
Will oral consent be obtained from participants?	Yes
Please explain your rationale and detail how oral consent will be obtained	Most of the time, oral consent will be supplementary to written consent. Where interviews occur in person the consent form will be explained at the beginning of the interview (having already been provided beforehand by email) and the participant will be given the chance to ask questions. If they then sign the consent form agreeing to the interview, I then will ask something similar to "are you ready to proceed with the interview?" to ensure that they are comfortable and feel ready as well as having technically signed their consent. This effectively a combined written/oral consent. If I encounter any circumstances where a participant is unable to return a signed consent form to me before the interview (e.g. the interview is on the phone and the participant is some distance away) I will go through a more detailed oral consent process: I will read out the information on the consent form and ask the participant for

	their response to each question (e.g. do you understand the information provided and your role in the interview? etc.).
In the case of research in online spaces or using online technology to access participants, will consent be obtained from participants?	Not applicable
In the case of participants with disabilities, learning difficulties or mental health problems, will arrangements be made to ensure informed consent?	Not applicable

Section 5: Travel risk assessment

Will the researcher(s) need to travel for the purposes of their research?	Yes
Will they be travelling to a non-EU destination?	Yes

Section 6: Counter-Terrorism and Security Act 2015

Does your research concern groups which may be construed as 'terrorist' or 'extremist'?	No

Section 7: Conflict of interest

Does your research involve a conflict of interest as outlined above?	No

Section 8: Research Proposal

Please attach a copy of the research proposal	Placeholder information sheet.docx, type application/vnd.openxmlformats-officedocument.wordprocessingml.document, 11.1 KB

Dear Rosalind,

Your Research Ethics form submission ("Scientific openness as a new epistemic virtue" [ID: 212768]) has been reviewed.

Category: Approved

Level: 1

Additional comments:

None

Appendix D: Sample information sheets and consent forms

Sample information sheet for scientists

Interview information: study of openness in science

What will the interview be about?
The interview will explore your opinions and experiences of "openness" in science. Your research and career, ways of working, and any experience you have with institution/funder policies will be relevant. The interview will last for about an hour and with your permission, be audio recorded. **There are no right or wrong answers: I am interested in <u>your</u> perspective.**

What are the benefits of participating?
Your comments will contribute to a deeper understanding of openness from scientists' perspectives, written up in my PhD thesis and communicated in academic and public fora. Ultimately, this understanding could inform and broaden scientific, public and policy conversations about the meaning of openness in science.

Who is doing this research?
I, Rosalind Attenborough, am the researcher for this study and I will conduct the interviews. I am a PhD candidate in Science, Technology and Innovation Studies at the University of Edinburgh. My undergraduate degree was in biology. My supervisors are Dr Niki Vermeulen and Dr James Mittra. My project is funded by the UK Economic and Social Research Council.

Will I be anonymous?
Yes, unless you choose otherwise. By default, records will refer to you by number or pseudonym rather than your real name. I will also redact details (e.g. job title) that identify you uniquely. I will consult you before publishing or archiving any contextual details that could suggest your identity. However, if you would prefer that your real name appear alongside your comments, you can request to waive your anonymity.

What will happen to the interview data (audio recordings and transcripts)?
I will keep your interview data securely for the duration of my analysis. After you have had a chance to view the transcript, you will be given choices about whether and how your data is archived. With your permission, I will retain your anonymised interview transcript for my future research on this topic.

Are there any risks?
This research is low risk: the interview will occur in a safe, professional environment and is not expected to cover sensitive topics. I will provide you with a copy of the transcript after the interview, and later I will let you know which quotes I plan to use. There is some chance that even if quotes are anonymised, you may be identifiable to readers particularly familiar with you. If this concerns you, I can work to minimise this possibility.

Can I withdraw from the interview?
Yes. You can withdraw at any time before or during the interview. And, if you decide afterwards that you would prefer not to participate in the study, or if you would like to retract specific comments, please contact me by [one month from the interview date] and I will arrange this.

Many thanks considering this information. If you have any questions, please feel free to contact me (Rosalind Attenborough) at s1562082@sms.ed.ac.uk.

Sample consent form for scientists

Consent form: study of openness in science

Please indicate Yes (Y) or No (N) in the appropriate boxes to signify whether you understand and agree with each of the following statements.

	Y/N
1. I have read and understood the **Interview information** sheet, and I understand my role as a participant.	
2. I understand that my participation is **voluntary** and that I may **withdraw** from the interview.	
3. I understand that I will be **anonymous** in any information reported from the interview, unless I request otherwise.	
4. I agree to **audio recording** of the interview for the purpose of transcription. Transcription will be done confidentially by the researcher (Rosalind Attenborough) or a professional transcription service.	
5. I understand that after I have been sent the interview transcript, I will be offered a choice about **the archiving of data from my interview***. No data will be archived without my consent.	
6. I give permission for the researcher to **securely retain an anonymised interview transcript** for her own future research.	
7. I agree to participate in this study.	

Name of participant: ____

Signature of participant: ____

Date:____

*The following choices will be available regarding data archiving: (1) no archiving; (2) secure online archiving of an anonymised transcript by the UK Data Service or similar; (3) public archiving of the transcript, anonymised or not according to your preference. This choice will be offered after you have viewed the transcript so that you can make an informed decision, and I will be available to answer any questions you may have at that time.

Sample information sheet for advocates/policymakers

Interview information: study of openness in science

What will the interview be about?
The interview will explore your opinions and experiences of "openness" in science. I am interested particularly in experiences you may have had - and knowledge you may have gained - while advocating forms of scientific openness, or creating/implementing policies around openness. The interview will last for about an hour and with your permission, be audio recorded. **There are no right or wrong answers: I am interested in <u>your</u> perspective.**

What are the benefits of participating?
Your comments will contribute towards new, reflective understandings of openness in science from a diverse range of perspectives, written up in my PhD thesis and communicated in academic and public fora. I hope these understandings will open new conversations and have conceptual and practical meaning for scientists, policymakers, and open science advocates.

Who is doing this research?
I, Rosalind Attenborough, am the researcher for this study and I will conduct the interviews. I am a PhD candidate in Science, Technology and Innovation Studies at the University of Edinburgh. My undergraduate degree was in biology. My supervisors are Dr Niki Vermeulen and Dr James Mittra. My project is funded by the UK Economic and Social Research Council.

Will I be anonymous?
This depends on your choice; you will be consulted after the interview. You may be experienced with publicly discussing these issues, and prefer to have your real name/organisation associated with your comments. If not, I will protect your anonymity by removing your real name and other unique identifiers from interview quotes. You can also choose to be anonymous selectively, e.g. if certain topics are more difficult to talk about publicly.

What will happen to the interview data (audio recordings and transcripts)?
I will keep your interview data securely for the duration of my analysis. After you have had a chance to view the transcript, you will be given choices about whether and how your data is archived. With your permission, I will retain your anonymised interview transcript for my future research on this topic.

Are there any risks?
This research is low risk: the interview will occur in a safe, professional environment and is not expected to cover sensitive topics. I will provide you with a copy of the transcript after the interview, and later I will let you know which quotes I plan to use. There is some chance that even if quotes are anonymised, you may be identifiable to readers particularly familiar with you or the issue of openness. If this concerns you, I can work to minimise this possibility.

Can I withdraw from the interview?
Yes. You can withdraw at any time before or during the interview. And, if you decide afterwards that you would prefer not to participate in the study, or if you would like to retract specific comments, please contact me by [one month from the interview date] and I will arrange this.

Many thanks considering this information. If you have any questions,
please feel free to contact me (Rosalind Attenborough) at s1562082@sms.ed.ac.uk.

Sample consent form for advocates/policymakers

Consent form: study of openness in science

Please indicate Yes (Y) or No (N) in the appropriate boxes to signify whether you understand and agree with each of the following statements.

	Y/N
1. I have read and understood the **Interview information** sheet, and I understand my role as a participant.	
2. I understand that my participation is **voluntary** and that I may **withdraw** from the interview.	
3. I understand that quotes from the interview will be reported either **anonymously** or **with my real name/organisation**, depending on my choice. I can also choose to be anonymous selectively, e.g. on sensitive topics. These choices will be offered after the interview.	
4. I agree to **audio recording** of the interview for the purpose of transcription. Transcription will be done confidentially by the researcher or a professional transcription service.	
5. I understand that after I have been sent the interview transcript, I will be offered a choice about **the archiving of data from my interview***. No data will be archived without my consent.	
6. I give permission for the researcher to **securely retain an anonymised interview transcript** for her own future research.	
7. I agree to participate in this study.	

Name of participant: ____

Signature of participant: ____

Date:____

*The following choices will be available regarding data archiving: (1) no archiving; (2) secure online archiving of an anonymised transcript by the UK Data Service or similar; (3) public archiving of the transcript, anonymised or not according to your preference. This choice will be offered after you have viewed the transcript so that you can make an informed decision, and I will be available to answer any questions you may have at that time.

Appendix E: Interview transcripts

Some interviewees opted for their transcripts to be made publicly available (or archived online in a controlled access way), and/or to be named. I have not yet uploaded the relevant interviews to data archiving services (I will complete this after submission). However, I include two interview transcripts here as samples of method and data. These interviewees opted for their transcripts to be made publicly available. Ernie is anonymous, and Jenny (Professor Jenny Graves) is named.

Interview transcript: Ernie

- Opted for the transcript to be publicly available with pseudonym
- Made some edits to the transcript, for flow and confidentiality of third parties
- Recruited by "cold" emailing; interview took place in-person, October 2017
- UK university context

So to start out with could you tell me briefly about the kind of research you're currently involved in?
Yeah, we're currently involved in studying what happens to the chromosomes when cells divide. So we're interested in how the DNA - the chromatin transforms its structure from the interphase nucleus to make the mitotic chromosomes. We're interested in how the mitotic chromosomes attach to the mitotic spindle - which is what guides them in their movements during mitosis - and we're interested in how that process is regulated, in particular by one complicated protein kinase.
Okay, right. And is that... As part of that, how would you define the discipline or disciplines that you're working within?
I'm a cell biologist. And considered to be an expert on chromosome structure and mitosis.
Right, okay. Is that how - Is that the most comfortable identity you would give yourself, cell biologist?
Sure, sure.
Has that been stable throughout the years, or is it something you've changed over your career?
I started as a molecular biologist, but I started a long time ago when molecular biology encompassed lots of things that included cell biology. But I've considered myself a cell biologist for many years now, yeah.

Okay. So your initial training was...molecular biology. Is that how it was talked about at the time, or was it...?
It was just called biology [laughs].
Just biology, yeah.
Yeah. But I did my - my undergraduate I didn't particularly train very much in biology, I did that only when I went into postgraduate work, and I did that at [institution], and the department there was very molecularly oriented. Molecular genetic orientation to that department at that time.
Is there something distinct about that time of being a molecular biologist that's different from being a cell biologist? In terms of...
[Pause]. Not really, a lot of what I did could also be lumped under microbiology. I studied viruses. I studied bacteriophages, viruses that infect bacteria. But I was doing structural studies of them by x-ray diffraction analysis and also doing genetic studies. It was kind of a complicated mixed PhD with many different aspects to it. And today all of those things are covered in our textbook in cell biology that we have published.
Okay, right. And so - sort of, being in all these different disciplines, I'm guessing that that means you're a member of maybe one main scientific community or several scientific communities where you do your work - where you sort of review other people's work. What would you say those are? The scientific communities you're in.
Mostly things having to do with chromatin and chromosomes and cell division.
And is there any particular word or words you would use to describe those communities - some people say the community is big or small or competitive or friendly, or has some kind of character...
I'd say that those communities tend to be very competitive and not always particularly friendly.
Hm, right okay. Well I'll ask you about openness now, and - just to start off with, I'm wondering what first comes to mind when you think about openness in science, or whatever you thought this interview...would or should be about?
Well it's your interview so I didn't [laughs]...I didn't have any particular ideas for what it should be about. I mean, there's the whole issue of openness in publishing, there's the issue of openness in how you interact with others - I mean those are the two major things that I think of when I think of openness. There's this new thing coming along where - wanting people to make all their data open. But that's, in my opinion, particularly groups of people who sort of have an axe to grind and don't think about - really about how what they're arguing for will impact many other people who work in different ways from them.
Mm. Do you want to expand a bit on that part of it?

Well, so a lot of the people...there are people who are arguing that all - basically lab books, all lab books should be electronic. At one extreme there are people who are arguing that the, the content of those electronic lab books should be made accessible to anybody at any time - to everybody, basically. Right from the moment they're created. This is...completely unrealistic for a number of reasons. It's - the first reason why it's unrealistic is, I have my own ideas and my own methods that I want to use to answer questions. The fact that I can convince people that my ideas are better than other people's ideas, is how I get grant funding. If everything I do is immediately made open to all of my competitors, then where's the advantage in me having ideas, I may as well not have ideas, I may as well just get good at...surfing the web and finding what other people have done and not yet even published and then jumping on top of it. And so I, I think that that's really...the people who do that don't think about the fact that competition is extremely important in getting funding. If I can't convince people - if I can't convince the [funder], that I am the right - so whenever you write a grant, you have to have a really important problem, you have to have a really powerful way to study that problem, and you have to be able to convince a committee of people that you are the best person to study that problem. Better than anybody else. Time after time, people will not get funded, even though they've identified a good problem, they have a good method, and then the committee sit there and say, "Oh yeah but that, that group off in America's already doing this. They were always going to do this better, and this applicant can't compete." So you have to have that edge, and if you make everything open, where's the edge? You lose the edge.
Mm-hm.
Alright so that's one thing. The other thing is... People want, they want to have all the data out there, well if I'm doing genomic sequencing - and so what I do is I isolate some DNA, and then I do whatever I do to it, and then I send it off to a facility and I get back a few terabytes of sequence data which I'm then going to work with. That comes back as a big huge file, and it represents a lot of - you know, maybe months' worth of work or whatever. Sure, I can make that available to everybody when I decide I want to make that available. But what if what I'm doing instead involves taking thousands of individual pictures of things with a microscope. Am I then supposed to put these - you know - figure out a way of formatting in which to store these thousands of pictures and then make them all available to everybody? That's going to take as much time or more than it takes get the data in the first place. Where - where's the allocation of time for doing that? I'm - you know, I'm supposed to answer a question within a certain amount of time, and the [funder] – they don't give me any money for a secretary, and they wouldn't give me - they wouldn't say, "Oh yeah well six months out of each year can be spent in the process of archiving your data so that you can make it open." But if that's what it took then it simply means that I'm losing six months of research time, so for people whose data involves collecting many scraps and different pieces from all over, and all different types of data, to then have to assemble all of that and make it available to everybody in a comprehensible way, is a huge, huge time burden. And people arguing for open data often don't think about that. So the people who are in favour of this open data tend to be people who produce huge datasets, which is like one big lump, and you can release the one big lump. But when you produce thousands of little datasets, now you're talking about a whole different range of - in the work involved.

Mm, okay, so these people who you - you think have a sort of particular axe to grind with open data issues, is that how you imagine that arises, that they have a different kind of data to deal with? Their sort of workflow is completely different? Is that where you think this comes from?
A lot of the people, their workflow is quite different, yeah, yeah. Uh... Yes, yeah.
Okay. So does the whole - when I talk about the issue of openness in science, does it feel like it's something relevant to you...?
Oh, sure.
Is it something that you...like you feel a particular way about, like sort of interested, or passionate, or worried, or annoyed, or...
Well, one of the things I enjoy most about science is going to meetings and visiting other places and talking to people about my work and having them talk to me about their work. And people who are constantly playing a game, that you know - you show me yours and I'll show you mine, and I'm only going to show you - I'm going to figure out how much what you told me is worth and I'll show you that much - I mean that's just...that's a drag. So I cultivate relationships with people who I trust and then we just talk freely about ideas, and I do that with people who I compete with as well as people who I'm collaborating with. And... But the competitors that I do that with are competitors who I know and trust.
Okay. How do you know if you can trust someone, in the context, is it...?
You just... Lifelong experience gives you some hints for how you, you know, who you can trust and who you can't, but sometimes you make mistakes, and there are people you can't trust. Others, you can. I mean I... Years ago, I felt that I was stabbed in the back scientifically by one of my best friends. Many years ago. We used to go out every Friday and have dinner together, and have a drink, and just socialise and have a good time. And then... And we were collaborating a project, and then without - on a project - and without saying anything to me he wrote a grant to compete with me.
Hm.
At the same time as he was doing all of this, and so that really hurt. But...that didn't stop me from having other very close friendships and collaborations with people, and thankfully that's never happened again.
Did find that experience formative in any way, did you...
Well... It certainly influenced my future interactions with him, I mean I still interact with him years later and we are still friends and enjoy each other’s company socially. But even though he has come a long way in his career and is now super-successful and would be extremely unlikely to need my ideas ... Because of that residual memory, I don't go out of my way to tell him things that I'm thinking about doing that I think he might be interested in working on.

Mm, okay. And what about with other people, do you...are you more...?
Well with most of the people I interact with, there are just some people who you...just the word is out there. You know, we were talking about - I went out to dinner with a seminar speaker last night and the name of another person - not the person I was just talking about - came up, and two different people round the table said "Oh," you know, "he's a terrible shark, I wouldn't - nobody has anything - nobody wants anything to do with him", so there are people who just have this reputation. I mean one person once told my students that the only good reason to go to a scientific meeting was to steal ideas. I think he thought he was being funny, but that is actually what some people do [laughs].
Wow.
You know, I - personally, that doesn't interest me. If I can't have my own ideas, then I should retire. 'Cause what science for me is about is the creative process of having ideas, and then seeing if you can figure a way - first of all, they have to be good ideas, and then you have to see if you can figure out a way to see if they're true or not. That's what the fun is.
Mm, okay. Did you have any - thinking about this way of learning who you can trust and deciding what ideas to share, and how to be open with other people, do you - did you have any - Yeah, is there any aspect of your training in this, did you have any influential mentors or people who showed you how this kind of thing is done?
Not really. I mean my PhD training was in the bacteriophage, in the virus assembly community, and there was a - there was a specific meeting for those people at [location/institution] lab that we used to go to, and we used to socialise with other people. But I don't - but there wasn't a big history of the group leaders collaborating with one another. Uh... I think they, I think they talked about - at that meeting you would talk about unpublished work, and it would - and people wouldn't steal your work. So maybe to that extent there was, but... No, that just came...it's just obvious that if you can collaborate and become friends with people who have complementary expertise to you, then your life is going to be richer. Now I collaborate a lot with people in [country], and that... that involves a very particular type of process. Because most of the [nationality] people who I have worked with and know, do not place a high priority on respecting boundaries on things. If they see something, they believe that everyone has a right to work on it, and they'll work on it. So the way I collaborate with those people is we talk in general about what we're working on, and then basically we just end up both doing what we were going to do anyway. The important key is that when we have finished up an overlapping project, we get together and we talk about it and we put together what we all have and we always publish together. But we don't say - I'm going to work on this and you're not going to work on that, because that just doesn't work. They work on it anyway. But they don't do it maliciously, it's just the way they, it's just the way that the people I've interacted with are. And I've had very long term friendships and interactions with a number of people in [country]. And they've been friendships that have really enriched my scientific life.

So, and when you're collaborating with other people, are there - those boundaries are sort of...sort of exist, and they are unspoken or...?
Well... Not necessarily, I mean at the moment we're working on a big - we're collaborating on a, we have - I mean, I have a number of collaborations going at the moment - there's a big one where we have a paper which is, we're revising in [prominent journal], and there my lab is doing all the wet cell biology and biochemical work, another lab is doing the molecular biology analysis, and a third lab is doing very high level mathematical modelling of the data. None of the labs would really be competent to do what the other labs are doing, so it's a perfect collaboration in a sense, because everybody knows what their niche is. And then, the paper's going to have three - the paper has three first authors, and three corresponding authors.
Mm, okay. I'm wondering - going a bit back to the issue of openness in general, is there a time when you started hearing the word "openness" or "open" in relation to science - is it something that you've always...thought and heard about, or is...?
Well, yeah, no. No, I - I think that started, I mean if I would think about openness, I would think, well primarily - the first big conversation about this was about open access in publishing I think. And in my understanding, this was basically because the University of California couldn't pay Elsevier's bill. Elsevier - because Elsevier bundled all of their journals together and they charged, and they were charging the University of California all this money, and California had a budget crunch, and people couldn't really afford to pay it, so a lot of them became real activists for saying, "Well we think this should be open, publishing should be open", but I think it was this community of people, primarily in the Bay area, who started pushing that. And because they were very influential, very powerful, very good scientists, they got noticed, and then other people joined in with the idea, and of course now that's - it's pretty much - except that I mean if you publish with the Wellcome Trust or Max Planck or other big societies you're expected to publish things open access. Which is - that's fine with me, I don't see necessarily why anybody should have to pay. Just have to, you know, just have to figure out a way to make it work for publishers.
Yeah.
I mean I don’t know, but I think PLOS makes a tonne of money, actually. Because they just charge these big publication fees to the authors. So they don't, they don't do it by selling subscriptions, they do it by making the authors pay - and as I'm sure you know, there's this world of predatory publishers out there now who are just trying to get money out of people that publish papers.
Mm, yeah, indeed. Is that... Yeah, there were a few things that came up there - but, so you're aware of the, the sort of expectations or policies of the university or funders - do those sort of impact on you directly?
Well, so - I'm a [funder name] fellow, so I have a kind of a slightly odd attitude about the [home university name]. I mean they host me, and they administer the payment of my salary, but actually it's the [funder] who pay my salary. And, so I've alwa- ... Actually, there's a great book which you probably never could find, because it's, I think it's long out

of print, but it's an amazing book called "Academic Gamesmanship" - it's a really very sarcastic view of how to be successful in an academic career. And one of the points that the guy made in that book is, the most important thing is to figure out what you're going to be assessed for, and then do that first. And - you'd just be amazed at the number of people who don't do that. You know, people get hired, people come here as [funder name] research fellows, and there's all this buzz buzz buzz around the university - all these people saying, "Well you know, some day it's going to be really important if you know how to teach, you really need to teach", and so they convince a lot of these people to get involved in teaching courses. Then, four years down the line they have to renew their fellowship. In that renewal, the *only* thing they're judged on is their research output and ideas. So *all* the time that they spent teaching, wasn't doing anything that was going to help them to get their fellowship renewed. And if they don't get their fellowship renewed, the university can say, you know, bye-bye - you're gone. Well, you know how to teach, maybe you can find a job at a teaching university [laughs], I mean it's - people - you have to pay attention. So if the [funder] say "we want everything published open" - I do what the [funder] says [laughs].

So that's - that's something - do you have to put much effort into that, thinking... Do you, like, always - sometimes - publish open access? Or...

No we always publish open access.

Is that - in an open access journal, or is that with a repository system, or...

No, we... no, so we're very lucky because the [funder] then gives a pot of money to the [home university name] and I contact the [library staffer name] in the library and I say, "[name], we have this paper that's coming out in [journal], which - I mean in one case several years ago the [society name] journals were pretty slow actually at getting involved with open access publishing, and I think we published one of their first open access papers and it took poor [staffer name] about 20 or 30 emails back and forth for them to figure out how they could take our money. Even though they had this option "open". So there is somebody who will do it, and the [funder] will pay for it. All I have to do is to...is to make sure that that option gets ticked. Now where it becomes a bit of a pain is in some of these collaborations where we're not the communicating lab, but I'm an author on the paper. Those papers are also all supposed to be open, and other people don't necessarily have the money to do that. You know, if it's three - three or four thousand dollars [£3000-4000] a paper, they don't necessarily have that, so then we have to try to work out arrangements...where [home university name] can be billed for it. But the [funder], I mean - they're, they have been a bit of a pain, recently because now there are certain types of open access license that they like and other ones that they don't like. Like they don't like the open access at PNAS. But they're not going to tell me not to publish in PNAS, but they - they refuse to pay, the way PNAS publishes open access - well it's open access! I don't know what it is, it's some technicality over who owns what. But it is open access. No, they won't pay. What a pain. Then we have to figure out how to...you know, what to do about it. So sometimes it's quite - sometimes it is a drag. But in general, it - actually, we're blessed, it goes pretty smoothly, and the guy [library staffer name] who's in charge of doing it, he's great. I hope he retires after I do [laughs]. He's one of those people at [home university] who does a great job, which is...[removed for institutional anonymity] - it's definitely not 100%.

Hang on I'll just check what time we're up to, because I want to - yep, make sure we stick within time. So going back a bit to the open data part of things, you - you gave me some of your views on that, before. So I'm assuming - does this mean that you don't practice that kind of open data...?
No.
Do you do anything particular with - what form does raw data take for you, in your research?
Raw data is lots and lots of image files... I mean the other thing about this is it's completely impractical. Alright. Suppose we make a movie of dividing cells. We make that movie over a weekend. Every thirty seconds, we sample five locations, and we take a stack of images. That movie is two terabytes. Suppose we make ten movies. How am I supposed to put 20 terabytes online? Whose building network is going to transfer 20 terabytes of data? I mean, you know, just like - it's just not realistic. So imaging data - I mean, imaging is really interesting, imaging is sort of - has the potential to break the back of the computing revolution. Ever-faster, ever-bigger, ever-smaller, all this - but we can produce *massively* ever-more amounts of data. Just *huge* amounts of data. So we have huge amounts of image data, those are on a server. We have, for some of our publications the Journal of Cell Biology, for example, has something called the Data Viewer. And if you want, you can put all your raw data into the Data Viewer. You can either do it when you submit your publication - which seems a bit odd to me, because your publication might get rejected. But you can certainly do it after the publication has been - after it's been published, and we've done that. So, you know, if people want the raw - I don't have any problem with giving people access to the raw data, but I think that people should have access to the raw data that's in the publication, and that's backing up what's in the publication. And if I had to take 500 pictures, and I come up with 20 pictures that end up - pieces, bit and pieces of them are in the publication, it's a lot more practical to make those, the raw files for those 20 images (which actually might end up as 80 individual original source images if we are working with four colours) available than all the 500 (which could be 2000 source images!), all the ones that didn't work or were not informative.
Mm, okay.
So I prefer - I prefer a model where we...we make the raw data available for what we - for the data that we publish.
Mm. Do you see anyone doing something like that, in your field?
Well people do do that through the JCB data viewer and through Omero servers – we are not experts at that yet, but are looking into it. I don't know...if any other journals are doing it. I mean they had to get grant money to do this I think, and - you know, it's expensive - it's a very large server which somebody has to maintain.
Okay. So you don't think it's likely to happen in your field, that it becomes standard...

Oh I think it probably - eventually something will happen. I'm not sure. I mean eventually we will go to electronic lab books in my lab - at the moment people use paper. Most people in the [funder] centre use paper. And...one of the reasons we're not moving to electronic lab books, is that there's no standard for what to choose. I'm not going to make my group learn how to use an electronic lab book and then have it come from, down from on high that we chose the wrong one, and we have to use a different one, and of course nothing is going to be compatible - you know, so we just... I'm just being very conservative on that. We'll see what happens. Eventually if a directive comes from the [funder], and it tells us which electronic lab book we have to use, then we will, we will figure out how to try to make it work.

Mm. But that would be...a sort of private one...to your lab...

Well, yes, but once you have electronic lab books then you can decide what you'll show. But I don't see any reason to show people data prior to acceptance in a journal. Except for my collaborators. And my friends. And people who I trust. You know... What I own, and nobody else can have, are my ideas. They come from inside my head, I generate them through the happiness and misery of my life, that's where the ideas come from. When I graduated university, I was going to be an artist I wasn't going to be a scientist. But I found that the process of designing beautiful experiments actually satisfied the same creative urge in me, but I do have sort of an artist's view of ideas - that ideas area really...really critical, and they're unique, and they come from your soul. And other people don't own them. So even in a totalitarian society - you just can't make people share ideas. I'll share my output. I'm happy to share my output, but my ideas are mine.

Have you heard of, sort of, preprint servers?

Mm-hm.

Do you have particular views about them?

Well, I have three papers on BioRxiv.

Oh right, okay.

Yeah.

Is that something that you - when do you put them on there, is it...

Well - one we put onto BioRxiv because we weren't quite sure what we were do it - we weren't terribly worried that anybody else was going to come along, it was a computing...it was a computing paper, and it was something that a student actually figured out during a - during a PhD rotation which ended up being published. And so we put that onto BioRxiv and then that was quite interesting, because a journal editor contacted us and said "Would you submit that paper to our journal, if you haven't submitted it somewhere else?" And it ended up coming out in that journal. That's one. Another one was a paper that we've submitted around, to a bunch of journals - I have a very very talented senior [nationality] postdoc who finds it difficult to write in English and so she's always very, very slow. She should have written the paper a year ago, but she didn't and she got scooped. Some parts of her story got scooped, other parts didn't but a whole series of journals had decided to only focus on the parts that got scooped, and so it was editorially rejected a number of times. And part way through that process, we said - well this is ridiculous, I mean we have good data here and we may as a well sort of put a marker down that we've done it, and so we decided to put it on BioRxiv. And so that's on BioRxiv. And then this big paper that we're working on for [prominent journal], even though it's an incredibly competitive area, we decided that we wanted to have it known that we had done it, so the authors all agreed and put that on BioRxiv. That's weird because that's even been referred to in a pa- ... I didn't know you could refer to things from BioRxiv, but it's been cited.

Do you think that's - so that's a sort of new thing that - to put something in a competitive area out on bioarxiv, for you?

Well - if people want to say "I did it first"... I mean I could- You know, it would have had an impact on me in the past. For example, I discovered the first proteins to be discovered at the centromere region of chromosomes, so these are the proteins that direct chromosome segregation. I discovered them and I had an awful fight to try to publish these papers, so after I discovered them it was over a year from the first time I submitted the paper - it was something like four or five journals later before finally anybody would publish them. I mean, the paper got me elected to the [prestigious organisation], but nobody would publish it. And... Then, it turns out that a month before my paper finally appeared in print, another person published a very small part of that story. In a specialised journal. And they published their paper in December of a year and my paper came out in January of the next year. And even though...I have the letters saying that my paper had been submitted the previous January - I can't... There are a few people who like to try to take me down a peg and say "Well you got scooped" - well, I didn't get scooped. I just...you know. It was the pu-... So if I had been able to put that on BioRxiv ...that would have answered the question, right?

So do you see it as a positive development...

Yeah. Yeah. Definitely - if you're worried about getting credit for an idea, and if...scientific journals will still take your paper - I mean I'm kind of amazed actually given how - it causes all sorts of weird problems, you know... So, [prominent journal] is going to accept a paper that's been on BioRxiv and yet they're...and they're going to put an embargo when the information on the paper can be released - except the paper was actually published a year ago. [Laughs]. It's...interesting.

Yeah, yeah. Very interesting!

Yeah. Well we'll see what happens, hopefully. It better get into [prominent journal].

Yeah, yeah. Is there...you were saying before that you...you know, you work out who you can talk to, and who you can trust, and you don't like thinking about this sort of - you tell me something and I'll tell you something. But do you, would you say that you think about openness on, on an everyday basis - sort of, when you're talking to colleagues, or at conferences, or writing papers? Is it...

I'm just, I'm a...I just tend to be open with - unless, unless there's somebody with whom I feel that I better not be open, I try to be open with most people.
And what about with your students in terms of - do you find yourself needing to give them advice about what to say or not say?
Yeah. Sometimes. I mean I - we had a visiting seminar speaker yesterday, and one of my students was gonna tell him something about his work, and he was very worried - you know, should I talk to him about this, what do you think, you know... I told him - first of all, I know the guy, he's very busy, he's not going to rush back and try to scoop us. Secondly, you've got a paper which you're revising for a journal, that's already been refereed and hopefully you'll be able to answer the referees' comments and get it published, in that case if you do, the paper will be published within a couple of months, so even a big lab isn't going to be able to start a project from scratch and beat you in a couple of months, so you don't worry about that sort of thing. But you don't feed people ideas...that they'll find irresistible. So you just... So, yeah, so I am not a big proponent of posters in meetings. So the people from my lab, we do posters, but we don't put our really interesting, tantalising, someday-to-be-realised projects on posters. What we put on posters are things that are either done or almost done. In general. There are exceptions, but in general that's what we do. I don't want to put something out there for everybody to see unless if we had to, we could beat them.
Okay. Yeah - just do another quick check of the time.
I mean we have actually - I mean we have taken advantage of that ourselves. We went to the cell biology meeting, one year and we had a project that we were working on and the project was kind of mid-way. We went, and saw a poster - and somebody had discovered the same protein in a different system, by a completely different way, and...and they were studying it. I talked to that guy, and he had a paper that was actually in press at [journal], and we came back and - I had a, they were very good students, very organised group of people, they finished all of the experiments they needed to do in the next month, and we actually beat the [journal] paper to publication cause [journal] was slow. But... I was open with the guy from the meeting. He knew that we had the same protein and we were both working on it. But it - it, that was an example where seeing something - we said well, we have to get off our backsides and really push on this. And we did. You don't necessarily - depending on how competitive you're feeling, you know, whether - depending on how long it had been it is until you have to get your grant renewed the next time, you may or may not want to be...to do for the good of the scientific community and stimulate other people to hurry up - maybe you'd rather beat them. So that's a complicated issue. When it comes down to fighting for your survival, you do fight – but you can do it honourably and respect other people. I always want to win on my own merits because what we do is great objectively speaking – never because we are better than someone else. I do not see science as a competition.
Yeah, yeah it is. I'm wondering if you've - a couple of situations, like have you ever faced a situation where you wanted to share more information than you did in the end, or you sort of regretted holding back, in any way? Or do you feel like it's generally been the right thing to do?

No, I generally share the information. I mean with the - so my two [nationality] collaborators, basically when I go to [country] I just - I talk to everybody in their lab, and people tell me everything they're doing, and when I talk to them I tell them what we're doing without holding anything back. And what comes out of that are, are... In one case, they weren't looking at their data in one particular way. Or there was another way to look at their data and I pointed it out to them, and we got a really nice big collaborative publication out of the changes that they made after I'd pointed that out to them. So that was really worth it.
Yeah. And if you - I mean we've spoken about potential instances, of this, but have you ever felt like you regretted sharing as much information as you did, you sort of...
Well, back in the distant past with that one person who wrote the grant to compete with me, I regretted - but on the other hand I needed him, my lab were not experts at molecular cloning, DNA cloning at the time, and this was in the early days of DNA cloning actually, and he was an expert at it, and we needed to clone something. So it was a marriage of convenience, which then became a friendship, but then unfortunately our interests clashed. Other than that...I don't think so. I don't think I've regretted it. Looking back from the perspective of many years later, I am glad that we worked so closely and productively together. There were huge benefits for both of us as well as some costs. But in hindsight, those costs were worth paying and friendship has survived.
Okay. And so overall do you think - in your field which you describe as quite competitive - do you think there is a consensus about how people ought to behave with regard to openness? Or are people very different in the way they approach things?
Different people are different. Sure. I mean there are people who talk about everything and don't...and just figure that it'll all be okay in the end, and there are people who are very very close-mouthed about what they say, and all they say are things which - they talk about things that are in press or published.
Where would you put yourself there, do you think?
In general if I'm talking at a meeting, I'll talk about stories that are - that I think are almost finished. But not necessarily in press yet. So...probably a little more open than some, and...probably, probably more open than most actually. I love to tell a story, but it might not be wise to tell a room full of writers the outline for the story you plan to write someday but haven't started yet.
And do you feel like anything should be different than how it is, in terms of how people behave around openness? Would you prefer...
Well, so the only issue about openness that actually we didn't talk about, is openness in refereeing papers. I...I strongly believe that all refereeing should be open. And that people should know the identity of the referees. I think that - you know - the misery, the major misery of my life professionally has been dealing with vicious referees. I mean, we were - we were the first people to develop a system to do something that sounds impossible, which is we developed a cell-free system using ground-up cells to study cell death. You would

think, well how can you study cell death if everything's dead to begin with - but our system really revealed a lot of interesting things about apoptosis which is a type of cell death. The first reviews we got of that, we sent the paper to the journal [name], and we were told "There's nothing you can do that will make this paper suitable for publication." That was actually true, somebody wrote that down! And [journal] rejected our paper, and in the end they didn't know whether to take it or not take it, and they asked an expert, who came and told me later, they said, "How do we know if this is right or not?" And he said, "You don't know if it's right or not, you only know what you see." So they decided to reject the paper. So it got published in the [different journal name], and it became pretty influential. But then the original prestigious journal published three knock-off papers by big-name labs who just copied us. And because our paper hadn't blown up and turned out to be wrong... So the pioneer gets screwed over [laughs] and the copy cats get the big high impact publications. And that was - it's the quality of refereeing. In my first story where I identified the centromere proteins - people had never heard of me before, I was young at the time, they didn't know who I was, other people had tried to do this and they'd failed, and I got comments like, you know, "You couldn't possibly know the answer to this, you can't possibly be right, so-and-so's already tried this and they didn't get it to work, so I can't see how this can be working", you know, just - comments you wouldn't put down if you had to sign your name to them. And I frequently sign my reviews. If people don't like what I say, then they're free to tell me, but I never review papers vindictively. And I never ever tell people that they have to do a bunch of experiments just to do a bunch of experiments because it's a stupid - that's, it's not, in the end, it's their paper, it's not my paper, it's either good enough to get published or it's not good enough to get published. So the closed...I mean, students and post-docs waste so much of their lives - particularly people who want to publish in high impact journals - they waste so much of their lives doing make-work demanded of them by these anonymous referees. So... And if everybody had to be open, nobody would stab anybody - people wouldn't get even with you because you rejected their paper, because...you could - you know - you...they can reje-...you know, everybody knows who everybody else is and people wouldn't say the nasty things that they say if their name was going to be associated with it.
Yeah. Do you think that it would be viable for junior scientists to review openly like that, do you think it would work?
If everybody was doing it, yes. If the review process w-, if it was accepted that the review process was open - yeah.
Okay.
I mean how is somebody going to get even with you - they're going to get even with you by rejecting your paper? I guess they could get even with you by trying to stop you getting money or trying to stop you from getting a job or some really horrible thing, but that wouldn't - I would hope that even most cynical people would regard that as an overreaction, but if, you know, you think, if I reject their paper they're going to reject my paper? But if they reject your paper they'd have to sign it too. And if they did it for reasons that were just not sound, then everybody would know they were just doing it out of vindictiveness, and so they wouldn't do it. But I don't think it'll happen. Where I hope - what I hope will happen is that the process will break down, actually. And one way to break the process down is for

people to put their reviews online and then invite people to analyse them with text editor programs, and I think over the course of time text - text recognition programs will be able to identify a lot of the people who write the anonymous reviews anyway.
Okay.
[Laughs] So I would be quite up for outing anonymous referees by identifying who they are.
So when you talk about the whole thing breaking down, do you mean publishing, or...
Not publishing, no, just the - this business of the protection of being anonymous and being able to say lots of vicious things.
Oh okay, so it will technically no longer be possible to be anonymous.
Well I think we are approaching a world where it's going to be, it's getting harder and harder to be anonymous. You know, people can walk around with a camera and take a picture of you and then they could figure out who you are. If they've got access to the right software they certainly can. The government can.
Yeah, yep.
So that, you know, that - Actually there was a really interesting, very strange book called Super Sad True Love Story, about a world where people are walking arou-...it's written maybe ten years ago, in a very prescient, it's written about a worl-...people are walking around with these things which are the equivalent of iPhones, and you walk into a room and you just take pictures of people and then you, you can immediately see everything about them, how much they're worth, everything. This is not in the impossible future.
Yeah, not so distant maybe.
Right, so why should we hide when assess somebody's work, why should we hide who we are.
Yeah. I just have a couple more questions.
Okay.
Just sort of stepping back from all of this - or maybe this is actually my final question, yeah. We'll see. Just stepping back from all this, I'm wondering if you could think about scientists, or a scientist that you really admire, or someone that you've looking up to in your career, maybe someone that you've worked with...
Mm-hm.
Something like that. What do you think the qualities are of that person, what makes them a good scientist, for you?

Well, one of the people who I most admire is one of the guys who I worked with as a PhD student. He wasn't my - he wasn't my titular advisor, but I worked in his lab, and we published papers together, and he was - he still is - he's in his 70s, upper 70s and he's still extremely active and he's just intellectually brilliant. He doesn't compromise. He loves science. He could be pretty tough, but then we sat together writing a paper once - which was an article in [prominent journal] actually - and he just sat there, and I, what I remember about the dialogue that we had when we were writing is that he was...he let me speak first. And then he would sort of modulate what I said, instead of telling me how he could say it better. He let me do things in my own words. And so it was a really, really sensitive experience. And I see him at meetings, and he still is really admired and - by the young people. And, you know - so he's scientifically excellent, but he's able to appreciate other people's contributions and make other people other feel that they're worth something.
Do you think that those qualities or what he did have anything to do with - does that relate in any way with openness as we've been talking about it?
I would think...I think he's sort of a standard...he has sort of a standard attitude to openness I think, he talks about work when it's, when they have a picture for what they're doing to do, you know, when they know where things are going to go. I don't think... He's a crystallographer, so a lot of - crystallographers, until you have crystals you don't have really - you can talk all you want, but you don't really have anything, if you don't have crystals you can't get a structure, if you're a crystallographer. And if you do have crystals then probably nobody else has the same crystals so you're probably okay, so - and that's a kind of community which is a fairly big community, where it's fairly safe to talk about things because it's hard to scoop other people. It's not like when it involves just growing some cells and seeing what they do. But I've worked with - I mean one of the really great people I worked with was the Nobel Prize winner [name], he was my [working relationship descriptor], and he was absolutely paranoid about secrecy. He wouldn't tell anybody anything. And he always felt that other people were trying to steal his ideas. So - but he was still a great and influential scientist and became [leadership role] of the [prestigious organisation], and all - and on and on. But that would be an aspect of him that I recognised even back when I was a post-doc that that was something that I didn't want to copy.
Mm. Why didn't you want to copy it, did you...
Well, because it's just unpleasant. I mean, and actually...it probably stopped a Nobel... It probably stopped a Nobel Prize from being awarded. So he and another guy did the first crystal structure of a [molecule] that was [specific molecule] a real landmark. And he became convinced somehow that the other guy had stolen his idea. Now I knew actually both of them, and - and I'm sure the other guy didn't steal his idea. But [name 1] was convinced that [name 2] had stolen his idea. When [name 2] would come to the [institution] lab at [city] they would put a padlock on the model room where they were building the models so they couldn't sneak in - as if he was going to sneak in and steal their ideas! And so basically the field just became so muddied by this - accusations - that I think the Nobel committee said - forget it, we're going to stay away from it. But it was the first crystal structure of a [molecule]. You know, it could very well have been a Nobel prize. So, you know, that was an example where just this... It wasn't that either of them wasn't going to get

credit. They both - they were both really admired people. So, you see that and you think: oh no, it's better to be open and take a few lumps along the way. But on the other hand, you know, I have to be careful, I can't go out and talk about... If my postdoc is working on something and she's discovered that this protein kinase is involved somehow in regulating cell division, and there's seventeen thousand papers on this protein kinase and nobody, none of them are about cell division because people haven't recognised this, I'm not going to go to one of these big huge labs with 30 or 40 people in it and talk to them in detail about what she's doing, before her paper is in press, because they could flatten us like a bug. So, as a lab head you have to be careful to protect the people who work for you, so you do have to sort of - there are some lines you have to tread, the things that you can say and things that you can't say. But it's better to be open.
Mm, mm-hm. Do you think there's any aspect of, like - if you think about this, do you think there's a moral question about openness or is it mainly practical, or...some other way of deciding about what to do?
Well I - the whole way I live my life with all my interactions with people is that I try to...I want to feel that I'm morally right. You know. Treat people...treat people right, and... But people don't see that, because they don't always necessarily recognise that being treated right can sometimes mean being treated quite harshly. As a scientist, if I think that you are following the wrong path in your work, it is my obligation to tell you that, even if it hurts your feelings. How can it be kind to let you go in what I think is the wrong direction? Sometimes [laughs] somebody can be quite critical of you, and you think they're being really mean, but you don't realise that actually they're bending over backwards to be as nice as they can. I mean, you know, you have to be...if you're a scientist, you have to be honest. But you know, in my experience, people are emotional animals and they do not always appreciate the honesty...
Yep. Okay, is there anything else you wanted to add that didn't come up along the way?
No.
Well thank you - very much.

Interview transcript: Jenny Graves

- Opted for the transcript to be publicly available with real name
- Recruited via personal connection; interview took place on Skype, November 2017
- Australian university context
- Transcript below partially anonymised as a default but contextual detail can be re-included in archived version, since she opted to be named.

So to start with, could you just tell me – (I know that this is...because I know something of what you do, but) - could you tell me for the purposes of the interview, briefly about the research you're currently involved in?
Well, I use comparative genomics to look at how sex - sex chromosomes and sex genes - evolved, and how they function.
Mm-hm. And as part of that, how would you define the discipline that you fit into - or disciplines?
It has become quite a cross-disciplinary thing. I mean it started off - molecular biology and cell biology, but of course it's gone right into genomics and particularly into evolution, and in fact right through to human genetics as well, so - you wouldn't believe the meetings I go to.
So all sorts of things. Is- Do you have a primary identity, that you'd put yourself as, among those?
I guess I call myself "animal genetics and genomics". It's a pretty big specialty! Is pretty descriptive. Or "comparative genomics" would probably cover most of my bases.
Okay, so would you describe yourself as a geneticist or genomicist...?
Well, I see the two as overlapping so much it's really hard - genomics is just a new technique that we geneticists use.
And is that - In terms of your career, how consistent has this identity been for you? Have you changed how you would call yourself?
I've always called myself a geneticist. I certainly, until recently, never called myself an evolutionary geneticist, but that's what I'm often called now.
Mm. And thinking about all of this - sort of in relation to the discipline definition, people are part of different scientific communities, the sort of communities you go to meetings in, or review papers in. Are there distinct - Is there a distinct community or communities that you're part of?

Yes very much so, in fact, I've become very aware of their differences as I've moved from mammals to reptiles. Because it's a very different community. The herp community is very different from, for instance, the mouse community. The mouse community is very focused on human health, genetics does both, but the herp community are all about guys who love to go catch snakes in the outback! And they're very friendly to each other. It struck me quite forcibly when we got, started to get back review comments on our grants - they were just so incredibly friendly, I've never seen anything like that! And people took me aside and said, "Well the herp community supports each other." And they have made very conscious decisions to - for the good of the whole, the whole discipline, to support each other. And that really has been quite a major factor. I wasn't aware there were so many differences, you know, I'm part of the human genetics community and that's very medical. And it's just another planet from say, conservation genetics which is another group that I belong to, the conservation geneticists, you know, they live on the smell of an oily rag [laughs]. It's quite funny sometimes when we get... I organise a conference which has human genetics and genetics societies together, and the human genetics society - the two days of that was at the [name of upmarket hotel] in [place], whereas the genetics society was in the unheated bunkers of the [discipline] department at the university! [laughs] And we stayed at [name] college - again no heating! [laughs] So a very very different set of expectations in these communities, and I think the reptile community is - is different again. So I was just quite astonished to find what tremendous differences there were. Not just in their expectations and funding, but also in their attitude to each other. The mouse community is known to be extremely nasty to other mouse geneticists, whereas the herp community is extremely nice!
Well actually exactly what I was going to ask about next, the character of those communities. Is there a certain community that you've felt like you've spent most time in, and that was formative, and the kinds of - the character of that community?
Well the genetics community, and genetics as defined in the 70s, 80s, and 90s when I was a young lecturer - again, it was not well funded, it was very far from being relevant to human health, that really only came to be a factor in the 80s. So before that it was quite a small community, generally rather supportive and friendly. Then there was quite a schism in the 80s when the human genetics society was formed - and I'm talking about [country] specifically. And I think there was quite a good deal of ill-feeling on both sides, I think the human genetics community felt a bit ostracised from the genetics community - you know, we would laugh at their papers sometimes because they were based on one patient! And we thought these guys just don't know how to do science - well of course human genetics has come a long, long, long way since then, it's much more aligned to mainstream genetics. So I think the boot is sometimes on the other foot, because the human genetics community is extremely well-funded compared with anybody else. So I very much regretted that there was that schism, because [country name]’s a small country, and both the societies have got about 300 members which is really too small, it'd be a lot better if we could permanently get our act together, but I think that's just not going to happen because the lifestyle is *so* different. [Upmarket hotel] versus the [discipline] department! [laughs]
Oh, that's really interesting! Yeah. Would you describe any of these communities as more competitive than the others?

I think human genetics community is *way* more competitive. And by human genetics, I would include in that model organisms like mouse. I think that the mouse community is extremely competitive. I - I did a little bit of work on mouse, and some work on human genetics as you probably know, but it was mostly in [country elsewhere] not in [home country]. But I had the experience, which was quite hilarious, whenever I put in grant proposals about [topic of expertise], I would get one referee who kept on saying, "Why [is] Professor [own last name] using [specific animal type]? Why not use transgenic mice like the rest of the world?" And I'd try and patiently explain that we could probably find out things that you can't find out via transgenic mice. And this happened for two years in a row, and it got a little bit personal, such that I was able to say by number, this referee has crossed the boundary, [they] shouldn't be a referee. But the next time it was the same referee, but this time [they were] saying, "It's alright for Professor [own last name], she can publish anything she likes - us poor mouse geneticists!" [laughs] "It's very competitive!" [laughs] So I could see where [they were] coming from, obviously [they'd] been battered down by the mouse community, and [they were] trying to sort of - include me in it.
That is extraordinary. Wow. Okay that's great - I'd like to ask you about openness now. So, just very broadly I'm wondering if you could tell me what first comes to mind when think about openness in science?
Well, it depends on whether you're talking about openness in publication or openness in science. So I see those as two somewhat different matters. I mean - for me, I've been on the edge, most of my life, I haven't been sort of in the mainstream human genetics community or the mainstream mouse genetics community, and I'm generally not at risk. So I'm probably much more likely to be open about what I'm doing, what I'm thinking, because I don't have the same sort of competition as human geneticists and mouse geneticists. We work on the [specific animal] - we're really not very concerned about other people rushing in our turf there. The same was always true of [specific animal 2] - you know, if I could find one person in the room who knew what end was what in a [specific animal 2] I'd be doing really well, so it was always easy me to be extremely open about what I was doing, at meetings - you know, no holds barred, I told my students and post-docs, you know, talk to anybody and everybody you possibly can because we need input and we also want them to know that what we're doing is actually relevant to what they're doing. So my experience is possibly not very typical but it has been quite a charmed life for me because it makes it very easy for me to say what I'm doing, to gather comments, and - and if other people are not being open about exactly what they're doing with their mice or their human it doesn't matter very much to me.
Mm, mm, okay that's really interesting. And what about the openness in publishing?
Well, I see and I really appreciate open publications that I can access - I mean that's just been a real revolution that I think has benefitted everyone. The fact that I can be sitting here writing my book, and think, oh, I better get back to that paper and find out - you know, exactly what the author says because maybe it's not what the reviewers say the author said. And I can get anything I like online, it's just absolutely fantastic. I love the fact that everybody can get at my papers too, and I certainly publish in open access journals wherever I can because I think it's very important that people have access immediately. I also have a lot of contacts, particularly in India, and in Thailand, and it's again been a game-

changer for them that they have access now to lots and lots of journals that they didn't have access to previously. So I see that there's been a tremendous opening up of the flow of information.
Mm, okay. [Dog barks] Hello dog!
[laughs] [comments about dog]
And I'm wondering, does openness feel like an issue that's relevant to you? Is it something you think about?
It's something I really appreciate, it's not something that concerns me a great deal. For the reasons I outlined before, I do feel I'm somewhat immune from the risks, but I certainly appreciate the benefits.
Mm, okay. And to perhaps attack the subject from a slightly different angle, particularly when you're talking about openness in science, could you describe to me what it is - what it looks like when science is open? Like...
Well, what it looks like is...a lot like many of the meetings I go to, where there really seem to be no reservations, that people will talk about what they're doing and what they're thinking, that there'll be discussions that are very real. I guess I'm always - that's my usual kind of meeting. I am part of a number of international bodies, [genome-related org] comes to mind, as well, where there clearly are constraints, and some people are very hesitant to talk about what they're doing, other people refer very directly to competition - but this has not been my experience at all.
Mm, okay. And you were saying that you - when talking to students, you advise them to talk to other people - have there ever been situations where you've though, I better advise the students not to talk about this, or anything like that?
I actually have made a policy that generosity pays. And for me and I think for our whole group, that has generally been the case, that we get much more by soliciting other peoples' opinion on what we actually care about and what we're actually doing, than we lose. We have been burnt, I think two or three times - you know, one student - no she was a postdoc who came from [country] and she did some nice work, I suggested that she talk to [scientist name] about it, which she did and I think that was extremely helpful to her. I thought that she was ready to publish, but she never got around to it. You know, three years later she was outraged to find that [scientist name] had done those experiments and published them. And I had to say, well I'm sorry but you can't sit around for three years and expect people not to take the next step. So I felt very sorry about that, but I don't think that the problem was actually talking to somebody, I think the problem was not acting, on - in a timely fashion, and just letting it sit there and somebody was sure to do those experiments. And I've also been burnt a couple of other times myself, and again you just have to say well, no I think we gain a lot more than we lose. So, once or twice where, you know - I gave a talk at [institution name] and I guess in retrospect I could see somebody there who was more than just interested, who ran in and did the experiments and published before we did. But no that's not been my general experience and I think we've gained hugely by my policy.

Mm. And this sort of "generosity pays" policy, is that something that you came to at a certain time in your career, that you decided that that was the way to approach things?
I don't know that I came to it, I think it was always part of my background, because - you know - in the 50s, in genetics, it was just not so competitive. I never had to - my first research was on the [hypothesis name], you know, [genetic phenomenon] in [specific animal], and I was only too grateful to find somebody who was actually interested in what I was doing. So I guess I was brought up thinking that, you know, it's a big wide world there, and if I can get anybody interested in my work that's wonderful, so competition never really entered into it until much later and by that stage I guess I'd already had positive experiences about the benefits of telling people what I was doing, and soliciting their advice.
Mm, yeah, okay.
It only became, sort of - a spoken policy much later when I think we observed that the world around us wasn't necessarily as generous with their sharing of ideas and data than, than we were. And so I really had to think about - well, is this good or am I putting my students at risk?
Mm, yeah. Okay. I want to ask you about sort of, some specific forms of openness now, we've already talked a bit about open access - I'm wondering if I could ask again: so you said that you like to publish open access, is that a sort of rule for you, or something that you tend to do, or what's your thinking?
It's not a rule but you know, the journals that are open access have become very much top of our priority list, and we do, we do publish still in some journals like [subject-specific research journal], that - you know, they're hobbling towards open access. The Springer journals just haven't quite got there yet, and that's a rather unique journal because there aren't very many good [genetic sub-field] journals, so [genetic sub-field] stuff, we do tend to send to either [specific journal] or [specific other journal]. But otherwise it's generally one of the PLOS journals or BMC journals. But of course, the main thing we're looking for is who is going to read this paper. We need to - the first thing I get my students or postdocs to do is tell me where the literature is mainly being published, because you want to be in that stream of literature. And more and more, in some of the open access journals. And we have a lot more choice now than we used to have, too, so we certainly exercise that choice.
Mm. Okay.
Starting [indistinct?] Nature [laughs].
[laughs] And, is open data, or sharing the primary data behind your results, is that something that happens in your field?
In genomics of course, it happens pretty much automatically, any sequence we get goes immediately open on the website, so - you know, we can't publish any sequence without it being accessible. And that's been policy in genomics for a long time, and of course it - it has enormously advanced the field, I don't think anybody would tell you otherwise. So we

appreciate that and of course all the genome data is immediately published. Everything that we've got on the [specific animal] is already online. So people are already, even though we haven't published the data yet, people are already accessing it - of course, most of the people who are accessing it are actually authors on [specific animal] genome paper [laughs]. And that would have been the same, probably, for the [previously studied animal] genome, the [other previously studied animal] genome, again I think the interests of the authors are dual - you know, they're interested in the genome paper but they - they're all there for a very specific reason, some are interested in blood proteins, and so forth. So they're probably at the forefront of the people who are actually using the public data.
Mm. And do you have other types of data that are the basis for your research as well, apart from genomic data?
Well our lab, as you well know, has rather specialised in physical mapping, so we have tonnes and tonnes of mapping data. And there's no obvious repository for that data. In fact the original data which of course is quite voluminous because it's all photographs - I mean, it'd love it if there were some way of archiving that data that means that we didn't have to handle it [laughs] but there doesn't seem to be that. [Name of collaborator] who I'm sure you'll remember has tried to get some sort of systematic archiving at least of map positions if not actual original photographs, and I think that's a very good thing, I'd like that to be available as well, but there - so far - is no agreement on what's the best way to archive that data.
Mm, okay. Is there - do you come under any pressure from sort of policies, institutions, journals, to archive that kind of thing that's quite hard to archive?
No, quite the reverse, nobody cares at all [laughs]. I mean it would be a real benefit to us to be able to offload some of our data because it is very voluminous, but nobody wants it. I certainly don't want to throw out the original data, but it is quite difficult to store.
Okay. And another - another potential form of openness is preprints. I don't know if you've come across these in biology - sort of, sharing of articles, possibly at the point of submission, in online repositories?
Yeah, I've come upon that more as a reviewer than as an author, I mean we just published a paper that was - I was quite shocked to find, you know, we'd submitted a revision, and the next day it was online. And I thought, well uh, yes, I guess that's good! [laughs] I would have thought that, you know, somebody had to have a careful look at it before it went online, but - no no, there he is! And we were satisfied with it, but it hadn't been through a second review process, which is what I always thought the revision was aimed to do. So I was a little bit surprised to find that. But - we were happy about it, because we were already getting comments on it, so that's got to be good.
Mm. So you don't - would your...
Not sure whether original papers that have not been submi-...even heard...that have just been submitted, I'm not sure how valuable it is to have those online. You know, sometimes a paper will go through several iterations, and sometimes a paper - the premise, the whole

premise will be challenged - so I guess openness is great but I don't know that you want to be doing the editor's work for them, and the reviewers' work for them, you know. It almost becomes a popularity contest rather than a serious scientific debate. So I see it as rather undermining the whole review process.
Mm. Okay. I'm working towards my last page of questions now. I think this may be covered a bit by what we've already said, but: I'm wondering, have you ever faced a situation in which you've wanted to share more scientific information than you did in the end? Or have you always felt that you're able to share as much as you wanted to?
I've always been on the rash side. A couple of times I have wondered whether perhaps I've compromised something of value to me or to one of my group. But very rarely.
Okay.
...The two instances that I already mentioned, the one talk I gave where I could *see* that somebody was, was acting very specifically, you know, questioning me very specifically - how did you do this and how did you do that and what did you look for - and in retrospect I could see that [they were] already planning to rush in and do the experiments and publish them before I got a chance to do it. But that was really the only occasion where I've, I've felt sorry that I was so forward with my suggestions. It wasn't data as much as ideas.
Right. And knowing about it now would you - do you think you would have acted any differently, if you knew what was happening?
Actually I don't think I would have [laughs]. Again, these are ideas I was floating, which at the time - this was about the [specific problem studying a gene] - and so I was talking about our attempts over ten years to [study this gene]. And my conclusion is [description of conclusion]. And my idea was [description of idea], and that we had already done that, but that was sort of taken up by a group of [discipline name] people who just went straight to the then unpublished [specific animal] data, and in fact they, they flouted the Fort - the uh, the Fort Lauderdale Agreement, that - you know - you wouldn't use unpublished data for something like this. So we were holding back from doing that work and they weren't. I think, you know, I got something out of that, eventually, I mean we were right, [conclusion], it's a very interesting story. And I'm sorry that, that my student wasn't the first one to demonstrate this, but I think - you know - the whole field moved forward.
Okay, yeah. And this may be implicit in some of the things you've said, but do you feel there's a consensus your discipline about how open science should be, in this way we've been talking about?
I certainly feel that very strongly in the...amongst the group that I largely work, do my research work with now, but that's the herp group. And I think it's always been very open and very friendly, and very open about their ideas. I'm not so much involved with the human genetics, but I think there still is a good deal of proprietorial information, and a good deal of fear that somebody else is onto it, I mean there's only one genome, and everybody's got the same information. So - again, it depends on which community you're talking about - the herp community loves it - YES! Go for it. You know, [journal name relevant to

community], everybody reads that online before it's published, and I think that's quite accepted now. I'm not sure that that's the same in some of the other, more medical communities.
And do - do you think things should be different from how they are, in any of these communities? Would you push the consensus in a certain direction, in some areas, if you could?
I'm not sure pushing is really required, I think there's sort of a rollercoaster there, and everything is rocketing towards very much more openness, it's going to be quite difficult to resist if you want to resist it. I think that's a good thing. I just hope we don't sort of stumble in our rush and end up with [an] incredible mess, and that's - I guess - why I'm a little bit anxious about submitted papers being available, because I can see the huge potential for mess when things are revised, and "oh, well, I read it in its first iteration, oh well and I read it in its 17th iteration". And there's somewhat different conclusions, who's right here? You know, it's just another dimension of - of data messiness - that I'm not sure the community needs.
Yeah, so when you say there's a rollercoaster and we're rushing towards certain things, what kind of openness is that - is it the sort of online, preprint openness?
Absolutely, I see that as being irresistible. I think it's taken much longer than I thought it would. I remember, [a society] meeting back - oh, it would have been 20 years ago, I think, [name] gave the talk about "what's going to happen in publishing". And she was predicting within a couple of years that there'd be complete turn - that, you know, we would now be, be paying for displaying our wares, rather than getting our data scientifically reviewed. And I think everybody was - it was a bit of a wakeup call - but then nothing happened for years and years and years. So it really surprised me that it was so long and some of the publishers were so very much dragging their feet. And still are.
Mm-hm. Yeah.
I think, you know, even the [region/subject-specific journal], which is one of the most conservative journals that I ever publish in, has seen the writing on the wall, and decided they at least have to have the option, of open - of open access. And I think [subject-specific journal] has done the same thing now, as well, so they have a sort of partial policy, but I don't think that even that is going to be tenable in the next few years.
Mm. Are you aware of your funders or your institution having policies about open access?
They've been very quiet on this score, and certainly - I'm a member of [number of] institutions as you probably know, and I don't know that *any* of them have policies, they're just sort of trying to run to keep up and figure out what's going on. So the only interest that most institutions have in publications is they want *more*. And they want their [national research assessment] scores... I mean it's all become very much scores, and impact factors and stuff like that, so bibliometrics. Whatever they say, gets into decisions about funding of institutions - so they don't care whether it's open or closed or whatever, they just want their score to be higher.

Mm. Okay.
So I don't see them has having done anything positive at all.
Mm. Okay.
[National research/funding councils] - I haven't seen anything, any policies from [particular council], there may well be, but I'm not aware of them.
Mm. Okay. And I just have one last question for you if that's okay?
Sure.
This is taking a step back from all of these things we've been discussing. I'm wondering if you could think about a scientist who you know, who you really admire, who you maybe have aspired to be like in the past. And if you could describe to me what sort of qualities they have, and what - what you think makes them a good scientist.
Well there's - I mean I've never really had a mentor, but there's a scientist I've worked very closely with for many many years, who is distinguished by things that I'd like to think that I am also. And that is a tremendous energy, and a tremendous ability to integrate data from many sources. I'm not in the same league at all, but he's really gone from population genetics, to conservation genetics, to human genetics - and he can see a really really big picture. And that's something I aspire to, is being able to integrate such a lot of observations of genes and genomes and chromosomes, into a big picture of evolution. I think he's - somewhat similar to me...he's been in the [institution] in the [large country], and so has had all sorts of restrictions that I haven't had to face. I mean basically they wanted him to work on [health field] and he wanted to work on [contrasting field], so [laughs] there was bound to be some conflict there. But he managed to do everything, and managed to end up with what I think is a, a truly grand picture of the evolution of animals.
So it's - is it a kind of breadth of interest, or a sort of intellectual capacity, or both of those things?
It's both of those things and it's not just breadth, it's also depth. You know, he's got extremely deep knowledge of the human genome, he's worked on the human genome, he's worked on retroviruses, he's worked - he really invented the whole field of conservation genetics, so he's able to apply what he knows, he deeply understands the populations that he works on, which might be [examples of specific animal populations] and all sorts of things. So he just has a tremendous capacity and I found it very inspirational that you're able to bring things together and actually make them simpler to understand.
Mm. Yeah. And thinking about all of that, do you think that has anything to do with openness by any measure that we've discussed today, or do you think that that's different?
I know that some of the things he's worked on, particularly [health condition] has - you know - that's exactly the situation that you most want to avoid, where there is really is real

competition, there's real competition for funding, and for results, it's all very high profile. So he's certainly been in a milieu that I haven't been in, and had pressures that I haven't had to put up with. His response has generally been to talk more, not less. And [publish?] more, not less. He's been very big in genomics, and one of the founders of comparative genomics, and a big advocate for getting all the data online, immediately.
Mm. Is that something that you think has influenced you at all, that attitude towards - or that you've grown with that in some way?
Very much - I mean matched very much my attitude anyway, and so I found it very reinforcing that he was able to inhabit very much larger spheres than me using the same set of guiding principles, you might say.
Mm, yeah. Okay, that's really interesting! Thank you so much [name].
Oh you're very welcome Ros, I can see that, you know, you're into some very deep questions there, and I guess you'll get such a lot of different answers [laughs].
Mm, yeah, no this is - I really love doing these interviews.
Well good luck in your research!

Appendix F: Approach to interview analysis

Figure 9 summarises my approach to interview analysis. My methods of interview analysis included close reading and commenting (**Figure 10**); mind mapping (**Figure 11**); "coding" transcripts in *NVivo* (**Figure 12**); listing, visualising, and writing about themes identified through these activities (**Figure 13**); tabulating interviewees' responses to certain themes once established (**Figure 14**); and writing summaries of interviews in relation to these established themes (**Figure 15**). These processes bore some relation to grounded theory's progression from expansive coding, to a refined coding framework, to higher conceptualisation. They also shared with grounded theory an emphasis on iteration: repeated comparisons between emerging themes and the data on which they are built. However, using multiple improvised methods in addition to *NVivo* coding helped me to explore the data at different scales – from phrases to narratives and whole accounts – and to express my emerging interpretations not only as categories, but as visual associations and as nuanced, tentative forms of written reasoning. This allowed a non-linear, creative process: higher forms of conceptualisation could occur both at early and late stages, and interact with detailed, grounded observations. I did not apply all methods to each interview – some, e.g. mind mapping, were too time-intensive and detailed to be applied in every case. However, early, expansive methods built up my understanding to a point approaching theoretical saturation (Bryman 2012:421) at which analysis of new interviews reinforced my primary observations. Although I began to solidify core analytical categories during this exploratory phase, my overall findings remained flexible into the writing up phase.

Organised methods of analysis were accompanied and shaped by tacit processes of conceptual reflection that were an outcome of sustained, in-depth engagement with the interview data and the overall topic. Sometimes I was able to document these reflections, and often they were prompted *by* documentation: transcription, organised analysis, conference preparation, and writing. Sometimes they were ephemeral, or occurred outside time allocated for analysis: non-research conversations, presentations, and readings. Where I was able to document, remember, or discuss these reflections with others, they often sparked further reflections and helped me to conceptualise the data. **Figure 16** shows examples of captured reflections.

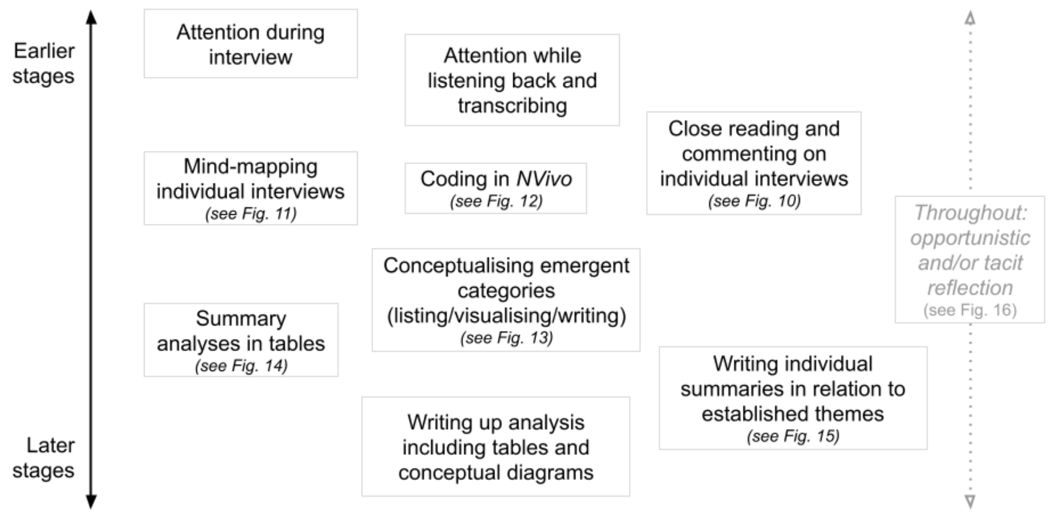


Figure 9 | Summary of my approach to analysis of interview data. Includes both organised and tacit methods, from early to late stages of analysis.

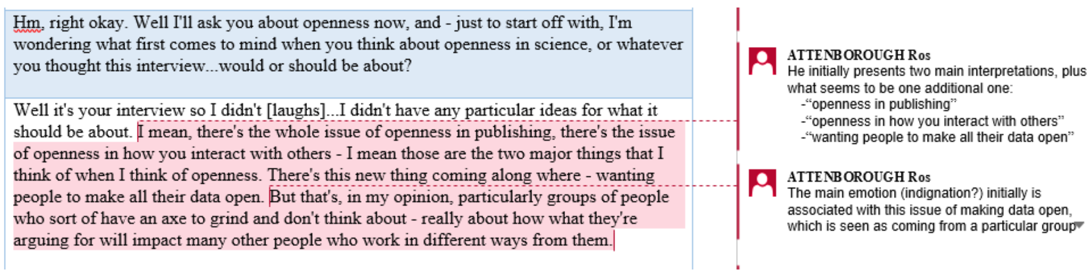


Figure 10 | Example of close reading and commenting. This method was useful for articulating and becoming sensitive to nuanced and tentative interpretations early in the process, but was too detailed to be applied in every case.

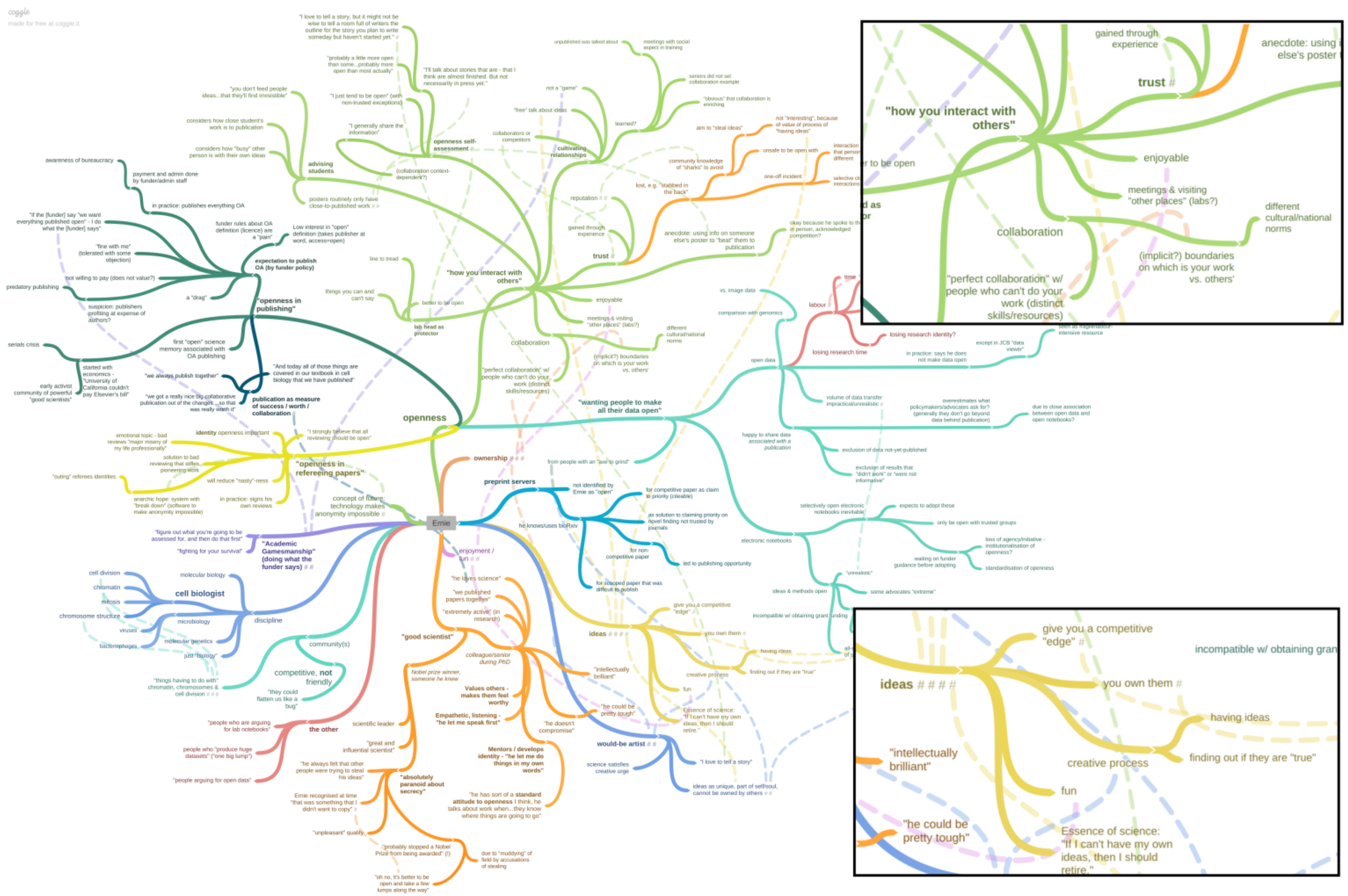


Figure 11 | Mind-map of interview with Ernie. One of several exploratory analytical methods; enlarged sections inset. See *Appendix G: Mind maps* for this and another mind-map in legible detail. Solid lines connect related themes, categories, and quotes; dotted lines cross-associate. Made in *Coggle.it*.

Figure 12 | Word cloud of codes categorising interview data in *NVivo*. Size and shade of each code indicates in approximately how often I applied it, both within and between interviews. Shows 70 of the most used codes among 270. Based on comprehensive coding of the first 16 scientist interviews. "Interactive disclosure" was my early phrasing for "interpersonal openness". This method was useful for identifying categories, but was limiting for capturing nuance and association. Made using *worditout.com*.

	As openness	Attitude or emotion	Closure	Consensus	On request	In relation to publication	Ownership or tie	Risk or negative consequences	Rules	Self-description	Stories	Trust	Value	Specific categories
Data practices	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	Context of data; Controlled access; Infrastructure & investment; Open but not open; Power; Relevance of openness; Stewardship & maintenance; Type of data; Using others' data; Volume of data; Sensitive data.
Interactive disclosure	↓	↓	↓	↓	↓	↓	↓	↓	-	↓	↓	↓	↓	Long-term influences; Advising others; Capacity, resources & security; Claiming priority & scooping; Cliques & closed circles; Collaboration; Freedom; Generosity; "Holding the cards"; Honesty; Personal attribute; Talking.
Open notebooks	-	↓	↓	-	-	↓	-	↓	-	↓	-	-	↓	Familiarity.
Open peer review	↓	↓	-	-	-	-	-	-	-	↓	-	-	-	Familiarity.
Preprints	-	↓	-	-	-	↓	-	↓	-	↓	-	-	↓	Familiarity; Physics.
Publishing practices	↓	↓	↓		-	-	-	-	↓	↓	↓	-	↓	Accessing publications; Copyright & licensing; Economics; Editorial, review & QC; Admin; Green & gold; Hybrid; Predatory publishing; Publishing

Figure 13 | Using codes created in *NVivo* to guide conceptual thinking. I mapped the association between codes representing openness practices (rows) and other codes (↓ indicates an association). This helped identify common and distinctive features across openness practices, in scientists' experiences. Full document in *Appendix H: Conceptualisation using NVivo codes*.

	1st response to openness includes:			Response includes spontanous mention of:			Shows good familiarity when brought up:		
	publishing/OA	*open data sharin*	*interpersonal ope*	*publishing/OA*	*open data sharin*	*interpersonal ope*	*publishing/OA as*	*open data sharin*	*interpersonal ope*
Adam	yes	no	yes	yes	no	yes	yes	yes	yes
Roger	yes	no	no	yes	no	no	yes	yes	sort of - opennes
Jude	yes	yes	no	yes	yes	no	yes	yes	yes
Gavin	yes	no	no	yes	yes	no	yes	yes	yes
Jacqueline	no	no	sort of (in negativ	no	no	sort of (in negativ	yes	yes	yes
Michelle	no	yes	yes	yes	yes	yes	yes	yes	yes
Mark	no	no	yes	yes	no	yes	yes	yes	yes
Ian	no	yes	yes	no	yes	yes	yes	yes	yes
Nicole	no	no	yes	no	no	yes	sort of (initial fam	no	yes
Thomas	no	policy docs ment	policy docs ment	yes ("a given")	yes	no	yes	yes	?
Ernie	yes	yes	yes	yes	yes	yes	yes	yes	yes
Neil	yes (thinks of pa	no	no	yes	yes	no	yes	yes	sort of (little talk
Yvonne	yes	yes	yes	yes	yes	yes	yes	yes	yes
Jenny	yes	no	yes	yes	no	yes	yes	yes	yes
Steve	no	no	yes	no	no	yes	yes	sort of (unclear o	yes

Figure 14 | Summary of scientists' accounts in relation to established openness categories. Records three measures of salience for each interviewee and category: mention of the category in first response; any spontaneous mention of the category; and familiarity when I raises the category. Useful for a comprehensive view across interviews, not for exploration/nuance.

#49 - Olivia ---

INTERPERSONAL OPENNESS IS PART OF HER VERY HOLISTIC FIRST RESPONSE. RELATES TO COLLABORATION. RECIPROCITY, BUT NOT GUARANTEED. FORMATIVE EXPRIENCE OF INTERPERSONAL COLLABORATIVE INTERACTION WITH COMPETITOR, WHICH WAS DISAPPROVED OF. DESCRIBES "OPENNESS IN SCIENCE" AS DAILY CONDUCT WITH GROUP, DEPARTMENT, INSTITUTION, FIELD. OPEN SCIENCE INSTEAD "THE PAPERS YOU PRODUCE". TALKS OPENLY ABOUT OWN RESEARCH EXCEPT ONE CONTRACT. SHE HAS SHARED INFO EVEN WHEN IT IS DISAPPROVED OF. SEEMS VERY CONFIDENT, EVEN ABLE TO NEGOTIATE/MANIPULATE SUPERVISORS. TRAINING HER OWN SUPERVISOR TO BE OPEN!

"I think another part of openness is, you know, being open to people's questions. Being open, not just within your group but to other groups. Don't be - I think being open, it's about collaboration. If you give…just, people have this sort of element of: oh, if you're not going to give me anything then I shouldn't give you anything and keeps this very closed off mentality within science. But actually if you go out there and say: oh you can have this, then quite often you get something back. Even in ways that you don't expect I think. Other doors open to you. That's a big part of openness."

"And actually presenting your work as well I think is part of openness. If you hide in the lab all day and hide at your desk all day and you don't get out there and talk about your research then no one can ask you questions so you can't really be open I guess."

Figure 15 | Late-stage analysis: summary-writing in relation to established categories. This note summarises Olivia's account in relation to interpersonal openness, and includes relevant quotes (more not shown). I recorded these summaries while writing up, to check my understanding with the data, and to ensure I had attended in detail to later interviews.

(a)

Interview thoughts

Very clear idea of openness as talking to others at conferences before publication

Seems clear that there is a kind of openness that is really about showing your wares and claiming your patch, even as it is simultaneously about creating collaborations and helping others

Also there seems to be a thing that is negotiated - there is an understanding that people are working on similar areas and are in competition. And that if you give them certain clues they can get there first. Seems almost as though there is a hidden code of conduct - that people often see situations where they could take advantage of info but they don't, and it's kind of about how into the area you are, etc....sort of like a good faith thing. Need to listen to those bits again.

(b)

- This is a very striking interview, especially for the strength of the expression and emotion that appears in it. I ended up coding large sections for "emotion" and especially "morality and moralistic language". [redacted] has strong beliefs about how scientists should behave, has clear red lines and is scathing about those who cross them. He also displays a positive passion for the purpose of his work beyond scientific goals. Among the scientists up to this point, he displays perhaps the clearest allegiance to openness as an virtue - though not necessarily an epistemic one? He is open "because it's the right thing to do" - he doesn't emphasise reciprocity like others. He's quite preoccupied by being portrayed as the "bad guy" in a situation where he felt tha the did the right thing.

The interviewee doesn't know what I mean by "openness" in science, and goes on to link openness with honesty, full and non-misleading disclosure of a scientific result (when it is communicated), trust, and "human" qualities. As I'm transcribing, I definitely feel though, that she is making this definition in opposition to some meaning of openness that is unsaid - these honest/trust issues seem to be the important thing to her, as opposed to something else undefined.

(c)

[interesting: her definition of open & about collaboration & sharing data in more personal way. & releasing at pub. not live, not licensing.]

(d)

research thought prompted by Zeba

Her old house metaphor for the patriarchy is very evocative, especially the idea of going back into that house and knowing it's rules, having the keys, being able to navigate, living but not complying.

[...]

Similarly, can scientists occupy the world that they feel restricts them in certain ways, makes them behave according to the "system" and its "currency"? Is there any way to participate in that system without reinforcing it, and still thriving? Are people doing that when they rebel? Is there a less risky way of doing it, a quiet disengagement that doesn't require being an anarchist, an activist?

Vulnerability is to shame as openness is to ...? (P. 64)
Is it a sense of worthiness allow scientists to share openly?
"Shame keeps us small, resentful, and afraid. In shame-prone cultures...administrators consciously or unconsciously encourage people to connect their self-worth to what they produce..."

Figure 16 | Examples of analytical reflections. Captured (a) immediately after an interview; (b) while transcribing and reviewing transcripts; (c) at a conference; (d) while reading unrelated texts that spark ideas: my friend Zeba's memoir (Talkhani 2019) and a book by shame researcher Brené Brown (2012).

Appendix G: Mind maps

Links to download detailed PDF of mind maps used to analyse some interviews. Alternative avenues will be explored for long-term sharing of these data in association with the thesis.

Ernie: https://www.dropbox.com/s/gm3bt0728dph62y/Ernie_mindmap.pdf?dl=0

Jenny: https://www.dropbox.com/s/umd9t52u1hb1qah/Jenny_mindmap.pdf?dl=0

Appendix H: Conceptualisation using *NVivo* codes

	As openness	Attitude or emotion	Closure	Consensus	On request	In relation to publication	Ownership or tie	Risk or negative consequences	Rules	Self-description	Stories	Trust	Value	Specific categories
Data practices	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	Context of data; Controlled access; Infrastructure & investment; Open but not open; Power; Relevance of openness; Stewardship & maintenance; Type of data; Using others' data; Volume of data; Sensitive data.
Interactive disclosure	↓	↓	↓	↓	↓	↓	↓	↓	-	↓	↓	↓	↓	Long-term influences; Advising others; Capacity, resources & security; Claiming priority & scooping; Cliques & closed circles; Collaboration; Freedom; Generosity; "Holding the cards"; Honesty; Personal attribute; Talking.
Open notebooks	-	↓	↓	-	-	↓	-	↓	-	↓	-	-	↓	Familiarity.
Open peer review	↓	↓	-	-	-	-	-	-	-	↓	-	-	-	Familiarity.
Preprints	-	↓	-	-	-	↓	-	↓	-	↓	-	-	↓	Familiarity; Physics.

Publishing practices	↓	↓	↓		-	-	-	-	↓	↓	↓	-	↓	Accessing publications; Copyright & licensing; Economics; Editorial, review & QC; Admin; Green & gold; Hybrid; Predatory publishing; Publishing decisions; Readership; ResearchGate; Role of publishers; Top journals & impact factor.
Rarely discussed types														
Digital interlinking	↓	↓	-	-	-	-	-	-	-	-	-	-	-	
In grant funding process	↓	↓	-	-	-	-	-	-	-	-	-	-	-	
Materials openness	↓	↓	↓	-	↓	-	-	-	↓	-	↓	-	-	
Methods openness	↓	↓	-	-	-	-	-	-	-	-	-	-	-	
Open code	↓	↓	-	↓	-	-	-	-	-	-	-	-	-	
Open source	↓	↓	-	-	↓	-	-	↓	-	↓	↓	-	-	
Personal frankness	↓	↓	-	-	-	-	-	-	-	-	-	-	-	
Process openness	↓	↓	-	-	-	-	-	-	-	-	-	-	-	

With research participants	↓	↓	-	-	-	-	-	-	↓	↓	-	↓	-	Rights and ownership.