Open science and codes of conduct on research integrity

Heidi Laine

University of Helsinki

heidi.laine@helsinki.fi

https://orcid.org/0000-0002-6658-0664

The purpose of this article is to examine the conceptual alignment between the ethical principles of research integrity and open science. Research integrity is represented in this study by four general codes of conduct on responsible conduct of research (RCR), three of them international in scope, and one national. A representative list of ethical principles associated with open science is compiled in order to create categories for assessing the content of the codes. According to the analysis, the current understanding of RCR is too focused on traditional publications and the so called FFP definition of research misconduct to fully support open science. The main gaps include recognising citizen science and societal outreach and supporting open collaboration both among the research community and beyond its traditional borders. Updates for both the content of CoCs as well as the processes of creating such guidelines are suggested.

Keywords: open science; research ethics; research integrity; science studies

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Open science[1] is one of the most ubiquitous research policy buzzwords of the 21ˢᵗ century. In Europe national open science policies are springing up like mushrooms after rain (Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, 2018; UNIFI, 2018; van Wezenbeek, Touwen, Versteeg, & van Wesenbeeck, 2017). The European Union is building a European Open Science Cloud (EOSC, 2018). Examples of initiatives outside Europe include the African Open Science Policy Platform, the USA-based Center for Open Science, and the OOO Canada network.

Open science, or aspects related to it, have sometimes been called a paradigm shift (Eisfeld-Reschke, Herb, & Wenzlaff, 2014; Salmi, 2015) and / or a revolution (Bartling & Friesike, 2014b; Dobbs, 2012), echoing Thomas Kuhn's ideas on scientific revolutions

---

[1]   For the purpose of this paper, science should be understood in its broadest definition, including all academic research.

(Kuhn, 1964). The urgency for more openness in and about science has been backed by a number of arguments, sometimes based on economic grounds, such as accelerating innovation, but just as often based on values. Openness and sharing is seen as the responsible thing to do, since most research is publicly funded, but also because transparency is considered to increase the reproducibility and reliability of results (Fecher & Friesike, 2014).

Professional ethics of researchers have been coded into guidelines on conduct at least since the post II World War Nuremberg Code on human experimentation. Research integrity (RI) emerged as a separate discussion from research ethics during the 1980s as a reaction to research misconduct scandals, and has since been codified into numerous codes of conduct (from now on referred to either as CoCs or codes), field specific, national, and international alike. In addition to detailing different forms of bad behaviour, these texts document the research community's views on what constitutes good and responsible behaviour in science (Löppönen, Mäkelä, & Paunio, 1991).

The purpose of this paper is to examine whether the current understanding of the part of scientists' professional conduct referred to as RI, and often also as responsible conduct of research (RCR), is in line with the ethical principles of open science. I will use four CoCs on RI as a proxy for RCR, asking, do these CoCs on RI direct researchers towards open science practices, or are they neutral or even negative towards open science? How do the codes approach the value of openness? Three of the codes are international and one national in scope. The chosen research method is concept analysis. A categorisation matrix used in the analysis is created based on samples of open science literature (due to the nature of the open science discussion consisting mainly of policy documents). Evaluating CoCs serves a dual purpose: it shows whether standing definitions of RCR and ethical open science principles match, while also allowing an examination of the supposed revolutionary nature of open science. Without going any deeper into Kuhn's theories (1964), CoCs can be seen as representing normal science, the status quo and the mainstream, against which the principles of open science can be reflected to see how big of a cultural shift open science actually is, and what are its most groundbreaking aspects.

Theoretical framework

Openness is widely considered to be one of the cornerstones of modern science. Historians of science date the principle all the way back to the 15th century and the dawn of scientific publishing (David, 2013). Robert K. Merton named communism, the common ownership of scientific goods, as one of the five norms of science (Merton, 1973). Bioethicist David B. Resnik defines the ethical principle of openness in science in his 1998 handbook on science and ethics as the responsibility of scientists to share data, results, methods, ideas, techniques and tools, as well as to allow other scientists to review their work and be open to criticism and new ideas. Resnik argues for the inclusion of openness to ethical principles both on moral grounds and to benefit scientific goals. To him, openness is the prerequisite for peer review. It prevents science from becoming dogmatic and biased, and

helps efficiency by fostering collaboration and the sharing of resources. He also argues that insofar as all people have a moral duty to help other people, and the sharing of data and resources constitutes a form of help, scientists have in addition to their scientific duties to be open, a general moral obligation to avoid secrecy (Resnik, 1998). Ruben Vicente-Saez and Clara Martinez-Fuentes present a contemporary definition of openness in science based on a systematic literature review: Open science is transparent and accessible knowledge that is shared and developed through collaborative networks (Vicente-Saez & Martinez-Fuentes, 2018). This is very general and thus does not exhaust the need for further definitions, but it can be used to recognise the key aspects of open science developments: (interlinked) knowledge, transparency, access, sharing and collaboration.

Ethical standards in science have two conceptual foundations: morality and science. Ethical conduct in science should not violate commonly accepted moral standards and it should promote the advancement of scientific goals. What makes a standard of conduct ethical rather than political is its application to individuals instead of groups. Ethical values and principles are connected as concepts: values are things that are evaluated as good or bad, either because they are seen as intrinsic or instrumental to the said good or bad, such as freedom, while ethical principles are normative directives that are derived from values, such as freedom from oppression as a human right. The situation is not clear-cut of course. Especially in the open science discussion the line between individual (ethical) and organisational (political) responsibility becomes blurred, and the concepts of value and principle can overlap and interchange. For instance, the ethical principles held by society at large can and should influence the values of science. Open science is one obvious example, another one would be gender equality. For the purpose of this article, I consider openness to be an instrumental value which serves the intrinsic values of communism in science (as defined by Merton, 1973) and helping other humans. Open access, reproducibility and verifiability of results, etc. are ethical principles derived from the aforementioned values (Elliott, 2017; Resnik, 1998).

Not all ethical issues necessitate standardisation and explicit guidance, but how can one recognise the ones that do? According to Joan E. Sieber, conflict can be one, if not the sole indicator: ethical issues can be identified as requiring codification when they cause clashes among values or interests. Sieber wrote, somewhat ahead of her time in 1994, that "matters such as informed consent, confidentiality of data, and data sharing involve ethics because they involve the perception of some form of good in science or society" (Sieber, 1994).

Since the emergence of RI as a separate concept, the term 'research ethics' has, confusingly enough, come to be an umbrella term signifying all good behaviour in science, as well as norms that apply specifically to researchers' interactions with research subjects. RI is the domain of researchers' professional ethics, similar to e.g. journalists' ethics. Breaches of research integrity, such as data fabrication, falsification, or plagiarism, are much less harmful than those of research ethics, since the former are usually victimless offences, hurting 'only' the robustness of scientific record and public trust in science. For this reason, research integrity is more autonomously defined and regulated by the community, while research ethics (again, a narrow definition) has closer links to legislation, at least in the

European tradition (Löppönen & Vuorio, 2013).

The purpose of professional ethics is to define what type of action is ethically appropriate in the conduct of the profession in question. The ethical sensitivity of a profession is determined by the number of independent choices and the power that professionals have in their activities. CoCs on professional ethics serve one or several of three purposes: 1) they are tools that help in ethical decision making, 2) they are a collection of principles that define a profession, and 3) they justify the existence of the profession as separate from other professions (Nieminen, Bargum, Karvonen, Väliverronen, & Ruuska, 2018; Räikkä, 2002). For example, in the Netherlands the national CoC on RI was initiated by Paul van der Heijden, former Rector Magnificus of the University of Amsterdam, who in a public speech suggested that Dutch universities should try to convince society at large of the worth of their efforts by making the principles of scientific conduct explicit in a commonly accepted, generic code of conduct for universities (Schuurbiers, Osseweijer, & Kinderlerer, 2009). But primarily CoCs in the scientific domain are normative research policies that are often located at the intersection of science, government and society (Rappert, 2007). CoCs on RI attempt to both describe and guide the way researchers behave as responsible professionals. Laetus O. K. Lategan has called codes, statements and declarations the software for academic integrity (Lategan, 2015).

Relevant previous research

In this chapter, I focus mainly on previous research literature on CoCs on RI and address open science research literature only briefly. I present open science-related literature, divided into sub-themes, more extensively in the chapter concerning content analysis results.

The research literature on open science is not abundant, unlike policy documents, reports, declarations and manifests. One of the most extensive treatises on open science is *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing* (Bartling & Friesike, 2014a). Its articles cover theory, history, case studies and technological as well as cultural aspects of open science. The most influential of this book's articles has been Benedikt Fecher and Sacha Friesike's introducing the five schools of thought on open science. The schools – infrastructure, public, measurement, democratic and pragmatic – describe different motivations and arguments behind open science development. The public and democratic schools are those with the strongest links to ethical valuations. The public school is defined in a way that is oriented towards citizens and their right to participate and be engaged, while the democratic school focuses on scientific knowledge as a resource that needs to be evenly distributed. (Fecher & Friesike, 2014)

Vicente-Saez and Martinez-Fuentes's (2018) systematic literature review represents a much more limited and mechanistic, albeit still valuable contribution to the field of research on open science. They searched the ISl Web of Science, Scopus, and a number of databases of such entities as the European Commission and the OECD for texts published

between 2006 and 2016 using open science as a search term, and came up with 1,220 hits. By including only English language publications in article format or policy-oriented studies, and eliminating overlap they narrowed the number down to 111, to which they later added 9 texts from between 1900 and 2005. The authors write that the abstract sift revealed the open science phenomenon to be imprecise, ambiguous, and not well-defined. According to them, authors mentioned open science without having a clear and common understanding of it, and that open science was used in various ways for different purposes (Vicente-Saez & Martinez-Fuentes, 2018).

One publication not included in the Vicente-Saez and Martinez-Fuentes's literature review is Whyte and Pryor's case study on open science in practice. It was probably not included because it is based on a conference paper, which was a basis for exclusion (Vicente-Saez & Martinez-Fuentes, 2018). I have found the article innovative and pioneering, and in addition to presenting empirical evidence on researcher experiences on open practices, interesting theorization is provided. Whyte and Pryor approach the concept of open science by creating a six-degree scale based on the level of research materials disclosure. The degrees from least to most open are: private management, collaborative sharing, peer exchange, transparent governance, community sharing, and public sharing (Whyte & Pryor, 2011). I have put this typology to use in my article concerning strategies against perceived risks of open science practices in two open collaboration research projects, in categorising and recognising the practice by practice level of openness in these projects (Laine, 2017).

Based on my search results on empirical research on CoCs I recognised two main trends. The first is mapping the landscape, with categorisations and groupings of CoCs (Godecharle, Nemery, & Dierickx, 2014; Kaiser, 1999; Resnik, Rasmussen, & Kissling, 2015; Vasconcelos et al., 2015). Research questions of this type tend to be variations of 'how many is', 'what is' and 'why is'. The second trend is examining researchers responses on and attitudes towards ethical issues and concepts (Giorgini et al., 2015; Löfström, 2012; Schuurbiers et al., 2009). There is also a genre of publications on whether a certain area or field should create a code of its own (Mwaka, 2017; Rossouw, Zyl, & Pope, 2014), as well as what could be called accompanying literature to published CoCs (Drenth, 2006; Karjalainen, Launis, Pelkonen, & Pietarinen, 2002; Löppönen et al., 1991; Steneck, Anderson, Kleinert, & Mayer, 2015). I have not included the latter categories to this literature review, but I have made use of them as secondary sources on the CoCs included in this study.

Throughout the literature a concern over fragmentation, lack of interoperability and varying understandings of central terms can be sensed. Many of the authors of the 2000s described an experience of ethical code influx (Bullock & Panicker, 2003; Rappert, 2007; Schuurbiers et al., 2009). One researcher interviewed by Schuurbiers et al. even referred to a "diarrhoea of codes". More recent studies on content and definitions, especially on research misconduct, raise the lack of harmony among codes as a concern (Godecharle et al., 2014; Räsänen & Moore, 2016; Resnik et al., 2015). Whether the perception of shift of focus from one decade to another is accurate requires a more systematic analysis of the literature than I have been able to do in this article. However, this might be an interesting

prospect for a future study.

Putting an exact number on the amount of CoCs on RI in the world is near impossible, and I have found no record of anyone trying to do that. Godecharle et al. found forty-nine CoCs in nineteen European countries alone (Godecharle, Nemery, & Dierickx, 2013; Godecharle et al., 2014). In their 2003 study, Bullock and Panicker evaluated 46 scientific societies from North America for ethical codes (Bullock & Panicker, 2003). Resnik et al. studied the research misconduct policies of the top forty R&D funding countries for 2014. They found that twenty-two of forty countries (55%) had a national misconduct policy, four countries (18.2%) were in the process of developing a policy, and four (18.2%) had a national research ethics code but no misconduct policy (Resnik et al., 2015).

These exercises in mapping the landscape have produced various ways of categorising CoCs. For example, Bullock and Panicker grouped them first according to enforceability (enforceable / non-enforceable). As a result of their analysis on the 46 codes, they ended up grouping them into three categories based on their purpose: 1) aspirational, 2) describing the common practice, or 3) regulatory and punitive. Bullock and Panicker found all of the codes had the same basic function and values, namely encouraging good practice, i.e. honesty, competence, and benignness. The practical application of these shared values into a CoC varied to a great extent according to the discipline in question but also in line with the issuing society's role and mandate in the field, sometimes versus other actors with overlapping roles (Bullock & Panicker, 2003).

Rappert made a distinction between 1) codes of ethics which he saw as being aspirational, even idealistic, 2) codes of conduct which aim at guidance, and 3) codes of practice that prescribe certain behaviour and are enforceable. His article from 2007 reflected on the major trends and developments at the time in ethical codification among the sciences. He also distinguished between universal codes that are very general and issued top-down; scientific society codes that are directed at members of a specified group; and institutional and workplace codes that can be even more specific (Rappert, 2007).

In his article Rappert went on to propose a criterion for evaluating the usefulness of codes. The criteria can be summarised by two revealing questions. Rappert's focus was on biosecurity specifically, but rephrasing the questions slightly by omitting direct references to biosciences shows that they can be useful for evaluating a broader range of ethical codes:

- Have codes helped clarify matters of uncertainty or disagreement?
- Have codes contributed to the existing understanding of what constitutes appropriate governance of science? (Rappert, 2007)

Godecharle et al. proposed yet another way of classifying CoCs, namely according to their focus either on norms or on values, which are then linked to two distinctive views on trust. According to Godecharle et al., these differences are at the root of the heterogeneity, and the resulting confusion, of European CoCs on RI. They write: "Values are universal and guide people in what or how they ought to be. Values are translated into norms, which are embedded in a specific context: situation, time, and place. Norms are subject to change. […] This distinction can also be applied to the context of research. The value of verifiability,

for example, is translated in certain norms, which can contradict one another." The first of the two approaches on trust emphasises the integrity and responsibility of the researchers, which resembles the value-based approach and an internal system of control. The second approach, which the authors see as becoming more prevalent, focuses on trust in control systems, linking it to norm-based approaches and external control systems (Godecharle et al., 2013).

Concerning researchers' attitudes and perceptions on RI and CoCs, Schuurbiers, Osseweijer and Kinderler published in 2009 a case study on the implementation of the Dutch CoC on RI, released originally in 2005 (Schuurbiers et al., 2009). Their research, which is based on researcher interviews, reveals the challenges of implementing a CoC in practice. The responses show that researchers can experience a gap between what is responsible in principal and what is practical. The interviewed researchers also felt that for all principles it was not sufficiently clear what a suitable practical application would be. In their conclusions, Schuurbiers et al. present what can be considered a negative checklist for effective CoCs: most scientists are not aware of its existence; there is no means of enforcement; it is too general to apply in practice; the moral dimensions of research decisions are separated from the practical context; and the division of responsibility comes across as unfair and unrealistic. They go on to suggest improvements both to the CoCs themselves and to the processes around them. The main challenge for CoCs is to remain close to the lived morality of researchers. The discussion known today as open science existed in 2009 and was recognised by the authors, just by different buzzwords: "This shift in practices goes by many names like Mode 2 science, post-normal science, or post-academic science […]. The point here is that neither scientific nor engineering codes of conduct quite cover the new types of research practices that have begun to emerge in the aforementioned areas of research." They conclude that attention needs to be paid to recent changes in the research context for codes to be effective (Schuurbiers et al., 2009). Whether the situation is different now, almost ten years on, is precisely the topic of this study at hand.

A 2015 study by Giorgini et al. yielded similar results to those by Schuurbiers et al. Giorgini et al. engaged 64 faculty members in an exercise on ethical decision making, after which they were interviewed. The written responses to the assignment were also analysed. According to the results, there were five primary themes regarding attitudes towards CoCs and other professional guidelines: 1) ignorance regarding guidelines, 2) use of ethical guidelines when a clear rule is available and consequences are substantial, 3) deference to professional norms over field-relevant ethical guidelines, 4) use of an internal compass in the absence of clear guidelines, and 5) thinking of guidelines as inflexible, idealistic, and representing a best-case scenario (Giorgini et al., 2015).

Most of the research on CoCs on RI, and on RI in general, seems to be policy oriented, with the primary motivation being the aim of creating better and more efficient regulation, instead of deepening an understanding of the phenomenon. This view is based on the observation that a large portion of the authors end up making practical recommendations on how to develop CoCs. For example, Bullock and Panicker recommended in their conclusion that if the primary goals of a code are to educate and inform action, an effective

ethics code should be user-friendly, specific and instructive. The enforceability of a code (for it to have 'teeth') requires the inclusion of a process for handling misconduct (Bullock & Panicker, 2003). Rappert concluded his article by stating that experience to date (in 2007 and in the field of biosecurity) suggested that the potential of CoCs to achieve laudable objectives should not be taken as a given. To Rappert, the question was not so much whether codes are good or bad policy options, but what commitments, means, motivations, and strategies are ready to be dedicated to making them meaningful. He criticised evaluating the success of enforceable codes based on content and implementation, at the cost of the processes of devising codes. As a practical recommendation he suggested that the processes should be considered equally interesting, and that making them meaningful should be considered alongside their content (Rappert, 2007).

Research materials and methods

This article is part of a doctoral research in progress, with the overall object of studying how RCR is defined in the research community. This general topic is approached through two case examples, the first being the Finnish Advisory Board on Research Integrity (TENK) and the other a radical open science practice, which I call open collaboration, and some of its Finnish practitioners (Laine, 2017). The context provided by the work on the doctoral dissertation has set the criteria for inclusion of CoCs in this study. As the CoC Responsible Conduct of Research and Procedures for Handling Allegations of Misconduct in Finland - Guidelines of the Finnish Advisory Board on Research Integrity 2012 (from now on referred to as the Finnish code) by TENK is the primary object of my interest, the other codes for this study were chosen based on relevance from the Finnish point of view. The Finnish code names four other codes: The European Code of Conduct for Research Integrity (2011), the Singapore Statement on Research Integrity (Research Integrity, 2011), the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (International Committee of Medical Journal Editors, ICMJE), and the Code of Conduct and Best Practice Guidelines for Journal Editors (Committee on Publication Ethics, COPE 2011). I have decided to include two of the aforementioned codes in this research, the Singapore Statement and the European code, plus one additional code, the Montréal Statement on Research Integrity in Cross-Boundary Collaboration (WCRI, 2013). The European code has been updated since the publication of the Finnish code in 2013, so I will review the updated 2017 version. The two codes focusing on publication ethics, the ICMJE code and the COPE code, are left out on account of their limited scope. Montreal Statement on Research Integrity in Cross-Boundary Research Collaborations is included because it is presented as complementing the Singapore Statement. Both the Singapore and the Montreal statements have the same origin, as they have been produced at the World Conference on Research Integrity (WCRI), although at two separate conferences. Tables 1–2 presents some of the key features of the CoCs. The source of information has been the code document itself, unless otherwise specified.

This research belongs to the field of social science history and thus approaches the

Code of Conduct	Who is telling?	Where	When	What	Why
Singapore Statement on Research Integrity	First drafted by a three-person committee consisting of conference chairs, then discussed among participants of the 2nd World Conference on Research Integrity.	2nd World Conference on Research Integrity in Singapore. "The World Conferences on Research Integrity represent effort to provide guidance for promoting integrity in research throughout the world."	2010	"The principles and responsibilities set out in the Singapore Statement on Research Integrity represent the first international effort to encourage the development of unified policies, guidelines and codes of conduct, with the long-range goal of fostering greater integrity in research worldwide."	"Publication of the Singapore Statement on Research Integrity is intended to challenge governments, organizations and researchers to develop more comprehensive standards, codes and policies to promote research integrity both locally and on a global basis."
Montreal Statement on Research Integrity in Cross-Boundary Research Collaborations	A similar drafting process to the Singapore Statement.	3rd World Conference on Research Integrity in Montreal, Canada.	2013	"[…] the following responsibilities are particularly relevant to collaborating partners at the individual and institutional levels and fundamental to the integrity of collaborative research. Fostering the integrity of collaborative research is the responsibility of all individual and institutional partners."	"[S]tatement builds on the Singapore Statement, extending it to collaborative research that involves more than one institution, academic discipline or country, as well as projects that involve some combination of academic, industry and government researchers." (Kleinert, 2015)

Table 1: Codes of Conduct (part 1)

phenomena with the primary task of describing and understanding, rather than creating new theories. However social science history is a theory-oriented discipline, so some theorisation is involved, both in methods and analysis, but I see it more as a tool than an aim in itself.

The three codes of conduct were analysed using deductive content analysis (Elo & Kyngäs, 2008). Content analysis is a method that has been used successfully to analyse and compare CoCs on RI in previous studies (Godecharle et al., 2014). The analysis begins with making sense of the data (Table 1–2). The next step is to create a categorisation matrix based on the existing theory or framework, in this case open science, which is tested against the backdrop of the data.

Open science is a nebulous term that encompasses almost any dispute concerning the future of scientific knowledge creation and dissemination. It touches upon every aspect of the research process (Bartling & Friesike, 2014b). Despite noble efforts to come up with a clear, scientifically deduced definition for open science (Vicente-Saez & Martinez-Fuentes, 2018), the term will likely remain as fluid as any other attempt to coin a complex system of practices, values and ideologies in one term. In fact, there is no definitive universal understanding of what science means, let alone open science. Fortunately the research

Code of Conduct	Who is telling?	Where	When	What	Why
Responsible Conduct of Research and Procedures for Handling Allegations of Misconduct in Finland – Guidelines of the Finnish Advisory Board on Research Integrity 2012	The editorial staff included the Chair of TENK, the TENK Secretary General, two TENK members and a secretary. The top institutional level of the Finnish research community was consulted through written statements commenting on a draft version.	Finland. Finnish National Board on Research Integrity TENK was founded by a governmental decree issued in 1991 to address ethical questions relating to research and to the advancement of research ethics in Finland. TENK coordinates a voluntarily binding self-regulatory research institution level process for handling allegations of research misconduct.	2012–2013	"The RCR guidelines provide researchers with a model for the responsible conduct of research. The effectiveness of these guidelines is based on a voluntary commitment by the research community to adhere to them, and to increase awareness of the principles of research integrity. The RCR guidelines apply to all academic disciplines in Finland [...]. The objective of these guidelines is to promote the responsible conduct of research and to prevent misconduct in research in all organisations involved in research work, such as universities, research institutes and universities of applied sciences."	2012 was a celebratory year for TENK: the 10th anniversary for the code and the 20th anniversary of the board itself. Updates to the original code from 2002 were relatively minor.
The European Code of Conduct for Research Integrity. Revised Edition	One lead drafting author. ALLEA Permanent Working Group on Science and Ethics, the revision included consultation among 23 public and private stakeholders in European research.	ALLEA is an organisation for European academies of science and letters, including beyond the EU.	2017	"At the European level, the European Commission uses the Code as a reference document in the Horizon 2020 Model Grant Agreement. The Council of the European Union has also reaffirmed the importance of the ALLEA Code in its Council conclusions of 1 December 2015, which presented research integrity as an essential condition for achieving scientific excellence and socio-economic impact."	"The document is based on 'The European Code of Conduct for Research Integrity' developed in 2011 by All European Academies (ALLEA) and the European Science Foundation (ESF). [...] The current revision is motivated by developments in, among others: the European research funding and regulatory landscapes; institutional responsibilities; scientific communication; review procedures; open access publishing; the use of repositories; and the use of social media and citizen involvement in research." (ALLEA, 2017)

Table 2: Codes of Conduct (part 2)

task at hand does not require a singular definition of open science. Instead, what is needed is an exhaustive a list as possible of categories of ethical principles associated with open science (I have defined what an ethical principle means in the context of this article in the theoretical framework section). Since open science is a science policy discussion, it seemed only suitable to look for definitions of the different elements of open science from policy documents, such as declarations, recommendations, manifests and the like.

I began creating the categorisation matrix (Table 3) based on the EU project the FOSTERs taxonomy on open science (Pontika, Knoth, Cancellieri, & Pearce, 2015), to which I added the Citizen Science & Open Collaboration and Science Communication. Both developments exist separately from the open science movement, but they represent many of the ethical principles of open science, such as making science more accessible. They have also been associated with the open science discussion of many bodies, such as the European Commission (see "Draft European open science agenda," 2016) and the Finnish Ministry of Education and Culture [okfn_2016]. I have also omitted open science policies and open science tools from the matrix, even though they appear in the taxonomy, since I felt tools were sufficiently addressed in connection with reproducibility and open methodology, while policy did not seem like a relevant category from the point of view of the individual researcher's behaviour.

After establishing these general categories, I went on to recognise subcategories. Since the number of subcategories was initially large and many of them were very specific, I deduced broader categories based on them. Instead of just retracing my steps back to the elements of open science as they are phrased in science policy discussions, I tried to deduce the underlying ethical principles behind what some might call open science policy jargon. I followed up this step with another round of deduction, coming up with what I found was a practical number of reasonably general but still meaningful categories. The final categories might seem obvious, but I felt it important to follow through these steps, instead of relying on my superficial understanding of open science elements, and also to make the process transparent and minimise the bias created by my own preconceptions as an open science advocate and activist.

In addition to defining categories I also needed to describe their meaning, as to be able to compare the CoCs category by category with the ethical open science principles. These definitions are presented in detail in the following section. After creating the categorisation matrix with category descriptions, I read the full CoC texts several times while reflecting on their content against the matrix. I also did key word searches on the texts. For the Finnish code I searched using both Finnish and English versions of the equivalent terms to make sure that nothing vital was lost in translation.

Content analysis results

Principles of open science and the categorisation matrix

The categorisation matrix has six categories for grouping the content of the CoCs: publication, research data, research methods, evaluation, collaboration, and communication. These are the main elements of scientific research that are the object of open science discussion and action. Table 3 shows the correspondence between the categories and open science principles, as well as a summary of the content analysis results. To evaluate the "open scienceness" of CoCs on RI I have recognised in both academic and other professional literature, such as policy documents, commentaries, etc., the main elements of open science and the aspects of the scientific research processes they address. I have come up with open science definitions for the elements in order to compare them to the definitions that rise from the CoCs. I have not attempted to create definitive or exhaustive definitions, which has been deemed near impossible, even unwelcome, by many discussants (Fecher & Friesike, 2014). Instead, my aim has been to display the spectrum of open science issues. Next I will proceed with presenting definitions of open science principles, against which the CoCs were compared.

Open access

According to open access (OA) activist and author Peter Suber, OA is something that authors of scholarly works can offer their readers, since they are unencumbered by the motive of financial gain. Suber goes on to define OA as "digital, online, free of charge, and free of most copyright and licensing restrictions" (Suber, 2012). According to Suber's description, in its broadest sense OA means that content is both *liber* and *gratis* to the end user. Internet pioneer Richard Stallman has famously clarified the difference between *gratis* and *libre* with the analogy of free beer versus free speech (Williams, 2002). The former represents *gratis* and the latter *libre*. Sometimes it feels that in OA policy action the *gratis* aspect of OA receives more attention than the *libre*. For example, the different routes of OA, such as the gold OA and green OA models, only consider the point of free access: is it directly from the journal, as in gold OA, or through a parallel archiving repository, as in green OA (for more on OA terminology, see e.g. (Suber, 2012). *Libre*, as in free speech, means that the text is free to be reused through for example content mining, which means that it should be accessible in the original context of publication, with accessible citation metrics, etc. [the_hague_declaration_notitle_2014]. Based on these considerations, I describe the underlying ethical principle OA as *freedom and ease of reading and reusing texts revealing research results*.

Open data

The FAIR data principles created by the FORCE11 community have since their publication in 2016 become the go-to standard on open data. For example, the European Commission has adopted them as part of their Open Research Data Pilot in the Horizon

Categorisation matrix category	Open science principle main category	Open science principle sub-category	Singapore Statement	Montreal Statement	Finnish code	European code
Publication	Open Access	Access to research publications	v	v	v	v
		Reuse of research publications	(v)			
Research Data	Open Data	Access to and reuse of research data metadata				v
		Access to and reuse of research data	v	(v)	(v)	v
		Importance of research data as a research output				v
		Transparency of research data as evidence	(v)			v
Research Methods	Reproducible Science	Transparency and reproducibility of research methods	(v)	(v)	(v)	v
		Transparency and reproducibility of research tools	(v)			v
Evaluation	Open Evaluation	Transparency of research evaluations				
		Content-based evaluation				
		Transparent peer review				
Collaboration	Citizen Science & Open Collaboration	Access to research processes				
		Access to research infrastructure and tools				
		Shared and reciprocated benefits of research				
Communication	Science Communication	Scientific knowledge in universally understandable format	(v)		(v)	(v)
		Proactive and targeted societal outreach				

Table 3: Open Science principles corresponding to categorisation matrix categories

2020 funding scheme (European Commission, 2016). FAIR is an acronym that is derived from the words findable, accessible, interoperable and reusable (Wilkinson et al., 2016). What is important to notice when considering the FAIR principles, is that they are directed in equal measure at the metadata as the actual data itself. Metadata is the equivalent to a set of ingredients and nutritional values at the side of a box of ready-to-bake flour: it describes the product in a way that makes it not only understandable, but usable as well. Without metadata it is anyone's guess what exactly the white powder in the box is and how it will behave. Another set of influential data principles published by FORCE11 are the data citation principles (Crosas, 2013; Finnish Committee for Research Data, 2018). They describe why citing data is important. In the context of this article, the most relevant principles are importance, meaning the value given to data as a research output; credit and attribution, meaning that researchers need to be recognised for their data work; and

evidence, meaning that an argument that is made based on data is only credible if the data is presented for evaluation. The rest are more technical in nature, describing how data should be identified; how citations should be constructed in a way that offers direct access to the data source; about creating a track record for the data; how data citations need to be specific (data sets can be extremely broad); and that citation practices need to take into consideration discipline and community-specific needs. Based on the above-mentioned principles and recommendations, I summarise the ethical principles of open data as *access to and reuse of research data metadata, access to and reuse of research data, recognition of research data as an important research output, and transparency of research data as evidence.*

Open reproducibility

Open reproducible research can be understood as open methodology. With open methodology the researchers create a link between the presented research results and the processes that were followed and applied in order to produce these results. In practice, this means making available the steps taken in a specific experiment, presenting the supporting tools, such as the materials, code, instruments, etc. and the concluding data after conducting an experiment (Pontika et al., 2015). The calls for increased reproducibility have been especially urgent in the field of psychology, which has experienced a turmoil widely known as the reproducibility crisis. The crisis began, but was not limited to, the research misconduct scandal surrounding a Dutch social psychologist Diederik Stapel (Stroebe & Strack, 2014). There have since been many measures to tackle the crisis, such as a report on open data by the Dutch Academy KNAW (Royal Netherlands Academy of Arts and Sciences, 2013), and the Reproducibility Project by the US-based Center for Open Science (Open Science Collaboration, 2015), as well as the publication of the TOP guidelines encouraging pre-registration of research (Nosek et al., 2015). Open source (code) and open hardware movements in science aim at making research tools reproducible in a very literal sense, so that not just the results can be reproduced, but also the actual tools, such as algorithms, software and physical equipment (Baden et al., 2015; Powell, 2012). The ethical principle of open reproducibility thus calls for *full methodological transparency.*

Open science evaluation

Open science evaluation (OPS) refers both to the notion that the evaluation processes and the data used to inform the evaluations should be transparent, as well as to broadening the range of things that are considered to be of merit, such as data outputs. The San Francisco Declaration on Research Assessment is one influential document that voices these demands. Although it does not name open science, it stems from the same considerations as is broadly embraced in open science discussions ("DORA – san francisco declaration on research assessment," 2012). There is also discussion about rewarding researchers for open practices through funding and recruiting, but this is not so much an ethical principle as a transition phase science policy measure (O'Carroll et al., 2017). The ultimate aim of open

science is to make all science open, therefore in the end eliminating the need to create incentives on openness for openness sake. In addition to the content versus platform problematic that is at the core of DORA, the perceived opacity of peer review processes has received criticism from open science advocates. The open science peer review oath (Aleksic et al., 2015) calls among other things for researchers to approach peer review processes as transparent, supportive and collaborative dialogues among equals. The oath challenges the status quo of anonymity by encouraging reviewers to sign their reviews. Based on these considerations, I define the ethical principles of OPS as *transparency of research evaluations, content-based evaluation instead of reliance on proxy metrics, and transparent peer review.*

Citizen science & open collaboration

Citizen science (CS) as a practice and discussion predates, or depending on the chosen definitions and time-frames, at least parallels that of open science (see e.g. Irwin, 1995). CS overlaps open science in as far as it can be seen to make demands on increasing transparency, inclusivity, and participatory practices in scientific processes. Traditionally, CS has been considered mainly a way of crowdsourcing and outsourcing the gathering of observational data (Palacin-Silva et al., 2016), but many organizations especially those representing and researching with indigenous communities are demanding that the concept and practice of CS be broadened (Stevens et al., 2014). Open and Collaborative Science in Development Network (OCSDnet) published in 2017 a manifesto to promote inclusive open science for social and environmental well-being [open_and_collaborative_science_in_development_network_ocsdnet_open_2017]. In the manifesto, they criticise the open science movement for diverting from its original goal; adding inclusivity of science; focusing on issues that do not sufficiently challenge the established paradigm; and benefiting solely the global north. The OCSDnets manifesto presents a set of seven principles for inclusive and collaborative science: science as a knowledge commons; cognitive justice through diverse understandings of knowledge; situated openness that addresses the questions of context, power and equality; every individual's right to research through participation; equitable collaboration and co-creation between professional researchers and non-professionals; inclusive and accessible infrastructures built on open-source technologies, and finally; knowledge as a pathway to sustainable development. Similar ideas are voiced e.g. in the Finnish Open Citizen Science recommendations (2016), the Sana CoC, and the 10 principles for CS by the European Citizen Science Association. Open collaboration (OC) is an emerging open science research method that resembles CS in its reliance on crowdsourcing participants and expertise. The main difference between CS and OC is the level of professionalism and expertise required to participate in OC projects, in which participation is not merit based, but due to the nature of the research questions requires in practice a certain level of academic experience and acquaintance with scientific work. Key features of the OC method include open co-authorship, remote online collaboration and immediate online sharing of all research outputs. Notable examples of OC include the Polymath Project (Smith et al., n.d.), the

Somus Project, the NMR Lipids Project (Laine, 2017), and ALD History (Ahvenniemi et al., 2016). Based on previous considerations I formulate the open science values of citizen science and open collaboration as *access to research processes, infrastructure and tools based on collaboration instead of institutional affiliation, and equally shared and reciprocated benefits of research among research participants and stakeholders.*

Open science communication

In the context of the open science discussion, science communication is considered to mean communicating science to non-expert audiences. Studies on popular conceptions on so-called grand challenges such as climate change show a big disparity between the views of the general public and the research community (see e.g. Funk & Rainie, 2015). For science communication to be successful, the minimum requirement is that the content is understandable, meaning avoiding technical jargon, and is proactively offered to the audience. Unfortunately, different publics are not flocking to researchers to hear about their latest ventures, so researchers need to be proactive and work at recognising the stakeholders who could benefit from their work (Brownell, Price, & Steinman, 2013). Popularising science is one aspect of science communication, but not the only one. It also includes researchers being available to journalists and the media in the form of interviews and consultations, and to decision makers and businesses to inform their decisions. Science communication recommendations by the Committee of Public Information in Finland state that science communications is important not only for disseminating research results, but also to broaden the public's understanding of the operating methods of science and research (Nieminen et al., 2018). From these considerations I have concluded that the ethical principles of open science communication mean *scientific knowledge offered for broad societal consumption in a universally understandable form, in a way that is proactive and targeted.*

Evaluation of the codes of conduct

Here I will present the detailed results of the content analysis. Table 2 summarises the appearance of open science principles in the CoCs. I have tried to be benevolent in my interpretations and readings, meaning that if I have felt that the text can in any reasonable way be interpreted to support the open science principle in question, I have marked it as a positive "hit" on the matrix. The amount of benevolence required is reflected in how I have marked the ticks in the matrices: a simple v without brackets means that I have felt the CoC supports the principle clearly enough, while a v with brackets means that arriving at the conclusion of support has required some stretching on my part. Before I go through the results category by category, a few words on the general appearance of the terms 'open' and 'transparency' in the CoCs.

Not surprisingly, the most recent of the codes, the European code, makes the most explicit commitment to openness as a general principle. The term is used six times in the text, in the contexts of rewards by institutions, publishing, data, and communication

among research collaboration partners. Transparency is even more prevalent with a total of nine mentions in the contexts of honesty as an underlying moral principle; the handling of alleged violations to RCR and communicating about the procedures for such processes; informing about conditions of data access; communicating with research partners; publishing in general; disseminating and publicising research results; and evaluating research by others.

The Finnish code uses the term 'open' only once, in connection with disseminating science, more specifically publishing and communicating. The term 'transparency' does not appear in the text.

The Singapore Statement includes the term 'open' once, in the context of sharing data and findings, which according to the code should happen not only openly but also promptly. However, both openness and promptness are secondary to establishing priority and ownership. Transparency does not appear in the text.

The Montreal Statement uses the term 'open' five times and 'transparency' twice. Openness is connected to communication among collaborating partners, dissemination of results, declaring funding sources, and discussion among partners about assumptions relating to practices. It is noted in the statement that openness of results can be limited by an agreement between the partners. Transparency is one of the sub-categories under the heading 'responsibilities in managing the collaboration'. It is linked to the way research is conducted and to disseminating results.

Publication

All of the codes focus extensively on publication, although none of them define explicitly what is meant by a publication. It feels safe to assume that the term is used in the traditional sense, excluding e.g. data outputs, videos, blogs, and publications popularising science. The European code defines publishing as something that should happen in an open, honest, transparent, and accurate manner. This also applies to what is called interpretations of science. The Finnish code states, rather complicatedly, that "When publishing the research results, the results are communicated in an open and responsible fashion that is intrinsic to the dissemination of scientific knowledge." This does not seem to refer directly to the actual publications, but instead to the way results are communicated *in* the publications. To better understand the intent of the text, I went back to the Finnish language version. My understanding based on the Finnish wording is that the values of openness and responsible communication are characteristic to scientific knowledge, which should guide action when publishing research results. The Singapore Statement does not directly advise on the process of publishing and the availability of publications. Sharing findings ("openly and promptly") is mentioned in a way that seems agnostic towards the mode of sharing. This choice of words could be deliberately liberal, since publications are mentioned elsewhere in the text. Reuse of shared findings is not mentioned, but the term sharing, rather than publishing, can be understood as allowing broader rights than just viewing.

Research data

The European code pays data management almost an equal amount of attention as publishing and is also in this sense the most advanced of the four CoCs. The code calls for data infrastructures that support multiple types and uses of research data. A call for data to be as open as possible, and as closed as necessary, is made, which can be translated to mean data being ideally open by default. FAIR data principles are encouraged (findability, accessibility, interoperability, reusability, see e.g. Wilkinson et al., 2016). The European code is the only one that makes any references to metadata (data about data, something that is crucial for reuse). Metadata is mentioned in the code in connection to research infrastructure and the kinds of content they should support. Metadata is also omnipresent in the FAIR data principles, so by supporting them, the code gives strong, albeit indirect, support to openness of metadata. The European code is also alone in recognising data as a first-class research output, requiring that "[r]esearchers, research institutions and organisations acknowledge data as legitimate and citable products of research".

The Finnish code does not take any direct stance on availability or research data. Just like in the publication category, the wording leaves room for interpretation. In the description of good conduct, only two mentions of data are made. Firstly, methods for acquiring data should conform to "scientific criteria" and be "ethically sustainable". Secondly, researchers should comply to "standards set for scientific knowledge" when recording the obtained data. These criteria and standards are not explicitly explained in the code. In the description of bad conduct, data is approached in a more practical way. Presenting false research data is mentioned as an offence, not only towards the research community and the research record, but also to decision makers, which could be interpreted as implying that data should be shared beyond the research community.

As mentioned above, the Singapore Statement advises on open and prompt sharing of all research findings, without making any distinctions between publications, data, or any other outputs, as long as priority and ownership have been established. Establishing priority and ownership are not necessarily in contradiction with open science as long as establishing priority is not used as an excuse for an indefinite embargo period, "just in case". In fact, establishing ownership, if it means licensing data, is in the spirit of FAIR data principles, since unlicensed data, even if it is freely available online, lacks reusability and interoperability. The Singapore Statement also recognises the importance of accurateness of research data for the research record: "When misconduct or other irresponsible research practice is confirmed, appropriate actions should be taken promptly, including correcting the research record." This could be read as a nod towards the importance of transparency of data as evidence.

The Montreal Statement advises against agreements that unduly and/or unnecessarily restrict dissemination of data.

Research methods

The European code explicitly supports the reproducibility of science, though not by mentioning the openness of methods. The code recognises the importance of proper research infrastructures for reproducibility, placing the responsibility for providing this on research institutions and organisations. Supporting reproducibility is made the researchers' responsibility in reporting their results.

The Finnish code does not acknowledge reproducibility or openness of methods as long as the above-mentioned "scientific criteria" and "standards set for scientific knowledge" cannot be interpreted as such. Still, in the description of bad behaviour, the code lists reporting research results and methods in a careless manner, resulting in misleading claims, as a disregard for the responsible conduct of research.

The Singapore Statement advices onresponsible methods and good record keeping but does not elaborate on the level of their openness: "[r]esearchers should employ appropriate research methods […] and report findings and interpretations fully and objectively", and "keep clear, accurate records of all research in ways that will allow verification and replication of their work by others".

The Montreal Statement encourages openness about methods among collaborating partners but makes no mention of making methods public beyond that sphere.

Evaluation

None of the codes mentions openness in the context of research evaluation methods and processes, be this for transparency of evaluation criteria, open metrics, or open peer review. Only the European code encourages a plurality of recognised research outputs in the evaluation: "Research institutions and organisations reward open and reproducible practices in the hiring and promotion of researchers."

What the codes do recognise as an openness-related issue is the need for transparency concerning possible conflicts of interest in evaluation. The Finnish code advises that "[r]esearchers refrain from all research-related evaluation and decision-making situations when there is reason to suspect a conflict of interest". The Singapore Statement for its part asks for the researcher to "disclose financial and other conflicts of interest that could compromise the trustworthiness of their work in research proposals, publications and public communications as well as in all review activities".

Citizen science & open collaboration

The codes do not recognise citizens' rights to participate in the research process, or even its possibility. None of them mention citizens, either as individuals, or as a stakeholder group. This is the case even with the European code, which names the European Association on Citizen Science (ECSA) as one of the consulted stakeholder representatives.

In the Finnish code, for example, the only stakeholder groups that are mentioned alongside the research community are decision makers and the general public. Both of these groups are mentioned in the context of research misconduct and the disregard

for responsible conduct, naming decision makers as possibly disaffected by false science claims. Misleading the general public is named an irresponsible practice, placing it in the least severe category of bad practice in a scale that from worst to least offensive goes: research misconduct, disregard for responsible conduct, and irresponsible practice. Societal interests in research are mentioned only in the context of research misconduct, stating that resolving misconduct is in the research community's and society's mutual interest. This dichotomy between the research community and society is also visible when the code discusses different roles that a researcher can have: "[…] researchers also need to comply with the practices listed above when working as teachers or instructors, when applying for research positions or for research funding, as well as when functioning as experts in their field both inside and outside the research community" (TENK 2013).

Non-researchers are taken into account only as research subjects that need protection. When the Singapore Statement mentions societal considerations, it does so only in the context of considering the possible harm that a particular research might cause: "Researchers and research institutions should recognise that they have an ethical obligation to weigh societal benefits against risks inherent in their work."

Open collaboration is not taken into account either. The demands made in all of the codes for equally shared responsibility among all authors when publishing research results is problematic for open collaboration. I will return to this point in the discussion part of the paper.

It is worth mentioning that the Montreal Statement, which focuses entirely on principles of collaboration, does not seem to recognise any type of open or crowdsourced collaboration.

Science communication

Communicating research beyond the research community receives only minor attention in the codes. None of the codes offers any positive advice or incentives on making science understandable, let alone appealing to audiences beyond the research community.

Even though the Finnish code and the Singapore Statement do not directly discourage science communication, they might as well, since it is only mentioned through warnings. According to the Singapore Statement "[r]esearchers should limit professional comments to their recognised expertise when engaged in public discussions about the application and importance of research findings and clearly distinguish professional comments from opinions based on personal views". The Finnish code names presenting deceptive or distorted information concerning one's own research results or the importance of those results as an irresponsible practice. The code demands that the same level of meticulousness, accuracy and whatever else is understood by "scientific criteria" and "standards set for scientific knowledge" that are applied to the research process proper, is also observed "when working as teachers or instructors, when applying for research positions or for research funding, as well as when functioning as experts in their field both inside and outside the research community. [Principles also] apply to teaching materials, written and

spoken statements, evaluations, CVs and publication lists, as well as to societal interaction in both printed and electronic publication channels, including the social media."

The European code, which does have a chapter on "publication and dissemination", focuses almost entirely on what seems to be quite traditional scientific publishing, and only as the last bullet point states, similarly to the Finnish code, that "[r]esearchers adhere to the same criteria as those detailed above whether they publish in a subscription journal, an open access journal or in any other alternative publication form".

The Montreal Statement also seems to be of a mind to restrict breadth of communication, since the following is the only passage that can be taken to address the issue: "Collaborating partners should come to an agreement on who has the authority to speak on behalf of the collaboration."

Discussion

I found none of the evaluated CoCs to be in blatant contradiction with the ethical principles of open science, but only the European code of conduct can be said to actively support and give guidance on open science. The fact that the European code recognises open science in a detailed way is not surprising, considering that the open science discussion was already in the science policy mainstream in Europe during its drafting, and was a high priority for the European Commission, which has since adopted the code for projects funded through its Horizon 2020 instrument. The remaining codes do recognise openness as a value but offer very little in terms of defining what it means or guidance on how to practise it.

The one aspect where even the European code falls short of a full recognition of open science is in crossing the traditional professional borders of the research community, i.e. citizen science, open collaboration and science communication. If we see CoCs as representing what Thomas Kuhn called "normal science", the paradigm that is overturned in a scientific revolution, this, as I suggested in the introduction, would be where we draw a line of demarcation between old and new paradigms.

The European code acknowledges the need for regular updates, and I would be surprised if at least citizen science and science communication did not appear more prominently in its next version. But should the CoCs address all of the aspects of open science presented in this paper? What is an appropriate level of detail and specificity for a CoC on RI? This is a very challenging question to answer definitively. I suppose some might argue, and not entirely unfoundedly, that many of the elements of open science presented in this paper are too specific to be included in a CoC as such. However, whatever the verdict is, the criteria of inclusion should be consistent. Currently, it is not clear what besides historical reasons justify excluding data management, citizen involvement and societal outreach, when publishing is deemed fundamental enough to be treated at length.

When considering what a CoC is trying to achieve and what issues it should or should not address, the two evaluative questions proposed by Rappert come in handy: has the code helped clarify matters of uncertainty or disagreement, and has the code contributed

to the existing understanding of what constitutes appropriate governance of science? (Rappert, 2007). Based on these questions the codes evaluated in this study, or more precisely their upcoming updates, should take stances on all of the open science issues presented here, since they are undoubtedly matters of uncertainty and disagreement and affect our understanding of the appropriate governance of science in a significant way.

If the ethical aspects of open science continue to be left out of RCR guidance and ponderings, the research community risks losses on both fronts: open science as well as RI/RCR. Open science is just as much about values and ethics as it is about technology. Most of all it is about the role of science in society. It is perhaps the most all-encompassing value discussion that the research community has ever known, and the research integrity angle and community of experts risks being side-lined, and the CoCs becoming even further detached from the day-to-day realities of scientific work if open science is not addressed more thoroughly than is currently the case.

But it is equally the open science community's responsibility to embrace the RI/RCR concerns and potential within open science. For a revolution, open science is at the moment gaining too little ground where it matters most, namely in the amount of research being made open. When looking at some of the key indicators on open science, such as the number of open access publications or researchers' data-sharing habits, the progress seems very modest, to say the least (Ala-Kyyny, Korhonen, & Roinila, 2017; European Commission, n.d.). Ethical concerns, 'doing the right thing', have been shown to motivate researchers into adopting open practices (Ali-Khan, Harris, & Gold, 2017; Laine, 2017), a finding that has yet to be fully recognised in discussions about researcher level incentives for open science.

On a more general level, the CoCs could benefit on one hand from less ambiguity and on the other hand from less rigid definitions. This sounds like a contradictory statement, but it is not. Both aims can be achieved at once. Codes could be formulated in a way that is not too heavily anchored in specific technologies and practices that risk becoming outdated or restrictive, such as journal articles, blind peer review or authorship criteria. There are more general terms available to address these issues, such as dissemination, evaluation and creatorship. For example, the examined CoCs currently indirectly define the practice of crowdsourced collaboration and open co-authorship as irresponsible by underlining the importance of the equal responsibility of all authors on the content of a research publication. Open collaboration is often conducted alongside other work as research funding models currently do not fully support open, multidisciplinary collaboration. In such cases authorship may be the only compensation for the effort. Offering authorship of an article is important in motivating participants to join a collaborative venture, and therefore it is beneficial to keep the authorship threshold relatively low. The group of participants may grow to become large in number and particularly in articles in the natural sciences, whose textual content may be limited, it is not appropriate or even possible for everyone to be involved in drafting the text of the article. Participants work in a network spread around the world and do not know each other personally. They do not have access to each other's research infrastructure to check measurement results, observations or samples, and thus cannot take full responsibility for the rectitude of each

other's actions. However, it can rightly be claimed that in these respects in which open co-authorship fails to comply with traditional authorship criteria to the letter, it does comply with its spirit and, if conducted with care, is even more responsible. The prevailing definitions of authorship are drawn up from the point of view of a closed research process. Their intention is to guarantee that no-one's input is left unrecognised and no-one receives merit that they did not earn. The temptation and opportunity for misconduct is minimised by a shared ethical responsibility. Openness makes it possible to achieve the above aims in a manner that is more transparently authenticated than in traditional publishing. In an open setting, the input of all authors can be verified by anyone and the public nature of the work means that the likelihood of being caught committing misconduct becomes so great that attempting it would hardly seem worthwhile (Laine, 2018).

At the same time avoiding too vague expressions, such as "standards set for scientific knowledge", as in the Finnish code would benefit the applicability of the codes. It can be argued that if standards were self-evident and self-explanatory there would be no need for CoCs. Also, as shown by Schuurbiers et al. (Schuurbiers et al., 2009), too idealised and general definitions of RCR can alienate researchers from CoCs and thus limit their impact.

It is evident that including a paradigm in the making, like open science, into a practice-oriented policy document on how to do research, is an enormous challenge. Therefore, I suggest a critical evaluation of the paradigm of CoCs themselves. They tend to be, at least in the case of CoCs included in this study, relatively static top-down (meaning authored mainly by senior researchers) descriptions of existing norms. One way of ensuring that open science gets recognised in CoCs would be to open the drafting processes to include more voices from the field, especially from the ranks of early career researchers. The digital environment could also in the future, if it does not do so already, provide solutions for defining and discussing RCR in a way that is dynamic, iterative and foresight oriented.

Bibliography

Ahvenniemi, E., Akbashev, A. R., Ali, S., Bechelany, M., Berdova, M., Boyadjiev, S., … Yurkevich, O. (2016). Review article: recommended reading list of early publications on atomic layer deposition—outcome of the "Virtual Project on the History of ALD". *Journal of Vacuum Science & Technology A*, *35*(1). https://doi.org/10.1116/1.4971389

Ala-Kyyny, J., Korhonen, T., & Roinila, M. (2017). Tutkimusdatan avaamisen esteet: Haastattelututkimus Helsingin yliopistossa. *Signum*, *49*(4), 25–29. https://doi.org/10.25033/sig.69198

Aleksic, J., Alexa, A., Attwood, T. K., Chue Hong, N., Dahlö, M., Davey, R., … Vieira, B. M. (2015). An open science peer review oath. *F1000Research*, *3*, 271. https://doi.org/10.12688/f1000research.5686.2

Ali-Khan, S. E., Harris, L. W., & Gold, E. R. (2017). Point of view: Motivating participation in open science by examining researcher incentives. *eLife*, *6*, e29319. https://doi.org/10.7554/eLife.29319

All European Academies ALLEA, E. S. F. E. &. (2011). The european code of conduct for research integrity. https://www.allea.org/wp-content/uploads/2015/07/Code_Conduct_ResearchIntegrity.pdf

Baden, T., Chagas, A. M., Gage, G., Marzullo, T., Prieto-Godino, L. L., & Euler, T. (2015). Open labware: 3-d printing your own lab equipment. *PLOS Biology*, *13*(3), e1002086. https://doi.org/10.1371/journal.pbio.1002086

Bartling, S., & Friesike, S. (eds.). (2014a). *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing*. Heidelberg: Springer Open. https://doi.org/10.1007/978-3-319-00026-8

Bartling, S., & Friesike, S. (2014b). Towards another scientific revolution. In S. Bartling & S. Friesike (eds.), *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing* (pp. 3–15). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-00026-8_1

Brownell, S. E., Price, J. V., & Steinman, L. (2013). Science communication to the general public: Why we need to teach undergraduate and graduate students this skill as part of their formal scientific training. *Journal of Undergraduate Neuroscience Education*, *12*(1), E6–E10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3852879/

Bullock, M., & Panicker, S. (2003). Ethics for all: Differences across scientific society codes. *Science and Engineering Ethics*, *9*(2), 159–170. https://doi.org/10.1007/s11948-003-0003-3

Crosas, M. (2013). Joint declaration of data citation principles. *FORCE11*. https://www.force11.org/datacitationprinciples

David, P. A. (2013). The historical origins of 'open science': An essay on patronage, reputation and common agency contracting in the scientific revolution. *Capitalism and Society*, *3*(2). https://doi.org/10.2202/1932-0213.1040

Dobbs, D. (2012). Is the open science revolution for real? *Wired*. https://www.wired.com/2012/02/is-the-open-science-revolution-for-real/

DORA – san francisco declaration on research assessment. (2012). https://sfdora.org/

Draft European open science agenda. (2016). https://ec.europa.eu/research/openscience/pdf/draft_european_open_science_agenda.pdf

Drenth, P. J. D. (2006). Responsible conduct in research. *Science and Engineering Ethics*, *12*(1), 13–21. https://link.springer.com/article/10.1007/PL00022265

Eisfeld-Reschke, J., Herb, U., & Wenzlaff, K. (2014). Research funding in open science. In S. Bartling & S. Friesike (eds.), *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing* (pp. 237–253). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-00026-8_16

Elliott, K. C. (2017). An introduction to values in science. In *A tapestry of values: An introduction to values in science* (pp. 1–18). New York, NY, USA: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190260804.003.0001

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

European Commission. (2016). *H2020 Programme Guidelines on FAIR Data Management in Horizon 2020. European Commission*. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

European Commission. (n.d.). Trends for open access to publications. *European Commission - European Commission*. Text. https://ec.europa.eu/info/research-and-innovation/strategy/goals-research-and-innovation-policy/open-science/open-science-monitor/trends-open-access-publications_en

Fecher, B., & Friesike, S. (2014). Open science: One term, five schools of thought. In S. Bartling & S. Friesike (eds.), *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing* (pp. 17–47). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-00026-8_2

Finland, O. K. (2016). Open citizen science. https://fi.okfn.org/projects/open-citizen-science/

Finnish Committee for Research Data. (2018). *Tracing data: Data citation roadmap for finland*. Finnish Committee for Research Data. http://urn.fi/URN:NBN:fi-fe201804106446

Funk, C., & Rainie, L. (2015). Public and scientists' views on science and society | pew research center. http://www.pewinternet.org/2015/01/29/public-and-scientists-views-on-science-and-society/

Giorgini, V., Mecca, J. T., Gibson, C., Medeiros, K., Mumford, M. D., Connelly, S., & Devenport, L. D. (2015). Researcher perceptions of ethical guidelines and codes of conduct. *Accountability in Research*, *22*(3), 123–138. https://doi.org/10.1080/08989621.2014.955607

Godecharle, S., Nemery, B., & Dierickx, K. (2013). Guidance on research integrity: No union in Europe. *The Lancet*, *381*(9872), 1097–1098. https://doi.org/10.1016/S0140-6736(13)60759-X

Godecharle, S., Nemery, B., & Dierickx, K. (2014). Heterogeneity in european research integrity guidance: Relying on values or norms? *Journal of Empirical Research on Human Research Ethics*, *9*(3), 79–90. `https://doi.org/10.1177/1556264614540594`

Irwin, A. (1995). *Citizen science : A study of people, expertise and sustainable development*. New York, NY, USA: Routledge.

Kaiser, M. (1999). Development of international guidelines for research ethics: A commentary on "scientific misconduct: Present problems and future trends" (b. Mishkin). *Science and Engineering Ethics*, *5*(2), 293–298. `https://doi.org/10.1007/s11948-999-0021-x`

Karjalainen, S., Launis, V., Pelkonen, R., & Pietarinen, J. (eds.). (2002). *Tutkijan eettiset valinnat*. Helsinki: Gaudeamus.

Kuhn, T. S. (1964). *The structure of scientific revolutions*. University of Chicago Press.

Laine, H. (2017). Afraid of scooping – case study on researcher strategies against fear of scooping in the context of open science. *Data Science Journal*, *16*(0), 29. `https://doi.org/10.5334/dsj-2017-029`

Laine, H. (2018). Open co-authorship of scientific articles. `https://doi.org/10.5281/zenodo.1302235`

Lategan, L. O. K. (2015). Software for academic integrity: The role of research codes, statements and declarations in research ethics and integrity. In *Integrity in the global research arena* (1st ed., pp. 165–176). Singapore: World Scientific Publishing. `https://doi.org/10.1142/9789814632393_0021`

Löfström, E. (2012). Students' ethical awareness and conceptions of research ethics. *Ethics & Behavior*, *22*(5), 349–361. `https://doi.org/10.1080/10508422.2012.679136`

Löppönen, P., Mäkelä, P. H., & Paunio, K. (eds.). (1991). *Tiede ja etiikka*. Porvoo: W. Söderström.

Löppönen, P., & Vuorio, E. (2013). Tutkimusetiikka suomessa 1980-luvulta tähän päivään. *Tieteessä Tapahtuu*, *31*(1), 3.

Merton, R. K. (1973). *The sociology of science : Theoretical and empirical investigations*. University of Chicago Press.

Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation. (2018). Le plan national pour la science ouverte. *Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation*. `m.enseignementsup-recherche.gouv.fr/cid132529/le-plan-national-pour-la-science-ouverte-les-resultats-de-la-recherche-scientifique-ouverts-a-tous-sans-entrave-sans-delai-sans-paiement.html`

Mwaka, E. S. (2017). Responsible conduct of research: Enhancing local opportunities. *African Health Sciences*, *17*(2), 584–590–590. `https://doi.org/10.4314/ahs.v17i2.36`

Nieminen, R., Bargum, K., Karvonen, E., Väliverronen, E., & Ruuska, M. (2018). *Bold communication, responsible influence. Science communication recommendations*. The Committee for Public Information. `https://www.tjnk.fi/sites/tjnk.fi/files/recom_scicommunication_2018.pdf`

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. `https://doi.org/10.1126/science.aab2374`

O'Carroll, C., Rentier, B., Cabello Valdes, C., Esposito, F., Kaunismaa, E., Maas, K., … Vandelvelde, K. (2017). *Evaluation of research careers fully acknowledging open science practices rewards, incentives and/or recognition for researchers practicing open science*. European Commission Directorate-General for Research and Innovation. `https://ec.europa.eu/research/openscience/pdf/os_rewards_wgreport_final.pdf`

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. `https://doi.org/10.1126/science.aac4716`

Palacin-Silva, M., Seffah, A., Heikkinen, K., Porras, J., Pyhälahti, T., Sucksdorff, Y., … Junttila, S. (2016). *State-of-the art study in citizen observatories: Technological trends, development challenges and research avenues*. Finnish Environment Institute. `http://hdl.handle.net/10138/164810`

Pontika, N., Knoth, P., Cancellieri, M., & Pearce, S. (2015). Fostering open science to research using a taxonomy and an eLearning portal. In *Proceedings of the 15th international conference on knowledge technologies and data-driven business* (pp. 11:1–11:8). New York, NY, USA: ACM Press. `https://doi.org/10.1145/2809563.2809571`

Powell, A. (2012). Democratizing production through open source knowledge: From open software to open hardware. *Media, Culture & Society*, *34*(6), 691–708. `https://doi.org/10.1177/0163443712449497`

Rappert, B. (2007). Codes of conduct and biological weapons: An in-process assessment. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, *5*(2), 145–154. https://doi.org/10.1089/bsp.2007.0003

Räikkä, J. (2002). Ammattietiikan merkitys. In S. Karjalainen, V. Launis, R. Pelkonen, & J. Pietarinen (eds.), *Tutkijan eettiset valinnat*. Helsinki: Gaudeamus.

Räsänen, L., & Moore, E. (2016). Critical evaluation of the guidelines of the finnish advisory board on research integrity and of their application. *Research Integrity and Peer Review*, *1*(1), 15. https://doi.org/10.1186/s41073-016-0020-9

Research Integrity, W. C. on. (2011). Singapore statement on research integrity. https://wcrif.org/guidance/singapore-statement

Resnik, D. B. (1998). *The ethics of science : An introduction*. Routledge.

Resnik, D. B., Rasmussen, L. M., & Kissling, G. E. (2015). An international study of research misconduct policies. *Accountability in Research*, *22*(5), 249–266. https://doi.org/10.1080/08989621.2014.958218

Rossouw, T. M., Zyl, C. van, & Pope, A. (2014). Responsible conduct of research: Global trends, local opportunities. *South African Journal of Science*, *110*(1/2), 1–6. https://doi.org/10.1590/sajs.2014/20130103

Royal Netherlands Academy of Arts and Sciences. (2013). *Responsible research data management and the prevention of scientific misconduct: Advisory report*. Amsterdam: Royal Netherlands Academy of Arts and Sciences. https://www.knaw.nl/en/actueel/publicaties/responsible-research-data-management-and-the-prevention-of-scientific-misconduct/@@download/pdf_file/20131009.pdf

Salmi, J. (2015). *Study on open science: Impact, implications and policy options*. European Commission Directorate-General for Research and Innovation Directorate RTD. https://ec.europa.eu/research/innovation-union/pdf/expert-groups/rise/study_on_open_science-impact_implications_and_policy_options-salmi_072015.pdf

Schuurbiers, D., Osseweijer, P., & Kinderlerer, J. (2009). Implementing the Netherlands code of conduct for scientific practice—a case study. *Science and Engineering Ethics*, *15*(2), 213–231. https://doi.org/10.1007/s11948-009-9114-9

Sieber, J. E. (1994). Scientists' responses to ethical issues in research. In W. R. Shadish & S. Fuller (eds.), *The social psychology of science*. New York: Guilford Press.

Smith, E., Gunashekar, S., Parks, S., Lichten, C., Lepetit, L., Morgan Jones, M., & Van Der Peijl, S. (n.d.). *Open science monitoring: impact case study – Polymath*. European Commission. https://ec.europa.eu/research/openscience/pdf/monitor/polymath_case_study.pdf

Steneck, N., Anderson, M., Kleinert, S., & Mayer, T. (eds.). (2015). *Integrity in the global research arena*. Word Scientific. https://doi.org/10.1142/9369

Stevens, M., Vitos, M., Altenbuchner, J., Conquest, G., Lewis, J., & Haklay, M. (2014). Taking participatory citizen science to extremes. *IEEE Pervasive Computing*, *13*(2), 20–29. https://doi.org/10.1109/MPRV.2014.37

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450

Suber, P. (2012). *Open access*. Cambridge, Mass.: MIT Press.

UNIFI. (2018). Avoin tiede ja data. Toimenpideohjelma suomalaiselle tiedeyhteisölle. Universities Finland UNIFI. http://urn.fi/URN:NBN:fi-fe2018052424593

van Wezenbeek, W. J. S. M., Touwen, H. J. J., Versteeg, A. M. C., & van Wesenbeeck, A. (2017). Nationaal plan open science. https://doi.org/10.4233/uuid:9e9fa82e-06c1-4d0d-9e20-5620259a6c65

Vasconcelos, S. M. R., Sorenson, M. M., Watanabe, E. H., Foguel, D., Palácios, M., Vasconcelos, S. M. R., … Palácios, M. (2015). Brazilian science and research integrity: Where are we? What next? *Anais Da Academia Brasileira de Ciências*, *87*(2), 1259–1269. https://doi.org/10.1590/0001-3765201520150165

Vicente-Saez, R., & Martinez-Fuentes, C. (2018). Open science now: A systematic literature review for an integrated definition. *Journal of Business Research*, *88*, 428–436. https://doi.org/10.1016/j.jbusres.2017.12.043

Whyte, A., & Pryor, G. (2011). Open science in practice: Researcher perspectives and participation. *International Journal of Digital Curation*, *6*(1). https://doi.org/10.2218/ijdc.v6i1.182

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., … Mons, B. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific Data*, 3. `https://doi.org/10.1038/sdata.2016.18`

Williams, S. (2002). *Free as in freedom: Richard Stallman's crusade for free software*. O'Reilly Media.