



Values in responsible research and innovation: from entities to practices

Marianne Boenink & Olya Kudina

To cite this article: Marianne Boenink & Olya Kudina (2020) Values in responsible research and innovation: from entities to practices, Journal of Responsible Innovation, 7:3, 450-470, DOI: [10.1080/23299460.2020.1806451](https://doi.org/10.1080/23299460.2020.1806451)

To link to this article: <https://doi.org/10.1080/23299460.2020.1806451>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 23 Aug 2020.



Submit your article to this journal [↗](#)



Article views: 1565



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 6 View citing articles [↗](#)

Values in responsible research and innovation: from entities to practices

Marianne Boenink ^{a*} and Olya Kudina ^b

^aDepartment of Philosophy, Faculty of BMS, University of Twente, Enschede, The Netherlands; ^bDepartment of Values, Technology and Innovation, Delft University of Technology, Delft, The Netherlands

ABSTRACT

This article explores the understanding of values in Responsible Research and Innovation (RRI). First, it analyses how two mainstream RRI approaches, the largely substantial one by Von Schomberg and the procedural one by Stilgoe and colleagues, identify and conceptualize values. We argue that by treating values as relatively stable entities, directly available for reflection, both fall into an ‘entity trap’. As a result, the hermeneutic work required to identify values is overlooked. We therefore seek to bolster a practice-based take on values, which approaches values as *the evolving results of valuing processes*. We highlight how this approach views values as *lived realities, interactive and dynamic*, discuss methodological implications for RRI, and explore potential limitations. Overall, the strength of this approach is that it enables RRI scholars and practitioners to better acknowledge the complexities involved in valuing.

ARTICLE HISTORY

Received 7 June 2019
Accepted 17 July 2020

KEYWORDS

Values; responsible research and innovation; pragmatism; value change; ethics of technology; entity trap

Introduction

Responsible (Research and) Innovation (RI and RRI) have become fashionable terms among actors concerned with the alignment of science and technology on the one hand and society on the other. As many have noted (e.g. Guston et al. 2014; Rip 2014; De Saille 2015; Van Lente, Swierstra, and Joly 2017; Genus and Stirling 2018), the emergence of both concepts was preceded by a variety of other labels and activities, e.g. different strands of technology assessment (Constructive, participatory, or real-time Technology Assessment), ethical, legal and social impacts/aspects research (ELSI/ELSA), value-sensitive design, and socio-technical integration research (STIR). The promotion and gradual uptake of ‘RI’ and ‘RRI’ can be understood as both an attempt to combine the earlier activities into a more encompassing approach to align science, technology and society, and as an attempt to overcome their weaknesses. Unavoidably, the RI and RRI labels themselves have been interpreted and taken up in a variety of ways since they were first introduced, partly informed by the professional history and the domain of the advocate

CONTACT Marianne Boenink  Marianne.Boenink@radboudumc.nl  Department of Philosophy, Faculty of BMS, University of Twente, P.O. Box 217, Enschede 7500 AE, The Netherlands

*Present address: IQ Healthcare, Radboud University Medical Centre, Nijmegen, The Netherlands

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

at hand. Recently, Owen and Pansera (2019) extensively discussed the similarities and differences between RI and RRI, suggesting that the first concept has been taken up in the academic field, whereas the second has gained more ground in the policy domain.

Notwithstanding the variety and differences, in this paper we will discuss RI and RRI as similar in an important respect (and hence from here on refer to all of them with the broader term RRI, regardless of how they label themselves). All these proposals are, after all, concerned with aligning *societal values* on the one hand with developments in science and technology on the other (e.g. Von Schomberg 2011; Stilgoe, Owen, and Macnaghten 2013; De Saille 2015; Wickson and Carew 2014). As such, this aim is not new; the novelty is in the way the alignment is supposed to be brought about. Two characteristics stand out here. First of all, rather than proposing values as an external point of reference for assessing and minimizing potentially undesirable results of innovation, RRI proponents argue that values should *feed into and co-shape* innovations right from the start (e.g. Taebi et al. 2014; Von Schomberg 2011; Van de Poel 2009). This also implies that values should be considered early on in the innovation process, when developments are supposed to be relatively malleable. Secondly, which values should go into innovation is *to be deliberated by those involved in and potentially affected by the innovation* (e.g. Stemerding 2015; Stilgoe, Owen, and Macnaghten 2013; Owen, Macnaghten, and Stilgoe 2012).

As some have pointed out, RRI thus proposes a shift in the ‘division of moral labor’ in innovation processes (Swierstra and Rip 2007; Shelley-Egan 2011; Rip 2014). In the past, politicians and citizens roughly defined the boundaries of what innovations to pursue, for example by prioritizing funding for specific research agendas. Within these bounds, scientists and technology developers were thought to be sufficiently accountable if they tried to reduce the main foreseeable risks of the innovation they developed. RRI, in contrast, implies that stakeholders should, sometimes collectively and explicitly, sometimes in more limited settings and implicitly, deliberate on and decide about the values that innovations should contribute to. In all these cases, this means that RRI comes with a great deal of ‘value work’: values have to be identified, deliberated upon, weighed, balanced or prioritized. We problematize such value work by revealing its complex nature in the case of value identification.

In view of ongoing attempts to align societal values with developments in science and technology, it is at least curious that there is hardly any reflection in the RRI literature on what values actually are. The review we conducted of articles published in this journal till the end of 2018 shows that most authors, while contesting how best to identify, collect and discuss values for RRI projects and processes, seem to take for granted that there are things called ‘values’, and that these are somehow ‘out there,’ accessible for anyone who makes an effort. They presuppose what we would like to call an ‘entity model’ of values. In this paper, we first of all argue (following Wittgenstein [1953] 2009) that language may lead us astray here. Contrary to what much of the RRI discourse suggests, people do not experience values as out there, readymade, waiting to be identified. They experience values as the evolving product of human valuing processes, and as such values are living phenomena, interactive and dynamic. Our second point is that this character of values should be accounted for when thinking about and (re-)shaping the ‘moral labor’ involved in RRI. In particular, the identification of values may require much more complex hermeneutic work than is often acknowledged. Although our work is reminiscent of critiques that

have been voiced in this journal before (e.g. Stermerding 2015; Capurro et al. 2015; Di Giulio et al. 2016; Grunwald 2016), our paper goes further in offering a theoretical reflection on the character of values and thinking through the methodological implications of these reflections. Overall, we suggest that in order to do justice to the moral hermeneutics involved in identifying values to guide innovation, a practice-based approach to RRI is warranted, since such an approach acknowledges the evolving, interactive character of values.

To substantiate our claims, we first discuss how values are represented in the two mainstream approaches of RRI, by Von Schomberg (2011) and by Stilgoe, Owen, and Macnaghten (2013). We show that, notwithstanding their differences, both risk falling into what we would call an ‘entity trap’: they tend to treat values as pre-given stable entities, ready-made for reflection. We briefly discuss that this way of thinking about values has strong historical roots and that similar approaches abound in contemporary applied ethics. We argue, however, that this way of thinking about values overlooks the hermeneutic work entailed in identifying values and that it may unnecessarily limit ways of dealing with ethical issues in RRI. Next, we discuss an alternative, practice-based take on values, which approaches values as *the evolving results of valuing processes*. We highlight how this approach views values as *lived realities, interactive and dynamic*, and consider how each of these characteristics has implications for the moral hermeneutic work required when identifying values for RRI purposes. We then discuss methodological implications for RRI that follow from endorsing a practice-based approach to values, and we explore some of its potential limitations. This allows us to conclude that a practice-based approach to values can complement existing RRI approaches by bringing into view the complexities of valuing, as well as the hermeneutic work involved in identifying values for RRI purposes.

Values in mainstream RRI approaches

A survey of the publications in this journal quickly shows that since the emergence of the concept Responsible Research and Innovation two definitions and approaches of RRI have become mainstream in the literature: most authors refer to either Von Schomberg (2011) and/or to Stilgoe, Owen, and Macnaghten (2013), Owen, Macnaghten, and Stilgoe (2012), Owen, Bessant, and Heintz (2013). Pellé (2016), when reviewing conceptions of responsibility in RRI, distinguishes a third approach, which she associates with virtue and care ethics (referring, among others, to Grinbaum and Groves 2013). The set of papers she refers to voices ideas about values and value identification that show affinity with what we label a ‘practice based approach’ of RRI. These publications in our view do not amount to a full-fledged RRI approach yet, since they provide little general guidance how to stimulate RRI in specific settings. They do point to an important yet underdeveloped theme, however, which we aim to develop into a third approach of RRI in this paper.

The two approaches of Von Schomberg and Stilgoe c.s. overlap to a certain extent, but there are also clear differences. We are particularly interested in how they conceptualize values. Since none of the authors explicitly defines values, the best entrance point is to ask how they suggest going about identifying values.

Von Schomberg (2011) defines RRI as follows:

RRI is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the ethical acceptability, sustainability and societal desirability of the innovation process and its marketable products (in order to allow a proper embedding of scientific and technological advances in our society). (9)

He first mentions interaction between stakeholders ('societal actors and innovators'), with transparency as the main value. Von Schomberg views this procedure, however, mainly as a means to ensure that both the innovation process and its products satisfy certain *substantial* criteria, listed in the second part of the definition: 'ethical acceptability,' 'sustainability' and 'societal desirability'. These rather abstract criteria, subsequently labeled 'normative anchor points', are operationalized in terms of the fundamental values of the EU charter on human rights and the Treaty of the European Union. Von Schomberg more specifically mentions safety, privacy, sustainability, quality of life and gender equality, among others (2011, 9–10). A similar list of values currently plays an important role in the way the EU uses the RRI approach in its funding policy (The European Parliament and the EU Council 2013; De Saille 2015).

In the approach to RRI developed by Von Schomberg, then, the values that should guide innovation processes and products are already known from the start. They can be identified by checking political documents or checklists based on such documents. The main challenge of RRI is to realize them in a balanced and fitting way in specific contexts. This is also where public debate and stakeholder deliberation come in: these may help to determine which trade-offs are justified in the context of an innovation, e.g. what an acceptable infringement of privacy is (2011, 13).

To be sure, Von Schomberg does not suggest that the values underpinning RRI are universal and eternal. Even though the values on his list appear to be objective, fundamental and obvious, they are region- and time-specific and acknowledged as such: these are 'European' values. On the one hand, RRI practices in Von Schomberg's definition derive moral legitimacy from fundamental values that are widely recognized and shared, and considered worth pursuing regardless of the preferences and likings of the public or particular stakeholders. On the other hand, these values are rooted in regional treaties and the political decisions underlying them, implying that their validity claim is particularly strong for citizens living in the relevant geographical and political domain, but also that they could, in principle, be changed in due time. Thus, by granting emphasis to external features in the constitution of values, such as geography, timing and context, Von Schomberg's RRI approach implies that values are simultaneously objective (or robust) and dynamic.

Stilgoe, Owen, and Macnaghten (2013), in contrast, propose an approach to value identification that is almost empty of substance. They define Responsible Innovation as 'taking care of the future through collective stewardship of science and innovation in the present' (1570), putting the notion of collective, forward-looking responsibility center-stage. Like Von Schomberg, they explain that RRI concerns deliberation about the process and the products of innovation, adding that it is also about its purpose. They propose that RRI has four dimensions (anticipation, inclusion, reflexivity, and responsiveness – more recently referred to as the AIRR framework), which enable 'raising, discussing and responding to' concerns about the purpose, process and products of innovation (2013, 1570).

Stilgoe and colleagues hardly refer to substantial values, except ‘care’. The AIRR framework consists of procedural values (like inclusion and reflection) that are in turn meant to access the repositories of substantive values: stakeholders. The main assumption about values seems to be that mobilizing the relevant stakeholders (including scientists and engineers) helps to identify the values at stake to be reflected on, discussed, balanced and decided upon as research and innovation move forward. Stilgoe and colleagues do not specify how stakeholders’ values should be mobilized; instead, they list some of the many ways to do so and focus on value deliberation rather than on value identification. They also do not discuss how stakeholders would actually know which values are important to them since this is beyond their primary task of articulating and justifying a framework.

Compared to Von Schomberg’s approach, then, Stilgoe and colleagues do not presuppose that the values that should guide innovation are given from the start. While Stilgoe and colleagues imply that stakeholders are an important source of information on values at stake (as well as on other, non-normative concerns), they do not prescribe how to tap this source. Our concern is that their silence on this point suggests that value identification is unproblematic and thus could invite RRI scholars and practitioners to assume that stakeholders who are invited to reflect on and discuss a proposed innovation will simply be able to state what their values are. A review of publications in this journal shows that authors applying the AIRR framework to specific cases of innovation actually use a variety of methods to identify stakeholder values, often in informal or even implicit ways. A number of these indeed seems to assume that asking stakeholders to voice values relevant to new and emerging technologies is unproblematic (for example Cohen, Stilgoe, and Cavoli 2018; Hartley et al. 2019), whereas others show clear awareness of the complexities and challenges involved (for example Stemerding 2015; Capurro et al. 2015; Repo and Matschoss 2019).

There are, then, clear differences between Von Schomberg’s and Stilgoe and colleagues’ views on the identification of values at stake in innovation. What is similar, however, is that both approaches invite a conceptualization of values as ready-made entities, at once knowable and available for deliberation. Whether they can be found in political documents or by interrogating stakeholders, a recurring assumption in Von Schomberg, which Stilgoe and colleagues do not explicitly dispel, is that values like ‘autonomy’, ‘privacy’, or ‘honesty’ are entities. In contrast to this view, and in line with the later Wittgenstein, we want to argue that language leads us astray here: referring to values with nouns seems to invite thinking about them as things, out there, waiting to be recognized. Considering how we actually know what values like autonomy or privacy mean, however, should make us aware that these concepts are derivatives: they are the product of the activity of valuing. By focusing on the outcome, rather than the process, both mainstream RRI approaches tend to reify values, thus falling into what we call an ‘entity trap’.

This observation touches on very fundamental and long-standing discussions in meta-ethics; RRI scholars are definitely not the only ones, nor the first to talk about values in this specific way. The ‘entity model’ of values is both ubiquitous and highly contested in many forms of applied ethics (see, for example, the debate on ‘principlism’ in biomedical ethics: Beauchamp and Childress 2013; Harris 2003; Arras and Brody 2013). It is also closely related to an ‘objectivist’ view of values. As any overview of axiology (value theory) testifies, this way of thinking about values has a long history (Schroeder 2016; Frankena

1967; Hart 1971; Partridge 1986). Plato, among many others, is typically interpreted as arguing that ‘value’ refers to a quality or property that exists independent from human evaluators, but also from the specific objects it is attached to. This view thus implies that something can be valuable even if this value is not recognized by any human being. It also sees values as stable phenomena: even when a value term (say ‘justice’) is used in many different settings, there is a core meaning, an essence, that transcends these situations. This objectivist take on values in axiology is contrasted with an ‘immanent’ view, which claims that values are always part of and experienced in the world experienced by humans. This view is often associated, but need not necessarily coincide with ‘subjectivism’: the position, defended by Nietzsche and others, that value is ultimately in the eye of the beholder. This latter position implies that there is no stable ground for value claims, and that the importance, but also the meaning of a concept like ‘justice’ can vary from one person and one situation to another. Subjectivism is thought to lead to moral relativism, since values do not seem to offer any stable guidance here.

This is, however, not the debate we want to engage with in this paper; our aim is different. What is at stake in the debates between transcendentalism and immanentism, and between objectivism and subjectivism, are metaphysical issues with severe implications for the justification of values. This is not the place to sort out these longstanding metaphysical and meta-ethical debates. Instead, we want to explore what a practice-based approach of values could look like and what benefits it would offer for RRI. When doing so we will, moreover, focus on issues of value *identification*, rather than value *justification*, remaining agnostic regarding the question whether a practice-based approach of values can avoid moral relativism.

Our starting point is that, whether values have an existence (and hence a validity) independent of humans or not, people usually access them by interacting with the world. This means that the core meaning of values is constantly being (re-)experienced and worked out within these practices. These activities of finding and/or giving meaning, however, are overlooked when thinking and talking about values as readymade entities. This has several drawbacks.

First of all, understanding values as entities *underestimates the work required to produce them*. It invites RRI scholars and practitioners to gloss over the intricacies of value identification, for example by simply referring to a pre-set list of values, or by assuming that the meaning of values brought to the fore by stakeholders is sufficiently transparent to all involved. In sum, the moral hermeneutics of identifying values is ignored. While in the remainder of this paper we will focus on this first drawback, there is another one that we want to at least briefly point out. Understanding values as entities also *unnecessarily limits the range of ways to address value conflicts*. Resolving value conflicts in the entity approach often boils down to prioritizing (with safety trumping privacy, for example) or balancing them (giving up some privacy to enable a bit more safety). In both cases the assumption is that the values involved remain qualitatively unaffected. If we approach values in a more dynamic way, resolving conflicts can also imply that values are reinterpreted in response to the context at hand, and more creative ways to address ethically challenging situations may become visible.

To be sure, when we say thinking of values as entities is a trap, we also recognize that it is a trap that is not easily avoided. We are very much aware that it would be challenging to substitute talking in terms of ‘values’ by talking in terms of ‘valuing’; as Dewey (1922b)

noted, ‘values’ are *evaluative devices* that we can hardly do without. Consequently, we sometimes use nouns ourselves in this paper. Nonetheless, our starting point is that this way of talking and thinking about values comes with several drawbacks, and we want to show how an alternative conceptualization of values might help RRI scholars and practitioners to avoid or at least minimize these drawbacks.

A practice-based approach to values

There are several sources and traditions that could help carve out such an alternative approach, like virtue ethics and ethics of care. We take our cue mainly from authors associated with pragmatism (in particular Dewey) and what today is often called ‘practice theory’, since they explicitly tend to approach ethical reflection as a practical, rather than a theoretical challenge. As Reckwitz notes, the practice approach ‘encourages us to regard the ethical problem as the question of creating and taking care of social routines, not as a question of the just, but of the ‘good’ life as it is expressed in certain body/understanding/things complexes’ (2002, 259). A practice-based approach to values, then, inquires how values manifest themselves in our ways of doing, rather than treating them as separate, self-contained phenomena. It stresses the valuing *process* rather than conceiving values as *entities*. This also implies, as we will discuss later, that reflection on values is (and should be) tied to doing (see Section on Implications).

A practice-based approach to values highlights at least three interrelated features: the *living*, *interactive*, and *dynamic* character of values. We discuss them subsequently, and for each of them ask what hermeneutic work is required to make valuing activities available for reflection. As we will show, each feature invites one to ask a number of questions that may guide ethical reflection in a way that remains true to the practice-based character of values, in the context of RRI and more generally. Together, these questions constitute a starting point for what we call a ‘moral hermeneutics’ that helps to understand what is normatively at stake, how and why, in a certain case or situation.

Values as lived realities

As Aristotle already realized, valuing takes place when an agent considers how a phenomenon contributes to his or her ends (Hondreich 2005). This suggests that what people call a ‘value’ is always relational: a phenomenon is valued as something for someone. As such, valuing mostly takes place in the context of action and provides orientation to action. It helps one to respond to a specific situation in one way rather than another. Valuing can, but need not be based on a reflective process. In some cases, an actor (or a set of actors) may engage in an explicit and extensive deliberation process, for example, to determine whether to introduce a new technology in a specific setting (e.g. AI to prognosticate patients in intensive care). However, this is the exception, rather than the rule. Most human actions are based on automated processes and routines, and in these cases valuing can be inferred only from the actual action. If we drive to work by car every day, we do not consciously decide again and again that ‘comfort’ is more important to us than ‘sustainability’; we rather continue a habit – a habit which may have started as a response to the lack of public transport options between home and work.

Dewey (1922a, 1929) understands human experience as constantly responding to its sociomaterial environment, and thinks of values as both a guide to and the product of (individual and/or collective) experience. Similar views have been put forward in the young field of ‘valuation studies,’ which approaches values as lively, complex matters of concern, interest and care (De La Bellacasa 2011; Dussauge, Helgesson, and Lee 2015). The products of these experiences and interactions are often conveniently summarized with nouns like honesty, justice, autonomy, or privacy. However, Dewey argues it does not make sense to detach such values from living experience, as if they exist in an independent (cognitive or social) realm. Such terms are a form of shorthand; they get meaning only in relation to specific circumstances and challenges. For that reason, Dewey does not talk about values or ends, but about ‘ends in view’ (1922a, 232).

To be clear, this does not imply that valuing is a fully subjective and/or situational activity. Valuing builds on past experience; according to Dewey values derive from ‘experimental empiricism’ (1930, 256–269). People consider future actions by (implicitly or explicitly) hypothesizing whether they lead to positive or negative experiences. Habits, routines, protocols or laws are initially informed by both collective and individual reflective judgment on how such a hypothesis will work out in real life circumstances. Such routines and regulations, however, can always fall short of expectations because the next situation may be subtly, but crucially different from what was known thus far. Dewey suggests that values based on past experience create a default position from which we tend to proceed, but that we always need to pay attention to how these abstract values may appear differently in concrete situations. Whereas we should understand the abstract value (the entity) as a product of many past empirical experiences, for such a value to give guidance in the current situation its meaning in relation to this situation should be specified.

This mutual relation between abstract value and concrete specification can be illustrated by the following example. Over the years, many patients and doctors have come to value patient autonomy in medical encounters, even to the extent that many countries have formulated laws requiring respect for autonomy. As a result, most doctors today approach their patients with a *prima facie* aim to respect their autonomy. However, this does not relieve doctors from wondering how best to respect the autonomy of Mr. Smith when he is visiting with a minor complaint for 3 weeks in a row now. Or to make them ask how to respect the autonomy of any patient when the new patient information system installed by the insurance company severely limits the options for treatment. Following Coeckelbergh, ‘We should not love wisdom in a Platonic way but act and find it in the world; there is no wisdom outside activity and experience. As Diogenes knew, we have *live* wisdom and *live* value. Value is neither to be described nor to be created; it has to be lived’ (2012, 199, original emphasis).

If we approach values as lived realities, which implications would this have for the moral labor involved in RRI, in particular for the identification of values? It should be clear now that taking values from collective codes or asking stakeholders which values they have can be a very first step at best. More hermeneutic work is required. We need to explore which values are actually embedded in the practices involved (or potentially involved) in RRI (see also Boenink, van Lente, and Moors 2016). This requires, on the one hand, that we ask how values that are codified or that participants come up with have been realized (and contextualized) in a specific practice. On the other hand, we

should explore whether these practices imply additional valuations that participants may not be aware of. These questions should then be extended to the future, asking how current ways of realizing value could be changed by introducing a new technology. While trying to answer this question, we should think not only of new priorities between values, but also of shifts in meaning in separate values.

Values as interactive

As mentioned above, Dewey conceives of valuing as an activity that is part and parcel of an individual's (or collective's) response to her environment. This implies that the resulting values are always interactive: they are shaped by the (perceived) environment and, in turn, shape that environment. This environment can be material, but also social or cultural. This way of thinking about values as interactive is highly relevant to RRI and should ring a bell with its scholars, because much work in Science & Technology Studies and philosophy of technology stresses that humans and their socio-technical environment are co-shaping each other (Bijker, Hughes, and Pinch 1989; Ihde 1990; Rosenberger and Verbeek 2015). While humans are producing technologies, these technologies change humans and societies: their roles, identities, skills – and also their values. As Pols (2014) writes:

Norms, values and strivings for the good cannot be separated from activities and ways of knowing practices. People mutually establish relations, between each other and with the technologies they use, and use words to frame these relations and the problems they hence want to solve. These activities of establishing and framing relations are oriented towards achieving something good. (176)

Dewey's pragmatist approach suggests that in moral deliberation we are (or should be) looking for the best fit between actions and situation. Regardless of whether one agrees with this particular criterion for moral justification, the concept of 'fit' raises relevant points to better understand existing practices. Questions to ask when considering potential innovations could be: to what elements of the environment is an existing practice a response? What triggered a certain way of doing and why has it been considered 'good'? This is particularly relevant for longer standing routines and procedures. A strategy that could be used here is to inquire about the 'exemplary situation' of certain ways of doing or regulations (Nauta 2000). For example, for what type of patient, with which characteristics and with which type of family- and patient-physician relations, did current regulation of euthanasia in the Netherlands create the ideal situation (Kennedy 2002)? Which variations in the situation do, however, cause uneasiness? In a similar vein, comparing related but different practices may also help to find out which values are realized in one practice but not in another. Such comparisons invite reflection on how practices could be improved (see Pols 2014 for a similar suggestion). Both strategies help to understand why a specific practice at some point was thought to fit the circumstances. Moreover, they help to anticipate how this environment-practice relationship might become less fitting.

Similar questions can be asked about future developments, for example in relation to emerging technologies. If we have a good grasp of the fit between the circumstances in which a certain practice seemed the best way to go, we might be able to surmise how the introduction of new technological possibilities in that practice could shift the

constellation. For example, which aspects of the situation and/or which routines might be de-stabilized by the introduction of machine learning-based algorithms to monitor bodily functioning? Under which circumstances (e.g. for what type of patients and doctors) would the use of continuous bodily monitoring combined with algorithms predicting health problems work out really well? And what circumstances or characteristics would lead to problems?

Whereas the hermeneutic task in the section above was about *describing* (or reconstructing) how values are embedded in a specific practice, the hermeneutic task involved here is to *explain* the fit between the (perceived) challenge and the (perceived) 'solution'. Whether such explanations are sufficient to justify the practice is of course something that would need to be scrutinized in later, deliberative stages of RRI projects.

Values as dynamic and changing

The preceding observations already imply that values, contrary to what the entity model suggests, are not experienced as static. What we think of as valuable shifts from one situation to another and evolves with human practices. Whether or not people will stick to certain values remains to be seen; but even if they stick to certain value terms, their meaning tends to change over time. We easily overestimate the stability of the meanings attached to 'autonomy' or 'privacy'. This was, again, put forward already by Dewey. For him, current morality consists of a set of values and norms that proved to be fitting in past situations. This is the reason why most of morality tends to go unnoticed: it has been routinized and has a largely tacit character. However, this should not make us forget that these routines have a specific, historical origin and have evolved over time. And, even more importantly in the context of RRI, they may further evolve in the future. Routines, Dewey (1922a, 1929) argues, originate in the problematic situations of the past. This also implies they should be open to change when the valuations implied are less capable of orienting action, because they do not fit the situation anymore.

This observation is even more relevant today, in the domain of RRI, because technology is one of the major factors destabilizing moral routines (Swierstra, Stemerding, and Boenink 2009, 120). This is also the reason new technologies often provoke reflection on morality: what was hardly discussed before suddenly grinds to a halt. This reflection may lead to more extensive ethical deliberation about what would be a good (or at least acceptable) way of dealing with the new possibilities, and how (if at all) to incorporate these in existing practices. Often, new ways of doing are developed, which themselves may become routinized over time. The use of Google Glass, for example, enabled people to (re)define the meaning of privacy in a way that fits the practices with this technology (Kudina and Verbeek 2019). The emergence of drones for military purposes led to new forms of moral experience and moral responsibility assigned to/felt by the soldiers (Elish 2017). Swierstra, Stemerding, and Boenink (2009; see also Boenink, Swierstra, and Stemerding 2010) label this continuous evolution (in mutual interaction) of morality with technology 'technomoral change', and others have voiced similar observations (Verbeek 2011; Vallor 2016).

The third and last hermeneutic task at stake, then, is to inquire into the *history* of valuation practices, asking how current routines actually evolved into their current shape. This not only highlights the dynamic character of values (e.g. the evolving meaning of autonomy or privacy), but also helps to explain how changes of meaning came about. In the

domain of RRI such an inquiry can guide the exploration and imagination of how future developments (of technology and other elements of the current constellation) could lead to destabilization and to renewed constellations.

Hermeneutic work involved in a practice-based approach to values

As discussed in the introduction, RRI is often understood as proposing a shift in the societal division of moral labor, even though it is not fully clear what such labor entails when it comes to value work. Following common practice in applied ethics (for example Bolt, Verweij, and van Delden 2003; Mepham 2005) we distinguish at least three steps in this moral labor. First of all, the values at stake need to be identified; secondly, there is some form of reflection or deliberation about which values should be realized, to what extent and how; and thirdly, decisions need to be taken about the best way to proceed. We already noted that mainstream approaches of RRI do not pay much attention to explicating the first step. This makes sense if one accepts the idea that values are entities and can be found in political documents and/or by asking stakeholders. A practice-based approach to values, however, explicitly brings into view that this first step already requires a large amount of hermeneutical work.

Based on the features above, we can now distinguish at least three hermeneutic tasks, all of which are primarily geared towards an understanding of the moral complexity of current practices challenged by innovation. The first is descriptive/reconstructive, the second explanatory and the third one is historical. Each of these can be guided by a number of questions (see Table 1). Since RRI is not only about responding to present, but in particular to future challenges, these questions should be directed towards the present as well as the future. In the latter case, the hermeneutic activity needs to be supplemented with imaginative activity. Ultimately, the understanding of current practice(s) produced by the hermeneutic work we suggest, should also provide RRI scholars and practitioners with a sound starting point for imagining how a practice might in the long term evolve in response to proposed innovations.

Table 1. Types of hermeneutic work implied by a practice-based approach to values.

Values as	Moral hermeneutic task involved	Guiding questions
Lived realities	Descriptive/reconstructive	Asking about the specificities of valuation: What 'values' are embedded in current ways of doing? What is the specific meaning in this context, compared to other contexts and to their abstract counterpart? How might technology affect these meanings (reinterpretation, shifts in weight/relevance)?
Interactive	Explanatory	Asking about the particular problem: To what challenges are (were) current routines a response? What makes them fit (or not) with the environment/situation? Which changes in doing and/or the environment/situation could an innovation bring about and how would these affect the current (lack of) fit between them?
Dynamic	Historical	Asking about the history of valuation: How did moral routines evolve to their current shape and meanings (and why)? What changes can be imagined?

Implications for RRI methods

What implications would a practice-based approach to valuing have for the methods used to identify values in projects aiming at RRI? It is not our purpose here to propose a detailed methodology. Nor do we want to claim that certain methods should be ruled out. After all, we endorse the practice-based approach as an additional, complementary approach to the two mainstream approaches, rather than as a replacement. We will say more about the potential weaknesses of this third approach in the next section. Here, the question is what methods would fit, provided one opts for a practice-based approach to values in RRI.

As should be clear from the argumentation above, a practice-based approach would not set much confidence by analyzing political documents or doing value preference surveys or other forms of asking stakeholders directly about their ‘values’. These bring up abstract values, pointing merely to the ‘summarized’ results of valuation, whether by stakeholders themselves or by experts. Moreover, they suggest that each value concept has only one meaning, whereas different parties involved in a political treaty or different respondents to a survey may entertain different interpretations. A practice-based approach to value identification, in contrast, puts the active realization of values in practices center stage. It studies and builds on ‘valuation in action’.

First of all, the conceptualization of values as *lived realities* provides a natural starting point for finding and studying them: in human practices and experiences (with technologies). Theoretically analyzing the relative importance of privacy in a networked society, for example, will not suffice to guide action when considering the introduction or use of a new app for monitoring one’s health. After all, (deciding about) action is always situated. To reflect on what privacy means in specific practices and how these meanings relate to other valuations in the same practice, we need case studies and methods to study these practices, first of all from within (Kudina and Verbeek 2019). Which values are made to matter in existing practices, and in the attempts to innovate these (Boenink, van Lente, and Moors 2016)? And how do the concerns voiced by actors express specific valuations (Van de Poel and Zwart 2010)?

Answering such questions is not an easy task, as there is no direct access to values in practices (Mol 2002; Dussauge, Helgesson, and Lee 2015). Understanding how values are manifest in practices requires qualitative, interpretative methods. RRI scholars can build here on existing philosophical methods with an explicit hermeneutic component (see Grunwald 2016; Austgard 2008; Fleming, Gaidys, and Robb 2003; Fagerström and Bergbom 2010). This is also where some earlier publications on RRI, voicing critiques similar to ours, may come in, since they often discuss examples of projects in which the authors tried to come up with alternative ways for identifying values. For example, Stemerding (2015) reflected on the RRI challenges of involving stakeholders in value identification in the context of synthetic biology, related to ‘create[ing] spaces and processes’ (141) for collective critical deliberation. Capurro et al. (2015) set out ‘to test an approach for extracting public values concerning a virtually unknown, scientifically and politically contentious topic, namely advanced lignocellulosic biofuel’ (247) and reflected on the difficulties of doing so, urging to adopt methods for self-reflection in this process (Capurro et al. 2015, 259). Di Giulio et al. (2016) used a mix of qualitative methods to show how local specificities and vulnerabilities need to be included in articulating the inclusion and responsiveness norms that guide RRI practices. Finally, Groves et al.

(2016) similarly argue for the necessity to make stakeholder deliberation in RRI ‘as inclusive as possible of a diverse range of societal perspectives on the values implicit in socio-technical arrangements’ (5) and suggest a psychosocial approach to practices that would enable rich, inclusive deliberations. Without using the term, all these publications point at the extensive moral hermeneutic work required when identifying values to realize RRI. Introducing the practice-based approach to values helps to systematize and theoretically substantiate such efforts.

Finding an appropriate method for the case at hand may also require transcending disciplinary boundaries and learning from other fields of inquiry. For instance, the method of Interpretative Phenomenological Analysis (Smith, Flowers, and Larkin 2009) originates in psychology, but can be adapted to uncover the moral hermeneutics involved in the RRI practices with new technologies (Kudina 2019). Whatever method is selected, the RRI scholar should refrain from direct translation of the participants’ narrative into values. Hermeneutic activity requires that one brings into view the interpretative framework informing the researcher as well as the one(s) guiding the participants in the practice studied.

Secondly, the methods used in a practice-based approach to RRI should acknowledge the *interactive* character of valuing and values. This means that they should never study ‘values’ in isolation, but always take into account how these are shaped by and themselves shape identities, roles, routines and materials. The hermeneutic activity thus is focused not only on mental and social processes that are or can be expressed in language, but also looks at the role of materiality. The interactive character of values also implies that the RRI researcher should be self-reflexively aware that values are always co-constructed by the methods used to identify them (see also Pols 2014). Qualitative researchers often accuse those using quantitative methods of ignoring this phenomenon, but surely qualitative methods like focus groups are co-shaping outcomes as well. This is all the more relevant when the research is about a new technology that most participants have not experienced, or may not even have heard of. Here again, several researchers in the RRI field have started developing methods that try to avoid at least the most obvious forms of co-shaping results, while at the same time acknowledging that no method will be fully neutral (e.g. Felt et al. 2014; Felt 2018; Boenink et al. 2018).

Finally, investigating values in practice is always a *dynamic* activity. There is no stable vantage point from which to identify values. Not only because the practice studied and the position of the analyst interact with each other, but also because both are subject to change. This does not mean, however, that a practice-based approach to identifying values precludes deliberating and arriving at normative conclusions later on. As indicated above, comparative and historical research can put specific practices in perspective and help to formulate conclusions as to what is or is not likely to improve a specific practice (e.g. Swierstra, Stemerding, and Boenink 2009; Kudina 2018). Instead of assuming a moral point of view that transcends the limitations of human experience, a practice-based approach to identifying values contributes to the active construction of a collective viewpoint that is as inclusive as possible within the context at hand. This enables deliberations and decisions about innovation in terms of better or worse. The evaluation of innovation as morally acceptable or even desirable is a product of interpretative and reflective human activity.

As this discussion shows, a practice-based RRI-methodology would differ from the substantial and most procedural approaches in several respects. First, it would also include

methods that go beyond language. Secondly, when using discursive methods it would particularly attend to the way invitations for stakeholder input and deliberation frame what is actually brought up.

Putting a practice-based approach to values in perspective

The practice-based approach to values is not without limitations. To put it in perspective, in this section we will first discuss three potential weaknesses: (1) the apparent bias towards the present rather than the future, (2) the somewhat reduced role of stakeholders vis-à-vis experts, and (3) its time-consuming character. Next, in line with our pragmatist starting points, we will briefly reflect on the type of RRI settings for which our approach might be more or less suitable. We conclude with an example showing how the three approaches of values can also be combined.

First of all, critics might respond that since our approach focuses so much on positioning values in practices, it is not feasible in many RRI activities. After all, when dealing with *emerging* technologies often no practices exist yet. The first thing to note here is that this is a challenge for any method trying to anticipate and shape future developments. The mainstream approaches of RRI discussed above must face this issue as well: for the substantive approach, it is not evident how long-cherished principles relate to new situations, and in the procedural approach stakeholders can have a hard time to respond to questions about their values in relation to emerging technologies they hardly know about. The practice-based approach, in contrast, offers three routes to reflect on values in relation to innovation. First of all, the exploration of existing practices of valuing is still important, since it provides a baseline to assess what might change if a new technology is introduced in this practice. More importantly, if new technologies are experimented with in existing practices (as is the case in clinical research, for example), this opens up a space of 'liminal innovation' (Mertens 2018). These liminal practices, where a technology is used for research purposes, enable the exploration of how technology-in-use impacts valuation and vice versa. Finally, emerging technologies, even while not realized or experimented with yet, tend to be present in society in the form of anticipatory discourse. As Grunwald (2016) has pointed out, it is important for RRI scholars not to take these visions for granted, but instead to strive for a hermeneutic understanding of these technological visions and the related expectations, promises and fears. We would add that studying the ways different stakeholders relate to or appropriate such visions, connecting them to their past and current experience, helps to reveal their value configurations, as well as to anticipate potential areas of moral concern. The hermeneutic methods suggested above can thus also be put to use for liminal practices and anticipatory activities, even though in these cases it may be more challenging to maintain their rigor and avoid uncontrolled speculation.

Another concern regarding the practice-based approach of values might be that it seems to reduce the role of stakeholders relative to those of RRI scholars. In contrast with the procedural and substantial approach the practice-based approach sees stakeholders as an important, but not necessarily as the best source for identifying what is at stake when innovating a specific practice. RRI researchers and facilitators (whatever their disciplinary background) play an important role in explicating and interpreting valuing processes. To put this difference in perspective, however, we should realize that

stakeholder input never gives ‘direct’ access to values: the setup of stakeholder involvement is *always* formative for its results, also in the more ‘direct’ procedural approach. We see it as a strength, rather than a weakness, of the practice-based approach that the role of these experts is explicitly attended to. That being said, a practice-based approach does not give experts the final say in identifying values; rather, it requires a repeated interaction between stakeholders and RRI experts, ideally striving for a ‘merging of horizons’, to borrow a phrase from Gadamer ([2004] 1975, 305). Finally, let us stress once more that what we say here is about identifying, not about justifying which values should guide innovation. So even if the role of RRI scholars and facilitators in the practice-based approach is larger than in alternative approaches, this does not imply anything specific for the division of roles at the stage of justification.

A final, more practical objection against a practice-based approach to values might be that it tends to be quite demanding and time-consuming, whereas the insights may be harder to use in policy making. Uncovering the dynamics of valuing processes in relation to new technologies is labor-intensive and, because of this, often tends to focus on a limited set of cases, practices and/or situations. This can be perceived both as a benefit, because such hermeneutic studies usually open up a broad variety of thick, experientially rich valuations, and as a drawback, because they tend to provide micro-level insights that are less generalizable and may be harder to translate into political decisions. These considerations may indeed sometimes limit its appeal and applicability in RRI. On the other hand, we should keep in mind that the alternatives have their own drawbacks.

To reiterate, the point of our paper is not that the practice-based approach of values should replace the other approaches, but that RRI scholars, facilitators, and practitioners need to make explicit choices, weighing the pros and cons of each approach in view of the starting position, the aims and the resources of the RRI project at hand. To give an example: If the proposed innovation is supposed to redress a situation that is widely considered to be morally problematic, a substantial approach may be crucial to point out where the current situation falls short and to identify minimum moral requirements. If, however, an innovation interferes in a practice that seems largely acceptable, although not perfect, the practice-based approach may be more appropriate than the other two, because it facilitates the identification and evaluation of subtler changes. In other words, whereas the substantial and the procedural approaches can facilitate the formulation of minimal moral requirements, the practice-based approach has more to contribute when wondering how to morally expand beyond a minimum level.

Ultimately, when considering values in RRI, the challenge may be to find a delicate balance among all three approaches. We believe that the balance is difficult, but possible. To substantiate this claim, let us briefly discuss how a recent European project on RRI in the context of Artificial Intelligence (AI), AI4People (Floridi et al. 2018), actually combines a practice-based with more substantial and procedural approaches to values.

Started in February 2018, the AI4People initiative aims ‘to create a common public space for laying out the founding principles, policies and practices on which to build a “good AI society”’ (Floridi et al. 2018, 5). This resulted in the publication of the ‘Ethical Framework for a Good AI Society’ (Floridi et al. 2018), which proposes the principles of beneficence, non-maleficence, autonomy, justice and explicability to guide responsible development and use of AI. These principles are further specified into twenty recommendations that can be implemented in specific practical settings. To be sure, its authors did

not explicitly endorse a combination of the substantive, procedural and practice-based approaches as we have delineated them. However, traces of all three can be observed in the resulting ethical framework.

Even though the affinity of the AI4People's ethical framework with the principles of biomedical ethics defined by Beauchamp and Childress (2013) may suggest otherwise, a *practice-based approach* to values was actually crucial for developing and formulating it. The five principles were identified by analyzing a broad spectrum of both current and anticipated AI practices, focusing on opportunities and risks. Moreover, the authors understand values as dynamic (Floridi et al. 2018, 22). The meaning of the at first sight well-known principles of beneficence, non-maleficence, autonomy and justice is not taken for granted, but actively interpreted to mirror the current and anticipated practices of AI development and use. Moreover, the authors also add a fifth principle, explicability, because of the novel challenges and opportunities accompanying the development of AI, 'that its workings are often invisible or unintelligible to all but (at best) the most expert observers' (Floridi et al. 2018, 20).

The point of the framework is not, as the authors point out, to produce value statements (Floridi et al. 2018, 15), but to outline best practices based on how AI is currently developed and used. The twenty final recommendations therefore detail concrete practices deemed desirable and ethical (Floridi et al. 2018, 22). The development of the ethical framework for AI by AI4People thus exemplifies the practice-based approach by identifying values in concrete living practices and by considering their dynamism in response to the changing social and technological environment.

The *procedural approach* at first sight may not seem to be very present in the AI4People's Ethical Framework. It was designed in consultation with twelve experts, representing civil society, academia, EU leadership, business and the media – so participation of stakeholders in developing the framework itself was limited. However, the concluding recommendations are rife with concepts like 'assessment', 'development' and 'incentivization' (Floridi et al. 2018, 22), implying that the framework promotes procedural values in addition to the substantial ones. The action points, moreover, encourage deliberation with stakeholders within specific practices. Finally, the authors explicitly state that the recommendations are not meant to be perennial: '[R]ecommendations should be seen as a "living document". The Action Points are designed to be dynamic, requiring not simply single policies or one-off investments, but rather, continuous, ongoing efforts for their effects to be sustained' (Floridi et al. 2018, 22). Thus, further deliberation may lead to novel interpretations and evaluations. Overall, then, the framework also exhibits characteristics of a procedural approach.

Finally, the *substantive approach* is also visible in the discussion of the five core ethical principles. AI4People explicitly puts forth 'a European approach' (Floridi et al. 2018, 22) to AI, promoting '... the development of AI technology in a way that secures people's trust, serves the public interest, and strengthens shared social responsibility' (Floridi et al. 2018, 22, original emphasis). In addition to suggesting that these 'normative anchor points' are characteristic for Europe, the recommendations make multiple references to European organizations and institutions deemed to represent best practices regarding other technological developments, after which AI practices could be modeled. Thus, even though the framework does not make direct references to specific documents, ideologically, culturally and normatively it presents itself as a European framework. At the same time, the

framework acknowledges the limitations of the substantive approach to values that we have outlined earlier: while endorsing apparently robust, well-known values it also indicates that these values are always dynamic. Turning to the practices of AI use and development is necessary to contextualize and update principles identified before in the bioethical domain.

In sum, the AI4People initiative and its Ethical Framework for a Good AI Society suggest that it is possible to combine different approaches to values when designing a framework for the responsible development of a new technology. Moreover, they show how the three approaches can complement and enrich each other, leading to a comprehensive form of RRI that none of them could have realized on its own.

Conclusions

In this paper we argued that current approaches to RRI offered by Von Schomberg (2011) and by Stilgoe, Owen, and Macnaghten (2013), regardless of their differences, are both liable to what we call the ‘entity trap’: the tendency to approach values as ready-made things, somehow available (‘out there’) for reflection and deliberation. We consider this a trap, because by ignoring that values are realized in human action one easily glosses over the complex hermeneutic work required to identify what values are at stake in a practice. This hermeneutic work is actually an important element of the ‘moral labor’ that RRI seeks to redistribute.

We therefore proposed a practice-based approach to values in the context of RRI that is rooted in pragmatism. Such an approach stresses the living, interactive and dynamic character of values. It highlights how values result from, as well as guide human action, and are continuously reshaped in and by the interactions of humans with their material and social environment. It also brings into view what types of hermeneutic work are involved when identifying values in practices.

Taking a practice-based approach to values has several methodological implications for RRI scholars. It implies, first of all, that values should be identified and understood in human practices. Moreover, hermeneutic methods are required to make explicit the moral complexity at stake in these practices *and* in identifying which values to consider in the RRI process. Such hermeneutic awareness should also be self-reflexive: the methods used by the RRI scholar to identify values will unavoidably co-shape her findings.

The practice-based approach to values is not devoid of weaknesses. However, neither are the existing approaches to identifying values in RRI. We suggest therefore, in line with our pragmatist orientation, that the choice for one rather than the other should be guided by the aims, context and resources of an RRI endeavor. Moreover, creative combinations are possible as well. Whatever the choices made, RRI scholars and practitioners alike should realize that referring to ‘values’ with nouns like ‘autonomy’, ‘privacy’ and ‘honesty’ is likely to lead our thinking about them astray, disregarding the evolving practices in which they are lived and which they serve to orient.

Acknowledgements

This paper profited from feedback given on (presentations of) earlier versions at the meeting of the Dutch network on Ethics and Politics of Emerging Technologies, November 23–24, 2018; the

workshop on Values & Technologies at the University of Twente on May 17, 2019; the Technology & Values research group meeting at the University of Twente in July 18, 2019; and the PhD course on Responsible Research & Innovation organized by the Dutch Research School of Philosophy on January 8, 2020 as well as from input given by students in the 2018-19 Technolab course of the MSc programme Philosophy of Science, Technology & Society at the University of Twente. Thanks to the participants on all occasions for their stimulating and critical input. We also thank the anonymous reviewers for their helpful responses.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Marianne Boenink's work on this paper was financially supported by a grant from the Netherlands Organization for Scientific Research, funding programme Responsible Innovation (NWO-MVI), with contributions from the Hersenstichting, Twente Medical Systems international (TMSi), and Clinical Science Systems (CSS) [grant number 313 99 309]. Olya Kudina's work on this paper has been supported financially by the project Value Change that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No 788321, and by the Netherlands Organisation for Scientific Research (NWO) under the research program 'Theorizing Technological Mediation' [project number 277-20-006].

Notes on contributors

Marianne Boenink is Professor of Healthcare ethics at the Radboud University Medical Centre in Nijmegen, with a specialization in philosophy and ethics of emerging biomedical technologies. She has been leading NWO-funded projects on early biomarker-based diagnostics for Alzheimer's disease and EEG-based prognostication of patients in coma, among others. While working on these, she has been developing methods for involving stakeholders in early societal and ethical deliberations that take into account the pitfalls of assessing technologies still in the making.

Olya Kudina is a postdoctoral researcher at the Ethics/Philosophy of Technology Section of the Values, Technology and Innovation Department at Delft University of Technology, where she explores the relation between technologies and value change. In her research, Olya combines empirical explorations and philosophical analysis. Olya's research interests span across the fields of ethics of emerging technologies, (post)phenomenology, hermeneutics, bioethics, and design.

ORCID

Marianne Boenink  <http://orcid.org/0000-0002-7017-0043>

Olya Kudina  <http://orcid.org/0000-0001-5374-1687>

References

- Arras, J. D., and H. Brody. 2013. "Methods of Practical Ethics." In *International Encyclopedia of Ethics*, edited by H. LaFollette, 3254–3265. Oxford: Wiley-Blackwell. doi:10.1002/9781444367072.wbiee768.
- Austgard, K. I. 2008. "What Characterizes Nursing Care? A Hermeneutical Philosophical Inquiry." *Scandinavian Journal of Caring Sciences* 22: 314–319.

- Beauchamp, Tom L., and James F. Childress, eds. 2013. *Principles of Biomedical Ethics*. 7th ed. New York: Oxford University Press.
- Bijker, W. E., T. P. Hughes, and T. J. Pinch, eds. 1989. *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge: MIT Press.
- Boenink, M., T. Swierstra, and D. Stemerding. 2010. "Anticipating the Interaction Between Technology and Morality: A Scenario Study of Experimenting with Humans in Bionanotechnology." *Studies in Ethics, Law, and Technology* 4 (2), doi:10.2202/1941-6008.1098.
- Boenink, M., L. van der Scheer, E. Garcia, and S. van der Burg. 2018. "Giving Voice to Patients: Developing a Discussion Method to Involve Patients in Translational Research." *NanoEthics* 12 (3): 181–197. doi:10.1007/s11569-018-0319-8.
- Boenink, M., H. van Lente, and E. Moors. 2016. "Diagnosing Alzheimer's Disease: How to Innovate with Care." In *Emerging Technologies for Diagnosing Alzheimer's Disease*, edited by M. Boenink, H. van Lente, and E. Moors, 263–275. London: Palgrave Macmillan.
- Bolt, I., M. F. Verweij, and J. J. M. van Delden. 2003. *Ethiek in praktijk*. Assen: Van Gorcum.
- Capurro, G., H. Longstaff, P. Hanney, and D. M. Secko. 2015. "Responsible Innovation: An Approach for Extracting Public Values Concerning Advanced Biofuels." *Journal of Responsible Innovation* 2 (3): 246–265.
- Coeckelbergh, M. 2012. *Growing Moral Relations: Critique of Moral Status Ascription*. New York: Palgrave Macmillan.
- Cohen, T., J. Stilgoe, and C. Cavoli. 2018. "Reframing the Governance of Automotive Automation: Insights from UK Stakeholder Workshops." *Journal of Responsible Innovation* 5 (3): 257–279.
- De La Bellacasa, M. P. 2011. "Matters of Care in Technoscience: Assembling Neglected Things." *Social Studies of Science* 41 (1): 85–106.
- De Saille, S. 2015. "Innovating Innovation Policy: The Emergence of 'Responsible Research and Innovation'." *Journal of Responsible Innovation* 2 (2): 152–168.
- Dewey, J. 1922a. *Human Nature and Conduct. An Introduction to Social Psychology*. New York: H. Holt and Company.
- Dewey, J. 1922b. "Valuation and Experimental Knowledge." *The Philosophical Review* 31 (4): 325–351.
- Dewey, J. 1929. *Experience and Nature*. London: George Allen & Unwin.
- Dewey, J. 1930. *The Quest for Certainty*. London: George Allen & Unwin.
- Di Giulio, G., C. Groves, M. Monteiro, and R. Taddei. 2016. "Communicating Through Vulnerability: Knowledge Politics, Inclusion and Responsiveness in Responsible Research and Innovation." *Journal of Responsible Innovation* 3 (2): 92–109.
- Dussauge, I., C. F. Helgesson, and F. Lee, eds. 2015. *Value Practices in the Life Sciences and Medicine*. Oxford: Oxford University Press.
- Elish, M. C. 2017. "Remote Split: A History of US Drone Operations and the Distributed Labor of War." *Science, Technology, & Human Values* 42 (6): 1100–1131.
- The European Parliament and the Council of the European Union. 2013. "Regulation (EU) No 1291/2013 of The European Parliament and of the Council of 11 December 2013 establishing Horizon 2020 – the Framework Programme for Research and Innovation (2014–2020) and repealing Decision No 1982/2006/EC." *Official Journal of the European Union*, L 347/104 (En), December 11.
- Fagerström, L., and I. Bergbom. 2010. "The Use of Hegelian Dialectics in Nursing Science." *Nursing Science Quarterly* 23 (1): 79–84. doi:10.1177/0894318409353800.
- Felt, U. 2018. "Living a Real-World Experiment: Post-Fukushima Imaginaries and Spatial Practices of 'Containing the Nuclear'." In *New Perspectives on Technology in Society: Experimentation Beyond the Laboratory*, edited by I. van de Poel, L. Asveld, and D. C. Mehos, 49–78. London: Routledge.
- Felt, U., S. Schumann, C. G. Schwarz, and M. Strassnig. 2014. "Technology of Imagination: A Card-Based Public Engagement Method for Debating Emerging Technologies." *Qualitative Research* 14 (2): 233–251.
- Fleming, V., U. Gaidys, and Y. Robb. 2003. "Hermeneutic Research in Nursing: Developing a Gadamerian-Based Research Method." *Nursing Inquiry*, doi:10.1046/j.1440-1800.2003.00163.x.

- Floridi, L., J. Cowls, M. Beltrametti, R. Chatila, P. Chazerand, V. Dignum, C. Luetge, et al. 2018. "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28 (4): 689–707.
- Frankena, W. 1967. "Values and Valuation." In *The Encyclopedia of Philosophy*, edited by Paul Edwards, 8–229. New York: Macmillan.
- Gadamer, H.-G. (1975) 2004. *Truth and Method*. Translated by J. Weinsheimer and D. G. Marshall, 2nd ed. New York: Crossroad.
- Genus, A., and A. Stirling. 2018. "Collingridge and the Dilemma of Control: Towards Responsible and Accountable Innovation." *Research Policy* 47 (1): 61–69.
- Grinbaum, A., and C. Groves. 2013. "What is 'Responsible' about Responsible Innovation? Understanding the Ethical Issues." In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, edited by R. Owen, J. R. Bessant, and M. Heintz, 119–142. Chichester: Wiley.
- Groves, C., K. Henwood, F. Shirani, C. Butler, K. Parkhill, and N. Pidgeon. 2016. "The Grit in the Oyster: Using Energy Biographies to Question Socio-Technical Imaginaries of 'Smartness'." *Journal of Responsible Innovation* 3 (1): 4–25.
- Grunwald, A. 2016. *The Hermeneutic Side of Responsible Research and Innovation*. Hoboken, NJ: John Wiley & Sons.
- Guston, D. H., E. Fisher, A. Grunwald, R. Owen, T. Swierstra, and S. Van der Burg. 2014. "Responsible Innovation: Motivations for a New Journal." *Journal of Responsible Innovation* 1 (1): 1–8. doi:10.1080/23299460.2014.885175.
- Harris, J. 2003. "In Praise of Unprincipled Ethics." *Journal of Medical Ethics* 29: 303–306.
- Hart, S. H. 1971. "Axiology – Theory of Values." *Philosophy and Phenomenological Research* 32 (1): 29–41.
- Hartley, S., C. McLeod, M. Clifford, S. Jewitt, and C. Ray. 2019. "A Retrospective Analysis of Responsible Innovation for Low-Technology Innovation in the Global South." *Journal of Responsible Innovation* 6 (2): 143–162.
- Hondreich, T., ed. 2005. "Value". In *The Oxford Companion to Philosophy* (2nd ed., 941). Oxford: Oxford University Press.
- Ihde, D. 1990. *Technology and the Lifeworld: From Garden to Earth*. Bloomington: Indiana University Press.
- Kennedy, J. 2002. *Een weloverwogen dood. De opkomst van de euthanasie in Nederland* [A Well-Considered Death. The Emergence of Euthanasia in the Netherlands]. Amsterdam: Bert Bakker.
- Kudina, O. 2018. "Accounting for the Moral Significance of Technology: Revisiting the Case of Non-Medical Sex Selection." *Journal of Bioethical Inquiry*: 1–11. doi:10.1007/s11673-018-9891-4.
- Kudina, O. 2019. "The Technological Mediation of Morality: Value Dynamism, and the Complex Interaction Between Ethics and Technology." PhD diss., University of Twente.
- Kudina, O., and P. P. Verbeek. 2019. "Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy." *Science, Technology, & Human Values* 44 (2): 291–314.
- Mepham, B. 2005. *Bioethics. An Introduction for the Biosciences*. Oxford: Oxford University Press.
- Mertens, M. 2018. "Liminal Innovation Practices: Questioning Three Common Assumptions in Responsible Innovation." *Journal of Responsible Innovation* 5 (3): 280–298.
- Mol, A. 2002. *The Body Multiple. Ontology in Medical Practice*. Durham: Duke University Press.
- Nauta, L. 2000. "Exemplarische bronnen van het autonomiebegrip [Exemplary Sources of the Autonomy Concept]." In *Onbehegen in de filosofie. Essays*, edited by L. Nauta, 67–87. Amsterdam: Van Gennep.
- Owen, R., J. R. Bessant, and M. Heintz, eds. 2013. *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. Chichester: Wiley.
- Owen, R., P. Macnaghten, and J. Stilgoe. 2012. "Responsible Research and Innovation: From Science in Society to Science for Society, with Society." *Science and Public Policy* 39 (6): 751–760.
- Owen, R., and M. Pansera. 2019. "Responsible Innovation and Responsible Research and Innovation." In *Handbook on Science and Public Policy*, edited by D. Simon, S. Kuhlmann, J. Stamm, and W. Canzler, 26–48. Cheltenham: Edward Elgar. <https://www.e-elgar.com/shop/handbook-on-science-and-public-policy>.

- Partridge, E. 1986. "Values in Nature: Is Anybody There?" *Philosophical Inquiry* 8 (1–2): 96–110.
- Pellé, S. 2016. "Process, Outcomes, Virtues: The Normative Strategies of Responsible Research and Innovation and the Challenge of Moral Pluralism." *Journal of Responsible Innovation* 3 (3): 233–254.
- Pols, A. J. 2014. "Radical Relationality. Epistemology in Care, Care Ethics for Research." In *Moral Boundaries Redrawn. The Significance of Joan Tronto's Argument for Political Theory, Professional Ethics, and Care Practice*, edited by G. Olthuis, H. Kohlen, and J. Heier, 175–194. Leuven: Peeters.
- Reckwitz, A. 2002. "Toward a Theory of Social Practices: A Development in Culturalist Theorizing." *European Journal of Social Theory* 5 (2): 243–263.
- Repo, P., and K. Matschoss. 2019. "Considering Expert Takeovers in Citizen Involvement Processes." *Journal of Responsible Innovation* 6 (2): 119–142.
- Rip, A. 2014. "The Past and Future of RRI." *Life Sciences, Society and Policy* 10 (1): 17.
- Rosenberger, R., and P. P. Verbeek, eds. 2015. *Postphenomenological Investigations: Essays on Human–Technology Relations*. Lanham: Lexington Books.
- Schroeder, M. 2016. "Value Theory." In *The Stanford Encyclopedia of Philosophy (Fall 2016 Edition)*, edited by Edward N. Zalta, <https://plato.stanford.edu/archives/fall2016/entries/value-theory/>.
- Shelley-Egan, C. 2011. "Ethics in Practice: Responding to an Evolving Problematic Situation of Nanotechnology in Society." PhD diss., University of Twente.
- Smith, J. A., P. Flowers, and M. Larkin. 2009. *Interpretative Phenomenological Analysis: Theory, Method and Research*. London: Sage.
- Stemerding, D. 2015. "iGEM as Laboratory in Responsible Research and Innovation." *Journal of Responsible Innovation* 2 (1): 140–142.
- Stilgoe, J., R. Owen, and P. Macnaghten. 2013. "Developing a Framework for Responsible Innovation." *Research Policy* 42 (9): 1568–1580.
- Swierstra, T., and A. Rip. 2007. "Nano-ethics as NEST-Ethics: Patterns of Moral Argumentation about New and Emerging Science and Technology." *Nanoethics* 1 (1): 3–20.
- Swierstra, T., D. Stemerding, and M. Boenink. 2009. "Exploring Techno-Moral Change: The Case of the Obesity Pill." In *Evaluating New Technologies*, edited by P. Sollie, and M. Duwell, 119–138. Dordrecht: Springer.
- Taebi, B., A. Correlje, E. Cuppen, M. Dignum, and U. Pesch. 2014. "Responsible Innovation as an Endorsement of Public Values: The Need for Interdisciplinary Research." *Journal of Responsible Innovation* 1 (1): 118–124.
- Vallor, S. 2016. *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford: Oxford University Press.
- Van de Poel, I. 2009. "Values in Engineering Design." In *Handbook of the Philosophy of Science*, edited by A. Meijers, 973–1006. Amsterdam: Elsevier.
- Van de Poel, I., and S. D. Zwart. 2010. "Reflective Equilibrium in R & D Networks." *Science, Technology, & Human Values* 35 (2): 174–199.
- Van Lente, H., T. Swierstra, and P. B. Joly. 2017. "Responsible Innovation as a Critique of Technology Assessment." *Journal of Responsible Innovation* 4 (2): 254–261.
- Verbeek, P.-P. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago: University of Chicago Press.
- Von Schomberg, R. 2011. "Introduction." In *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields*, edited by R. Von Schomberg, 7–15. Brussels: Directorate General for Research and Innovation, European Union.
- Von Schomberg, R. 2013. "A Vision of Responsible Research and Innovation." In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, edited by R. Owen, J. R. Bessant, and M. Heintz, 51–74. Chichester: Wiley.
- Wickson, F., and A. L. Carew. 2014. "Quality Criteria and Indicators for Responsible Research and Innovation: Learning from Transdisciplinarity." *Journal of Responsible Innovation* 1 (3): 254–273.
- Wittgenstein, L. (1953) 2009. *Philosophical Investigations*. Translated by G. E. M. Anscombe, P. M. S. Hacker and J. Schulte, 4th ed. Blackwell Publishing.