Responsible innovation 3

A European agenda?

Asveld, Lotte; van Dam-Mieras, Rietje; Swierstra, Tsjalling; Lavrijssen, Saskia; Linse, Kees; van den Hoven, Jeroen

DOI
10.1007/978-3-319-64834-7

Publication date
2017

Document Version
Final published version

Citation (APA)
Asveld, L., van Dam-Mieras, R., Swierstra, T., Lavrijssen, S., Linse, K., & van den Hoven, J. (2017). *Responsible innovation 3: A European agenda?* Springer. https://doi.org/10.1007/978-3-319-64834-7

Lotte Asveld · Rietje van Dam-Mieras
Tsjalling Swierstra · Saskia Lavrijssen
Kees Linse · Jeroen van den Hoven *Editors*

Responsible Innovation 3

A European Agenda?

Springer

Responsible Innovation 3

l.asveld@tudelft.nl

Lotte Asveld • Rietje van Dam-Mieras
Tsjalling Swierstra • Saskia Lavrijssen
Kees Linse • Jeroen van den Hoven
Editors

Responsible Innovation 3

A European Agenda?

Springer

Editors
Lotte Asveld
Department of Biotechnology & Society
Delft University of Technology
Delft, The Netherlands

Tsjalling Swierstra
Department of Philosophy
Maastricht University
Maastricht, The Netherlands

Kees Linse
Technology Foundation STW
Utrecht, The Netherlands

Rietje van Dam-Mieras
Leiden University
Leiden, The Netherlands

Saskia Lavrijssen
Tilburg, The Netherlands

Jeroen van den Hoven
Delft, Zuid-Holland, The Netherlands

ISBN 978-3-319-64833-0 ISBN 978-3-319-64834-7 (eBook)
DOI 10.1007/978-3-319-64834-7

Library of Congress Control Number: 2017953936

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
Introduction: Responsible Research and Innovation for Sustainability

Lotte Asveld and Rietje van Dam-Mieras

Abstract This chapter begins with a short description of the societal challenges that constitute the context in which Responsible Research and Innovation concept is developing. Subsequently the emerging Responsible Research and Innovation concept is mapped out. It ends with an overview of the different chapters.

1.1 Facing Twenty-First Century Challenges

Facing and dealing with the huge twenty-first century's challenges asks for actions in which all planetary inhabitants are involved. Path-breaking solutions are needed conceived in terms of fundamentally different sets of technologies, institutions and social arrangements from those we have today. Innovative approaches on both short and long term and the involvement of many stakeholders with often conflicting interests are required. The inherent complexity and uncertainty of the challenges not only asks for incremental changes leading to optimization within established frameworks, but also for fundamental transitions resulting in structural change and changes in framing conditions (Rikers et al. 2012).

The type of innovations needed cannot be restricted to designing and evaluating solutions, but must also engage with a process of paradigmatic change. This strategic management challenge requires special approaches. Visions of sustainable futures have to be created, the dynamics of co-evolutionary change on several innovation fronts have to be handled, the inherent uncertainty of change has to be faced and the communication about options and their implications with stakeholders and decisions makers has to be organized. Loorbach and Rotmans (2006) state that such transitions require organization-exceeding innovations at the system level, which

L. Asveld (✉)
Department of Biotechnology & Society, Delft University of Technology, Delft, The Netherlands
e-mail: L.Asveld@tudelft.nl

R. van Dam-Mieras
Leiden University, Leiden, The Netherlands
e-mail: vandammieras@zeelandnet.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_1

are realized by a variety of agents and which fundamentally change both the structure of the system and relations among the agents and other stakeholders.

There is another aspect that deserves attention. If we accept that our present society is a globalising society, an important point to realize is that globalisation exceeds the traditional frames of reference societies have. Every culture has its own specific worldview which is an important factor in its societal set of norms and values. Asking ethical questions such as 'What is a good life in a moral sense?' in a global society therefore, quickly results in 'defending our values against theirs' (van der Wal 2003). In a global space a multitude of different, culturally determined moral convictions will be at stake. How can that be seen as a source for development rather than as a source for conflicts?

Innovative development pathways should no longer focus just on the techno-economic system that delivers economic growth, but on the whole social-cultural-ecological system embracing the natural world, the socio-cultural world and interactions between the two. The natural world is inherently complex which by itself leads to indeterminacy and uncertainty. The interactions between mankind and nature are increasingly mediated through powerful technologies and the socio-cultural systems that interact with the natural world are, in their own ways, just as complex and unpredictable. The threatening irreversible ecological change and the limited capacity of humanity to adapt or respond when ecological change undermines the very basis of human survival or quality of life constitutes a powerful case for a precautionary approach (Jansen et al. 2008).

Most people will agree that, if mankind wants to continue to live on Planet Earth, societal development should become more sustainable than it presently is. In spite of that bringing about sustainable development turns out to be extremely difficult. The Sustainable Development Goals agreed upon within the UN context in 2015 (sustainabledevelopment.un.org) give an overview of the broad spectrum of actions needed to deal with our present global challenges. The SDGs are:

- End poverty in all its forms everywhere
- End hunger, achieve food security and improved nutrition and promote sustainable agriculture
- Ensure healthy lives and promote well-being for all at all ages
- Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
- Achieve gender equality and empower all women and girls
- Ensure availability and sustainable management of water and sanitation for all
- Ensure access to affordable, reliable, sustainable and modern energy for all
- Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
- Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
- Reduce inequality within and among countries
- Make cities and human settlements inclusive, safe, resilient and sustainable
- Ensure sustainable consumption and production patterns

- Take urgent action to combat climate change and its impacts
- Conserve and sustainably use the oceans, seas and marine resources for sustainable development
- Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
- Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
- Strengthen the means of implementations and revitalize the global partnership for sustainable development

Dealing with the twenty-first century challenges asks for innovative approaches in all societal domains at all levels of scale.

1.2 Responsible Research and Innovation

Responsible research and innovation (RRI) is a research and innovation approach aiming to take into account effects and potential impacts on environment and society. It tries to anticipate and assess potential implications and societal expectations with regard to research and innovation at an early stage of research and innovation processes. RRI is increasingly used interchangeable with Responsible Innovation (RI), also in this volume. Especially when the innovation takes place in a business context, RI is used instead of RRI, although both refer to largely the same concept.

RRI has been part of several European Framework Programmes. An Expert Group on the State of the Art in Europe of Responsible Research and Innovation chaired by Jeroen van den Hoven describes RRI in the following way. *'RRI refers to the comprehensive approach of proceeding in research and innovation in ways that allow all stakeholders to be involved in the processes of research and innovation at an early stage (A) to obtain relevant knowledge on the consequences of the outcomes of their actions and on the range of options open to them and (B) to effectively evaluate both outcomes and options in terms of societal needs and moral values and (C) to use these considerations (under A and B) as functional requirements for design and development of new research, products and services. The RRI approach has to be a key part of the research and innovation process and should be established as a collective, inclusive and system-wide approach'*. (Options for Strengthening Responsible Research and Innovation, European Commission, Directorate General for Research and Innovation Science in Society, EUR25766EN, https://ec.europa.eu).

RRI thus is about making science with society for society. The RRI Tools community project for research, technological development and demonstration (rri-tools.eu), funded by the European Union's Seventh Framework Programme, approaches RRI from six different perspectives in which different aspects come to

the fore. In RRI for Policy Makers 'Address grand societal challenges', 'Increase public trust' and 'Build a responsible future' are important aspects. In RRI for Researchers aspects in focus are 'Incorporate other views', 'Evaluate the impact', 'Anticipate, reflect, engage, act', 'Share the process, make it worthy' and 'Share your responsibility'. RRI for Business and Industry is about 'Invite all relevant actors', 'Boost socially acceptable innovation', 'Find new business opportunities', 'Reinforce your customers' trust', 'Add value, secure your future'. In RRI for the Education Community aspects are 'Responsibility is a learnt behaviour', 'Stimulate curiosity', 'Contextualize science' and 'Empower future generations'. RRI for Civil Society Organisations and RRI for citizens both deal with 'Your voice and ideas are important', 'Co-create the future', 'Be informed, be critical' and 'the media are key actors'.

The RRI approach and the RRI Tools community project are aiming at involving all stakeholders in research and innovation processes from an early stage on, which looks like a promising approach. It is important to realise that the description of RRI and RRI-tools given above is mainly from a European perspective. As it would be rather naïve, however, to think that a European RRI concept and RRI tools could be seen as a 'cure all solution' in a global society. Dealing with uncertainty, tensions among nations for geopolitical and cultural reasons, and differences in access to resources and technology will lead to different perceptions of solutions and priorities in different parts of the world. Therefore the RRI concept and tools will need contextualisation for application in other parts of the world.

1.3 Responsible Innovation 3: A European Agenda?

This volume of Responsible Innovation, the third in a series, reports on the developments of RRI in the context of the Responsible Innovation Research programme of the Dutch Science Council (NWO-MVI), but is certainly not limited to Dutch research only. It brings together cases that shed a light on how the RRI approach can be applied to specific innovations, with reflections and overviews on a more general level and a discussion of what RRI might mean in non-western cultural settings. The chapters are a selection of the fourth NWO conference on Responsible Innovation held in 2015 in The Hague. An overview of its contents is given below.

The first part of the book deals with general, conceptual issues relevant to RRI.

Tsjalling Swierstra (Chap. 2) describes the political and technological trends that explain the current prominence of RRI in Europe and the US. He links the financial crisis, social media, the rise of citizen science and intimate technologies to the four prominent dimensions of RRI: anticipation, reflectivity, inclusion and responsiveness and identifies hurdles for the further development of RRI.

Job Timmermans (Chap. 3) provides a comprehensive overview of the current RRI landscape, comprising the main contributors, the terms by which RRI is perceived and the areas it is being applied to. The analysis shows which actors act as hubs in the different thematic or regional segments that together make up the

discourse. The study hints towards future directions of the discourse that are relevant to both policy-makers and RRI researchers.

Ulrike Felt (Chap. 4) fleshes out the structures of the academic working environment and how they impact on the capacity of academic researchers to become truly responsible researchers. While identifying these challenges this chapter aims to explore how academic researchers can potentially make sense of RRI and turn it into an academic core value.

Pim Klaassen et al. (Chap. 5) take an iterative approach to the conceptualisation of RRI, based on the EU-funded RRI tools project, moving back and forth between theory and practice. As an outcome of this approach, they suggest that RRI is best captured if in R&I governance attention is paid to the five p's of *Purpose*, *Products, Processes, Preconditions* and *People*, and that further elaborations on the meaning of RRI should happen in dialogue with attempts at practicing RRI.

Federico Vasen (Chap. 6) takes us on to new frontiers in this chapter where he explores how the framework of RRI can be elaborated in order to effectively impact on science, technology and innovation (STI) policy in the developing world, particularly in Latin America. To this purpose, he analyses the dominant framework of science, technology and innovation policies in the region and proposes topics that should be included in the RRI agenda to become more responsive to issues related to other geographical contexts.

The second part of the book deals with more applied issues, such as those related to methods, topics and actors. Sometimes the methods or analyses have been developed with reference to a particular field of technology, but all chapters bear relevance to the general concept of RRI.

Steve Rayner (Chap. 7) discusses how a framework for responsible innovation can be applied to the emerging and already controversial field of climate geoengineering. This specific technological field provides an opportunity to develop RRI according to a model of guiding societal principles and technology-specific protocols. Rayner relies on the Oxford principles of Responsible Innovation that can support innovations in contributing to societal goals.

Udo Pesch et al. (Chap. 8) describe the interplay between formal and informal assessment in the case of two plans for the implementation of energy technologies in the Netherlands, What these studies reveal is that different underlying values emerge in different discourses. The interplay of these discourses can lead to a progressive entrenchment of positions. To avoid such a situation, the symmetry between different claims for the public interest needs to be attended by policy makers.

Lotte Asveld (Chap. 9) explores how the concept of identity impacts on the capacity and willingness of individual actors to participate in a social learning exercise such as RRI. Asveld claims that identity consist of both an individual's moral framework and her social role. Taking these two aspects of identity into account can help to design more effective RRI exercises.

Klara Pigmans et al. (Chap. 10) show that the current focus on interests in stakeholder participation creates or exacerbates conflicts, while a focus on underlying values may prevent this. Additionally, they provide a modelling framework to support value driven stakeholders participation in which they relate varying conceptions

of values to interactions between stakeholders and the organisational structures they are part of. This can be of support to stakeholder participation in water governance and technology governance in general.

The third part of the book focusses specifically on RRI in a business context. Here the term RRI is mostly replaced by RI.

Lubberink et al. (Chap. 11) investigate how Responsible Innovation can be applied in the business context. To answer this, they study which aspects of Responsible Innovation are conceptually similar and dissimilar from social- and sustainable innovation. These approaches are already embedded in the business context. The insights obtained are used for conceptualising Responsible Innovation in a business context.

Blok et al. (Chap. 12) address the question of whether and where normative considerations play a role in the innovation process of companies. Results of this study suggest that ethical decision making does not take place throughout the entire innovation process, but might be located at a higher or strategic level in the company. The results have important consequences for the operationalisation of the dimensions of responsible innovation throughout the innovation process.

Noorman et al. (Chap. 13) examine the applicability of broad stakeholder engagement in corporate settings and in smaller scale technological projects. Based on their analysis, they argue that there is a need for the field of RI to explore additional and alternative ways to address issues of stakeholder commitment and inclusion, in order to make RI's deliberative ideals more applicable to the rapid, fluid, partial, and provisional style of deliberation and decision making that were found in corporate contexts.

Together the contributions in this book cover a wide range of the current developments in RRI, from the move to other cultural contexts, the availability of increasingly refined methodologies and conceptualisations, to the move to application in the business context. We believe this volume offers a valuable contribution to the current discourse on RRI.

References

Jansen, Leo, Paul Weaver, and R. van Dam. 2008. Education to meet new challenges in a networked society. In *Innovation in education*, ed. J.E. Larkley and V.B. Maynhard, 1–51. New York: Nova Science.

Loorbach, Derk, and Jan Rotmans. 2006. Managing transitions for sustainable development. In *Understanding industrial transformation*, 187–206. Dordrecht: Springer.

Rikers, J., G.R. de Snoo, and M.C.E. van Dam-Mieras. 2012. Higher education and sustainability in europe. In *Higher education in the world 4. Higher education's commitment to sustainability: From understanding to action*, GUNi series on the social commitment of universities, vol. 4, 113–127. London: Palgrave MacMillan.

Van der Wal, Koo. 2003. Globalisierung, Nachhaltigkeit und Ethik. Natur und Kultur. *Transdisziplinäre Zeitschrift für ökologische Nachhaltigkeit* 4 (1): 100–119.

Part I
Unravelling the Concept of RRI

Chapter 2
Economic, Technological, and Socio-epistemological Drivers Behind RRI

Tsjalling Swierstra

Abstract To help explain Responsible Research and Innovation's recent rise to prominence, I relate four of its defining features – anticipation, stakeholder inclusion, reflexivity, responsiveness – to relevant recent economic, technological and socio-epistemological developments. The economic crisis affected the way anticipation is practiced. The need to justify increased public expenditure brought with it a shift from anticipation as harm-avoidance to anticipation as the attempt to realize the common good. Stakeholder inclusion has received an impetus from the new social media that lend a voice and a face to distant stakeholders, and help foster a sense of mutual interdependence. The growing awareness that many forms of research and innovation fail to deliver on their societal promises has helped to broaden Merton's 'organized scepticism' to include the input of other disciplines and non-experts. Finally, the progressive entanglement of intimate technologies and the lifeworld, has led to new demands that research and innovation extends its responsiveness to impacts on the good life and on the good society. However, all these developments come with pertinent questions and problems that make the success of RRI far from certain.

2.1 Introduction

The concept of Responsible Research and Innovation is only a couple of years old. So, how to explain the present prominence of RRI (at least in some circles in the EU and USA)? I suggest four developments that help explain this prominence, relating them to four main features of RRI. In doing so I will also point out some problems that deserve further attention if RRI is to become a more permanent feature of the research and innovation landscape.

In an influential article Stilgoe et al. (2013) list four elements that define RRI. These elements largely overlap with the four steps that define the RRI approach as developed by the Dutch Science Foundation, the organizer of the conference behind this volume. The first element is *anticipation* – or being *pro-active* in the

T. Swierstra (✉)
Department of Philosophy, Maastricht University, Maastricht, The Netherlands
e-mail: t.swierstra@maastrichtuniversity.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_2

Dutch RRI approach. Look before you leap. Explore in advance the possible consequences of your innovation, including the undesirable and unintended ones. The second element is *inclusion of stakeholders* in the innovation process – corresponding with the requirement of a *valorisation panel* in the Dutch model. Those suffering or enjoying the consequences of an innovation – whether as intended users or not – should have a say in its design, development, and implementation. The third element is *reflexivity* – which can be found in the Dutch RRI approach in the form of a demand for *interdisciplinarity*. Here the aim is to avoid information deficits and tunnel vision, including technological fixes, by creating an environment where assumptions are made explicit and probed from different perspectives. Obviously, this reflexivity is closely tied to the inclusion of stakeholders, as it is their heterogeneity that provides the best guarantee for open-mindedness. In the approach of the Dutch Science Foundation, this heterogeneity is specified by the requirement that scholars from the humanities, sciences, and social sciences collaborate. Finally: *responsiveness* – or the demand for *valorization* in the Dutch approach. The first three steps are to no avail if technology actors and policy makers don't translate the results of the deliberations into the design, development, and implementation of technology. This is admittedly still the weakest link in RRI, but the requirement in the Dutch approach that industry carries part of the costs of a RRI trajectory goes some way to ensure that the outcomes of the deliberations will at least be seriously considered. Other such incentives are the requirement that these outcomes are public, and the fact that the stakeholders can be expected to protest if their efforts are not taken up.

How new are these four key elements? Not very. However, they do receive a new twist in the light of recent – overlapping and mutually reinforcing – economic, technical, and socio-epistemological developments.

2.2 Economic Crisis and Anticipation

Anticipation is far from new in the context of research and innovation. Technology Assessment, for example, is more than half a century old. Marketing is older (even if the label may not be). No technology firm that doesn't probe the future, e.g. in the form of a business plan, before embarking on an uncertain and costly innovation trajectory. Nor is anticipation for *moral* reasons new. Discussions about business ethics and corporate citizenship have been going on for decades. The slogan 'People, Planet, Profit' is already in danger of becoming a cliché. 'People' here usually refers to the moral values of safety, health, and employment. 'Planet' refers to the moral value of sustainability. 'Profit' is then presented as the precondition for achieving the first two values.

Notwithstanding these continuities, RRI does approach anticipation in a somewhat novel fashion: anticipation becomes *aspirational*. A development that helps to explain this shift within the practice of anticipation, is the recent economic crisis.

This crisis destabilized the role-division between private and public agents regarding anticipation, because now investments in research and innovation had to be justified differently.

In the context of technological innovation, anticipation was for a long time largely framed in terms of a division of tasks between private companies and public government. (Military technologies obviously don't fall under this regime.) Identifying the positive aims and values of technology development was supposedly the domain of private market transactions. It was up to the entrepreneurs to anticipate what people wanted (or to make them want it) and then to make a profit by catering to these wants. Decisions about what is (un)desirable are in this model delegated to anonymous market forces. The role of civil society is limited to providing technology actors with feedback through what Albert Hirschman (1970) dubbed the 'exit' mechanism. Citizens influence technology development not through collective deliberation, but by privately 'voting' with their wallet. The model admits that technology development can have collective or long term harmful consequences for which the market is blind. Within the model, it befalls the state to protect long-term interests by installing an accountability regime that incentivizes private firms to not explode people, poison them, or devastate the natural environment.

One flaw of this traditional role division between private and public sector is that it is too kind on entrepreneurs. In *The entrepreneurial state* (2013) Mariana Mazzucato argues convincingly that governments are in fact key players in technological innovation. Private firms prove to be averse to engage in risky, long-term, large scale projects. So, to avoid procrastination and stimulate innovation, governments step in to shoulder those costs and risks collectively. The burden of technological innovation is thus shifted from private entrepreneurs to ordinary tax paying citizens. Internet, for example, was not developed in Silicon Valley by daring visionaries, but by successive American governments. Similarly, the Dutch government massively invested in genomics and nanotechnology to create a space sufficiently safe for entrepreneurs.

Even though this enabling role of public risk taking is a permanent feature of modern market economies, it usually remains underappreciated and somewhat invisible. That changes, however, in times of economic crisis. Then entrepreneurs become even less entrepreneurial, and the need for a safe space thus becomes subsequently more pressing. But when research and innovation are visibly paid for by public funds, the afore mentioned role division between private and public parties can no longer be upheld. A new justificatory regime takes shape, which affects how possible impacts of research and innovation are to be anticipated. Publicly financed research and innovation is subjected to requirements of *democratic legitimation*. Governments but must justify their decisions to a public forum in terms of a common good. It no longer suffices to delegate goal-setting to the anonymous market forces. Justifications now must embody to a more positive logic: not what we want to *avoid*, but what do we want to *achieve*? Not: what is *harmful* technology? But: what is *desirable* technology? Which technologies should we invest in to make this world a better place? It is no coincidence that *Horizon 2020*, the European Science

Program focuses on so-called Grand Societal Challenges. The focus shifts from avoiding harm to conduce human and nonhuman flourishing – what René von Schomberg (2013) dubbed "the right impacts".

The market may be strong in determining what consumers factually want and accept – this is what marketing is about -, but it is weak in establishing moral desirability. It fails to distinguish between factual 'acceptance' and moral 'acceptability'. What we *should* want can only be decided on in a deliberative arena: via *voice* instead of *exit*. (Hirschman 1970) Furthermore, where harm is comparatively clear and non-controversial, positive goals like the flourishing of humans, animals, and nature, or like the good society, typically invite plural answers. Aspirational anticipation thus comes with an emphasis on ethical discussion.

The financial and economic crisis thus helped broaden RRI's anticipation, and make it more ambitious. But this role of a – contingent – driver like an economic crisis should also give us pause. What will happen to broad and aspirational anticipation when the crisis is over and the importance of government funding subsequently decreases? What are RRI's chances in technological domains that don't depend on public funding? When market mechanisms recuperate their previous prominence, will RRI turn out to be temporary fad?

2.3 New Social Media and Stakeholder Inclusion

Like anticipation, RRI's plea for stakeholders' inclusion in research and innovation is not new. Despite the considerable numbers of scientists and technologists clinging to the aura of science and technology to shield them from interaction with laypersons, science's ivory tower has been under attack since the sixties and seventies of the previous century. And the attack has only grown fiercer, drawing strength from societal sub currents like the gradual democratization of Western societies, the rising levels of education, and the more public character of scientific controversies with experts fighting one another in front of the citizenry.

The influence of information and communication technologies is a new factor, though, and it impacts on stakeholder inclusion. The new social media have considerably enhanced the capacities of citizens to challenge expert's accounts and to concoct and advertise their own versions of reality. If a scientist nowadays tries to instruct society what to think and do, the Internet immediately explodes with indignant reactions. And this strengthening of stakeholder power is not restricted to the citizens of the Western democracies. The new social media have made it less easy for the West to externalize, outsource and hide technology's negative consequences to far-away places. They can confront Western societies with the hidden consequences of their actions; offer a forum to distant stakeholders in developing countries; and – ultimately – foster a sense of global interdependence and membership of one 'community of fate' (Van Gunsteren 1998).

Although stakeholder inclusion is not novel, the new social media affect the hierarchies between experts and non-experts. Empowering the non-expert stakeholders,

and broadening the circle of stakeholders to include those presently absent and hard to represent, the new social media fuel an idea that goes at least back to Locke and Rousseau: those affected by power, have a right to co-determine how that power is exercised.

Of course, here too pertinent questions abound. The first category relates to the fact that the social media yet only partially succeed in lending voice to stakeholders and disclose hidden and distant consequences. The unfortunate textile workers in Bangladesh who, on top of being underpaid and overworked, also had their factory collapse around them, managed to become news in the Western media. There were outcries and policies were announced. But today their predicament seems already largely forgotten. They simply lack the power to make themselves heard more permanently. The example shows that being a stakeholder is not simply a moral category: it also involves a lot of hard work, also by other agents like NGOs, and enabling political, social and economic conditions. The success of the new social media still depends on a whole array of other factors.

Furthermore, the new social media so far fail to provide a technological environment that is conducive to deliberation – in any sense that remotely resembles Habermas' conception of power-free communication that informs most models of deliberative democracy. We are increasingly confronted with unexpected phenomena like post-truth, filter bubbles, fake news, Internet trolls, and a debilitating lack of civility in cyberspace. The new social media hide things as easily as they expose them; they lend people a voice but also shut them up. The dream of stakeholder deliberations where different types of knowledge get combined into one encompassing truth, often turns into the nightmare of fragmentation, tweets, discrimination, propaganda, gut feelings, and unreason.

2.4 Disorganized Scepticism and Reflexivity

RRI's third feature is reflexivity. The demand for reflexivity is, again, hardly new. It is actually as old as Socrates. Everyone suffers from tunnel vision, rationalization, and all other imaginable forms of sloppy thinking. (Kahneman 2011) Rationality depends on the kindness of others to challenge us and make us aware of our placid assumptions, blind spots, flawed arguments, and cognitive mistakes like *confirmation bias*. Helping is not always pleasant. Socrates was condemned to death. But nor is being helped. It is in equal parts enlightening and threatening to expose oneself to conflicting perspectives. That is why reflexivity needs a conducive environment: reflexivity must be facilitated. As early as 1942 Robert Merton identified *Organized* Scepticism as one of the defining features of true science. (Merton 1973) He identified the scientific community as the environment conducive of reflexivity. This community was understood as a safe environment, set apart from the irrationality, propaganda and power politics characterizing civil society at large.

What is (somewhat) novel in RRI is that reflexivity gets linked to its second feature: stakeholder inclusion. Reflexivity is no longer a feature defined in opposition

to the "common folk", but is made to depend on them. Whereas in the previous section stakeholder involvement was motivated by appealing to the democratic right to control the powers that control us, linked with reflexivity it primarily appears as an epistemic mechanism. Bringing scientists, engineers and stakeholders together in collaborative formats generates a broader and more encompassing mind-set that leads to more reliable, more robust, superior knowledge and technology.

Which development prompts this link between reflexivity and the inclusion of non-scientific outsiders? A plausible answer is that *science and technology have too often failed in their promises to civil society*. Scepticism is no longer the prerogative of the community of peers, but gets extended to the community of citizens. And this spontaneous, disorganized, scepticism is not restricted to science's truth claims, but extends to science's and technology's social benefits – even if their truth and efficacy is not contested. The Enlightenment promise entailed that science and technology would replace ignorance, uncertainty, evil, and impotence by knowledge, certainty, good, and control. Eventually, everyone would profit from truth and power. It is exactly that socio-epistemological promise of science and technology to realize the common good, that is now provoking a sceptical response from lay persons. They are the first to point out science's and technology's failures, for example drawing on local complexities resistant to generalizing narratives imposed from above. (Scott 1998) From the value-laden perspective of lived experience, the Enlightenment's promise often appears as a fata morgana, as an ever-receding horizon.

The more knowledge we gather, the more we become aware of reality's evasiveness, of its awe-inspiring *complexity*. Progress in the sciences does enhance our abilities to intervene. But it simultaneously makes us aware of how limited our powers are in the face of the infinitely intricate systems we call life and society. The more data we collect, the less confidence we seem to have to predict, prevent, promote the future. The more time saving devices we apply, the more we find ourselves pressed for time. The more practical powers we collect thanks to science and technology, the more undefined responsibilities and moral uncertainties we find ourselves faced with. (Jonas 1985) The proliferation of paradoxes like these undermine the Enlightenment hubris. It is this disillusionment with science's and technology's ability to solve life's problems that constitutes a major driver behind the renewed call for stakeholder-driven types of reflexivity.

An example is healthcare studies. Through self-tracking devices, citizens produce information on their well-being, lifestyle and physiological processes, which can then be shared digitally. These enthusiastic citizens generate data for science and technology which cannot be collected without their active participation. (Sharon and Van Zandbergen 2016) And as many diseases prove to be lifestyle related, health gains cannot simply be the result of new science and technology, but only of the subtle calibration of many heterogeneous elements and agents, ranging from medical drugs to advertisements, self discipline, patient communities, lifestyle advice and the built environment. Another example is the discussion about smart cities. The driving vision here is that processes in cities can become more sustainable, efficient and safe by using Big Data and information technologies. To avoid that this becomes a top-down, technocratic, project, municipalities seek out citizens to help shape their own smart cities.

And again, there are questions requiring exploration. For example: how to ensure real reflexivity? In practice, stakeholder involvement can be indistinguishable from an opinion poll. And where to draw the line between marketing research, for which firms should pay themselves, and enhancing reflexivity by learning from stakeholders, a goal worthy of being subsidized by the Dutch Science Foundation in the name of responsible research and innovation? Another issue is that science-citizens collaborations can generate considerable frustration. The reason is that it is often unclear which expectations are legitimate and which ones are not. The terms of the new social contract are as yet undefined. How to distribute the burdens and benefits of such a collaboration? An example here is 23 and Me, an American web-based company, that seduces people to share their genome and upload their lifestyle details, by promising that these contributions will help to develop new humankind-saving drugs. Only later these disinterested volunteers find out their data is used to file commercial patents. Behind the open access, sharing, façade there is actually a corporation making money with the data that were donated for idealistic reasons. Many people who uploaded their data felt betrayed because the terms of that deal were rather implicit.

2.5 Intimate Technologies and Responsiveness

The fourth dimension of RRI is responsiveness. Researchers and innovators should not only be willing to involve outsiders in the deliberative process, but also to act on the outcomes of that process. And again, responsiveness is not a distinguishing feature of RRI, as we may safely assume that Technology Assessment has always been motivated by the desire to guide research and development.

But there are indications that in RRI responsiveness gets interpreted more radically, or more broadly, than in – say – Technology Assessment. Due to a progressive entanglement of technologies and our lifeworld, societal concerns regarding emerging technologies have broadened, and it is this broadening that gives a new twist to the responsiveness requirement. The societal development working in the background here, is the increasing prominence of so-called 'intimate technologies'(Van Est 2014). These technologies challenge the still-dominant instrumentalist conception of technology, paving the road for a conception of technology that stresses its active role as co-shaper of our lives.

It is still common to frame technology as an *instrument.* Two promises then typically get attached: the technology will enhance our capability to intervene in the world, and it will make life easier by setting us free from undesirable tasks. Technology is thus presented as an enabling device that allows the user, by giving her more free time, to conduct her life as she chooses to. This instrumentalist perception of technology thus perfectly aligns with a modern, individualist, liberal view of society and the good life (Borgmann 1984).

In this conception of technology, all that counts is that it doesn't explode in your face, poison you, deplete and pollute the environment; and that people don't misuse this neutral tool for devious purposes. These are the concerns technology and policy

actors typically consider to be sufficiently 'hard' or 'objective' to be 'responsive' to. What generates this 'hardness'? 'Hard' concerns score high on three scales (Swierstra 2012, 2015). First, on the scale between qualitative and quantifiable, they are the latter. For example: you can calculate the chance Y that X people will die from a nuclear explosion. Second, on the scale of moral ambiguity to moral certainty, they again score high on the second pole. Hard concerns are typically about non-controversial instances of harm, appealing to widely-held values like health, safety, sustainability, economic growth and employment. We all agree that when people die from an explosion or from radiation, that is unequivocally bad – because harmful. And third, on the scale between co-produced by many (human and non-human) factors and clearly caused by an identifiable (human or non-human) agent, hard impacts are those are directly caused by the (failing or misused) technology. People died because the reactor malfunctioned or was misused by terrorists. In both cases, we know who is accountable: the manufacturer of the technology, or its misuser. Of course, we can and do debate the exact numbers, or who is to blame exactly for what, but no one disagrees that these are matters of serious, public, concern. Developed countries therefore have accountability regimes in place, that incentivize innovators to be 'responsive' to these hard concerns.

But in the last decades, we witness a new type of concerns coming to the fore. Or to be more precise: demands that were previously considered as too 'soft' to warrant public attention, increasingly resist such easy privatization. And it is these new concerns that now put new strains on the responsiveness requirement, and can help to redefine it.

These new concerns arise from the increasing entanglement of technologies and life world. This is a new development, even if it is a gradual one without any clear breaks. Since the beginnings of the Industrial Revolution, thinkers have worried that humans were enslaved by the machines that were meant to be our servants. Think of the iconic image of Charlie Chaplin in *Modern Times*, devoured by the machine, turned into an insignificant cog. But these thinkers were confident that they were diagnosing the dire fate of *other* people, out there, in the factories. They reflected from a safe distance on other, poor, exploited, people, who were instrumentalized/objectified/manipulated/alienated/etcetera by an impersonal technological regime. But they believed that their own lives, by contrast, were largely exempt from the pervasive and intrusive influence of technology.

In the present, that dividing line no longer exists because technology now pervades everyone's lives. Technology is no longer 'over there'; it is everywhere. Technology is now *in us*, for example as pacemakers, deep brain stimulation, or maybe simply in the form of medication for chronic diseases. Or it has become *part of us*, as in the case of 'mind walkers' – paralyzed people who through their brainpower and thanks to a computer interface, can now command an exoskeleton. It is *between us*, in the form of communication technology. ICT is reshaping our relations, opening up new forms of community on the one hand, while breaking up previous communities on the other (e.g. through information bubbles). Technology is increasingly also *about us*. Data about our most private bodily processes, mental states, and activities are stored in databases, and made available for

endless comparisons and statistical mining. Privacy, experts tell us, is no longer a practicable ideal. And finally, technology starts to become *like us*, as in the case of robots. Or think about the possibility of e-coaching, where part of your super-ego is delegated and outsourced to an application on your smart phone that support the willing spirit in its attempts to keep the upper hand over the weak flesh. The overall point is: technology is getting really, really ubiquitous and really, really close. It is increasingly 'intimate'. (Van Est 2014).

These intimate technologies typically raise concerns that are much 'softer' than the 'hard' concerns discussed before. In the first place, they are typically qualitative rather than quantitative. For example: how will the quality of our social relations be affected by the prevalence of new communication devices and software? How will our understanding of 'friendship' change under the influence of Facebook? Such a question doesn't allow for quantification, but requires a qualitative answer. Secondly, the impacts of intimate technologies are not clearly and unequivocally harmful. If we look at how friendship is impacted by digital communication, some people will probably argue that there are indeed changes, but that these are harmless, or actually constitute an improvement; even if others claim that Facebook turns the idea and practice of real friendship into a travesty. And third, even if your child turns into a Facebook or computer game zombie, this is not directly caused by these technologies. S/he clearly also has some own responsibility, as do the parents. So, there is no unequivocal causal link between technology and impact, as is proven by the existence of all those people who resist the temptation and do not turn into zombies. Maybe the technology normalizes, invites or enables certain changes in practices, values, ideas and identities, but the impacts are typically co-produced by the technologies and their human users together.

These intimate technologies thus belie the instrumentalist conception of technology. They are clearly not simply enabling us to be and do what we wanted to be and do before these technologies entered our lives. They are clearly not neutral devices enabling us to live our lives as we want to. Instead, they help define what a good life is, and who we want to be. We have to accept that morals and technologies mutually shape one another. Even while aiming to influence the design, development and application of technology, our values, norms, aspirations, duties, rights, responsibilities, conceptions of harm, and so forth, are in their turn shaped by these technologies. This is the phenomenon of technomoral change – one of the most interesting 'soft impacts' of technology. We may believe that we are the ones applying moral standards to technologies, but all the time technology is 'talking back' (Swierstra 2013) and is destabilizing our standards, inviting us to modify them. For example: the smart telephone was introduced to serve our then existing goals, but since its introduction and permanent innovation, it has also created new needs, routines, expectations, etc. We are not enslaved by the phone, but we are certainly married to it. And as in any marriage, the partners shape one another in a myriad of more or less visible ways. For example, deep brain stimulation started out as a technology to treat Parkinson's disease, but now we look whether it can also help treat depression, and if that were to be successful, then it will go on to redefine what is considered 'normal' or 'healthy' brain behaviour.

The rise of 'soft' concerns, propelled by the intimate technologies, places a strain on the established accountability regimes, that restrict themselves to hard impacts. As a result of the spread of intimate technologies, this restriction is rapidly becoming problematical, which is one of the drivers behind RRI. Responsibility in the context of RRI should be *comprehensive* responsibility. The promise of comprehensive responsibility is that technology development becomes more responsive to a broad array of values and concerns, and not only to a very limited, albeit very important, subset of those.

This broadening of the set of concerns that belong on the agenda of researchers and innovators, comes with its own challenges too. The biggest question is how to restyle public space, public discussions, so that we can articulate these soft concerns without them being immediately privatized, and to collectively deliberate on them. We as citizens of modern, pluralistic, liberal societies are so drenched in thin morality (Walzer 1995) that we by now seem to lack the thick, qualitative, vocabulary to name and explore soft impacts.

2.6 Conclusion

I have argued that the economic crisis affected the way anticipation is practiced. The need to justify increased public expenditure brought with it a shift from anticipation as harm-avoidance to anticipation as the attempt to realize the common good. Stakeholder inclusion received an impetus from the new social media as these lend a voice and a face to distant stakeholders, and help foster a sense of mutual interdependence. The growing awareness that many forms of research and innovation fail to deliver on their societal promises has helped to broaden Merton's 'organized scepticism' to include the input of other disciplines and non-experts. Finally, the progressive entanglement of intimate technologies and the lifeworld has led to new demands that research and innovation extends its responsiveness to impacts on the good life and on the good society.

However, all these developments come with pertinent questions and problems that make the success of RRI far from certain. In a sense, the economic crisis provided a window of opportunity for RRI, but we should devise strategies to ensure that RRI survives after that window has closed. It is encouraging that aspirational ethics is akin to the ethical explorations, provoked by intimate technologies, of how science and technology can contribute to the good life and the good society. As these technologies are here to stay, so is aspirational ethics. More broadly, in Western countries – after a prolonged period of liberal individualism and postmodern deconstruction – many experience a need of new aspirational narratives to guide collective action and to provide meaning to individual lives. That these narratives are now primarily available in polarising, xenophobic and belligerent forms, only underlines the need to develop more positive narratives. Narratives that will have little difficulty to incorporate RRI, as research and innovation remain primary

constituents of the good life and the good society. Another – huge – challenge is to develop practices, routines, and regulations to curb the destructive powers of the new social media. These range from better education in our schools about navigating the digital world, websites that reward civil behaviour and discourage verbal abuse and mudslinging, lawsuits against people who make digital threats, and the civil courage (and patience) of individuals to partake in discussion groups made-up of people they don't a priori agree with. Finally, we need to work on the terms of a new social contract between science, technology, and society, between experts and laypersons. Or rather, on different contracts for different types of collaborations. Only if there is agreement on mutual expectations, on rights and obligations, there is a chance that the new types of expert-layperson collaborations will not end in frustration and distrust.

The main point of the reflections offered here was not to provide solutions to the problems RRI is faced with, but to make us aware that the fate of RRI is entangled with other, economic, technological, and social developments elsewhere in society. And that these other developments need attention and work, if RRI is to become a success.

Acknowledgement I want to thank Lotte Asveld for her very helpful comments on an earlier draft.

References

Borgmann, Alfred. 1984. *Technology and the character of contemporary life: A philosophical inquiry*. Chicago: University of Chicago Press.

Hirschman, Albert Otto. 1970. Exit, voice and loyalty. In *Responses to decline in firms, organizations, and states*. Cambridge: Cambridge University Press.

Jonas, Hans. 1985. *The imperative of responsibility: In search of an ethics for the technological age*. Chicago: University of Chicago Press.

Kahneman, Daniel. 2011. *Thinking, fast and slow*. New York: Farrer, Straus and Giroux.

Mazzucato, Mariana. 2013. *The entrepreneurial state: Debunking public vs. private sector myths*. London: Anthem Press.

Merton, Robert. 1973. *The sociology of science*. Chicago: University of Chicago Press.

Schomberg, Von. 2013. A vision of responsible research and innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. Richard Owen, John Bessant, and Maggy Heintz. Hoboken: Wiley.

Scott, James C. 1998. *Seeing like a state: How certain schemes to improve the human condition have failed*. New Haven: Yale University Press.

Sharon, Tamar, and Dorien Zandbergen. 2016. From data fetishism to quantifying selves: Self-tracking practices and the other values of data. *New Media & Society*. doi:10.1177/1461444816636090.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 9: 1568–1580.

Swierstra, Tsjalling. 2013. Nanotechnology and technomoral change. *Etica e Politica/Ethics and Politics* 15: 200–219.

———. 2015. Identifying the normative challenges posed by technology's 'soft' impacts. *Etikk I Praksis – Nordic Journal of Applied Ethics* 9 (1): 5–20.

Swierstra, Tsjalling, and Hedwig te Molder. 2012. Risk and soft impacts. In *Handbook of risk theory*, ed. Sabine Roesser et al., 1050–1066. Dordrecht: Springer.
Van Est, Rinie. 2014. *Intimate technology: The battle for our body and behaviour*. The Hague: Rathenau Instituut.
Van Gunsteren, Herman. 1998. *A theory of citizenship: Organizing plurality in contemporary democracies*. Boulder: Westview press.
Walzer, Michael. 1995. *Thick and thin: Moral argument at home and abroad*. Indiana: University of Notre Dame Press.

Chapter 3
Mapping the RRI Landscape: An Overview of Organisations, Projects, Persons, Areas and Topics

Job Timmermans

Abstract Increased attention in politics and academia coincided with a rapid expansion of the RRI discourse. As a consequence, the proliferation of RRI approaches and projects has made it harder to maintain an overview of the discourse. Accessing and keeping track therefore is difficult, especially for newbies, of which there are many now that RRI is being engrained in R&I policies. To untangle the RRI discourse a landscape study was undertaken providing a comprehensive overview of the main contributors, the terms RRI is perceived and the areas it is being applied to. Deploying a qualitative research methodology 536 persons emerged from the sources, affiliated to 246 organisations that reside in 89 different countries. Of these, 312 are authors and 168 involved in 18 RRI projects. Also, the study revealed 14 areas of application, 17 features and 4 methods in relation to which RRI is currently being addressed. Furthermore, the analysis shows which actors act as hubs in the different thematic or regional segments that together make up the discourse. Lastly, the study hints towards future directions of the discourse that are relevant to both policy-makers and RRI researchers. Well-represented areas and terms of addressing may be interpreted as important and therefore in need of further attention, while underrepresented areas represent opportunities for further research or justify further policy attention.

3.1 Introduction

In recent years Responsible Research and Innovation (RRI) increasingly has been gaining attention both in the realm of politics as in academia (see among others: Von Schomberg 2012; Asante and Owen 2012; Grunwald 2011; Geoghegan-Quinn 2012). As a consequence, the discourse on RRI has expanded rapidly, both in terms of the number of actors involved and substantively, in terms of the range of topics and areas of application being addressed.

J. Timmermans (✉)
Wageningen University, Wageningen, The Netherlands
e-mail: job.timmermans@wur.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_3

RRI often is understood as an umbrella term encompassing a range of existing theories and practices (Owen et al. 2012; Grunwald 2011; Fisher and Rip 2013). For example, disciplines such as ethics, innovation studies, law, and Technology Assessment (TA) have been associated with RRI. At the same time, RRI is framed differently in different local settings. For example, in the UK RRI is focussed on the impacts of R&I as an extension of traditional ethical approval (Macnaghten and Owen 2011), while in the Netherlands the MVI programme (NWO 2010) experiments with a grass roots approach funding individual responsible R&I projects. Together all these developments result in there being several, partly overlapping frameworks and approaches to RRI, which are being developed next to each other.

This proliferation of approaches and projects has made it progressively harder to maintain an overview of discourse. As a result accessing and keeping track of the discourse is difficult, especially for newbies, of which there are many now that RRI is being engrained in R&I policies such as EU Horizon 2020 (European Commission 2013). To remedy this, this contribution aims at untangling the discourse by providing a state of the art of the current RRI discourse.

Discourse is understood here 'as the body of statements, analysis, opinions, etc. relating to a particular domain of intellectual or social activity, especially characterised by recurring themes, concepts or values' (Oxford English Dictionary n.d.). Rather than analysing the concept of RRI itself, this contribution aims at providing insight into the discussion embodied by the discourse. For this purpose, a landscape study was undertaken that results in a comprehensive overview of the main (groups of) actors involved in the discourse including the main contributors, research projects and institutes. What is more, the analysis aimed at surfacing different angles and perspectives taken by the actors as well as the ways in which these are interlinked.

RRI is a term that has only relatively recently been introduced (Grunwald 2011; Siune et al. 2009). As a result, the overall discourse on RRI is still traceable for this analysis despite its rapid expansion, for example, resulting from multiple RRI projects being initiated and, not in the least, from the start of the Journal of Responsible Innovation in 2014. The chapter sets out to do just this, namely to map the overall landscape of RRI as it stood at the end of 2014 when the chapter was developed.

The chapter claims to cover the most important contributors to the RRI discourse. Importance will be established quantitatively in terms of the activity of an actor, for example, by measuring its number of articles or cross-citations, and qualitatively in terms of the topic, method and area of application of an actor.

To guide the analysis of the landscape, four main questions are proposed:

1. What actors (individuals, projects and organisations) are involved in the RRI discourse?
2. What connections exist between these actors?
3. What are the areas RRI is applied to in the RRI discourse?
4. In what terms is RRI framed and perceived by the actors?

Question 1 enables surfacing who the contributors are in the discourse, what they are working on, and where they are located. As such it becomes clear who can be considered experts per sub-area of the discourse. In addition, answering this question provides insight into certain attributes of the actors such as geo-location, size

(organisations/projects), affiliation, stage in career, and gender (persons). In this way, a further specification of the discourse overall and its contributors is provided.

Question 2 untangles how the actors involved in the discourse are connected to each other. It shows who is working with whom, and who are the best 'hubs' in the discourse, i.e. actors that connect a sub-set of others actors, and thereby are best equipped to unlock a certain section of the discourse. Additionally, answering question 2 provides insight into the cohesion or lack thereof in (areas of) the discourse.

Questions 3 provides insight into what areas RRI currently is applied to or is suggested should be applied to. On the one hand answering this question will indicate what areas are currently felt to be important to the actors in the discourse and have their attention. On the other hand, it also allows inferring what areas are not yet being addressed by the RRI community.

Question 4 aims at unveiling how RRI is framed or perceived by the actors in the discourse. It discusses what dimensions or aspects associated with RRI are deemed important based on the (amount of) attention they receive by the actors. Similar to question 3, question 4 provides insight into the topics and aspects that are currently addressed and by whom, but also enables surfacing of gaps in the current discourse that may be in need of further attention.

The chapter starts by explaining the methodology employed in undertaking the landscape study (Sect. 3.2). Following the structure of the guiding questions it then outlines the findings of the landscape study for each type of actors (individuals, projects and organisations), and for the areas, features and methods implicated in the discourse (Sect. 3.3). The chapter finishes by summarizing the main findings and drawing some conclusions (Sect. 3.4).

3.2 Methodology

To gain an in-depth understanding of the current RRI landscape and answer the guiding questions, a qualitative research methodology was deployed. This section briefly outlines the methodology in terms of data collection, data analysis and its limitations.

3.2.1 Data Collection

In order to attain a comprehensive overview of the discourse as it stood at the end of 2014 a data collection was assembled containing relevant sources available at that time. For that purpose, first a broad search was performed across major, widely used portals for academic research such as Scopus, Web of Knowledge and Google Scholar. As starting search terms 'Responsible Research and Innovation' and 'Responsible Innovation' were used as well as their abbreviations 'RRI' and 'RI'. Second, via

Table 3.1 Sources in Data Library

Type of source	Number of sources
Journal paper	76
Monograph books	5
Book chapters	82
Policy documents & reports	13
Project websites	18
Total number of sources	194

snowballing (Biernacki and Waldorf 1981) the sources collected, additional sources were discovered by consulting the references in the sources that were found already. Third, RRI project websites were found by searching the EU Cordis website,[1] by a general online search, and by snowballing the other sources collected.

Following these steps, 194 sources were found for five different source-types (see Table 3.1).

This precludes sources discussing predecessors of RRI and fields that are akin or viewed as fitting under the umbrella such as social innovation or Corporate Social Responsibility (CSR), if they do not make an explicit link to RRI in their contents, i.e. mention RRI in their body of text or are included in a book dedicated to RRI. For example, papers on 'Responsible Development' or 'Stakeholder participation' that do not explicitly link to RRI, therefore, have been excluded.

Furthermore, projects that are considered examples of implementing RRI but that do not aim at contributing to the discourse on RRI as such were not included in the data. For this reason, for example, the Dutch MVI projects (Sutcliffe 2011; NWO 2010) were excluded. Although these projects do explore ways of implementing RRI, they do not for the purpose of developing RRI conceptually or theoretically. However, many of these projects produced documents such as book chapters and journal articles that do contribute to the discourse. Via these documents findings of projects implementing RRI then were included in the source library.

Lastly, conference papers were excluded, as these are not readily accessible to the researchers. Also conference papers represent work in progress rather than established work, and in case they were further developed they were likely to have ended up in other sources such as journal papers and book chapters that were included in this research.

3.2.2 Data Analysis

Next, the data collected were analysed to discern what the relevant actors are in the current RRI discourse, how they are connected, which topics they are addressing, and in what contexts they discuss RRI.

[1] http://cordis.europa.eu, retrieved 4-2-15.

First **thematic coding** (Benaquisto, L. in: Given 2008) was deployed, using predefined themes related to the four research questions. Themes were established both inductively and deductively. On the one hand codes and accompanying rules (see Appendix I) were established beforehand based on the research questions, for example, the codes 'person' and 'organisation' were developed to answer the research question 'What actors are involved in the RRI discourse?' On the other hand, within these predefined codes, a further categorisation of recurring codes was done inductively to further reduce the complexity of the dataset.

Second, complexity was further reduced by **classification of codes**. Classification of codes involved allocating attributes to specific types of codes. For instance, the attribute <occupation> can be allocated to the code 'person'. This then would allow distinguishing groups of persons based on their occupation, for instance, 'professors' or 'lecturers'.

Third, alongside thematic analysis, and to gain further insight into the main actors contributing to the RRI discourse, a **cross-citation analysis** was conducted. The cross citation of documents within the RRI discourse provided insight into which papers are most-cited and thereby what authors were most influential in the discourse.

3.2.3 Limitations

This study has some limitations. First, the RRI discourse has only recently emerged. Although the oldest source included here was published in 2003, most sources date from the last 4–5 years. This brief time span limits the number of sources available. Especially the number of monographs (2) and reports (16) is significantly lower than the number of book chapters (83) and journal articles (75). In addition, actors may not have had the opportunity to establish themselves within the discourse, for instance, by establishing connections to other actors. Vice versa, the group of actors that do feature in multiple sources tends to figure prominently in the different analyses of the landscape study. The Delft University of Technology and the University of Exeter or persons affiliated to these organisations, for instance, come forward in the analyses of persons, projects, organisations and publications. However, the 'usual suspects' reappearing across the analysis is not just the result of the novelty of the discourse; these actors indeed are at the centre of discourse as it currently is emerging.

Second, for the FaRin and SYNERGENE projects, it proved hard to ascertain what persons are involved in the project. As a result members of these projects are not included in the analysis of persons.

Third, some of the edited volumes included in the analysis involved actors (for instance organisations and authors) and themes (for example areas of application, features and methods) that re-occurred in multiple chapters of one volume. Since the analysis treats every chapter as an individual source, these edited volumes may cause certain themes and actors to be over-represented in the data. To counter this,

when possible, this bias is considered in the discussion of the findings. Also, the cross-citation analysis neutralises these effects as it only considers the best-cited authors and not the most productive (in terms of the number of publications).

Fourth, as stated above (see Sect. 3.2.1) the delimitation used for this investigation excludes certain types of sources such as conference papers, and literature from adjacent fields that lack an explicit reference to RRI. Furthermore, the investigation limits itself to what has been stated in the sources. For instance, if members of different projects have a professional relationship that does not appear in the sources, that connection will not arise from the analysis.

3.3 Landscape Study

To sketch the current landscape of RRI in this section the findings are synthesised and re-structured accordance to the four guiding questions. Each question is addressed in a separate sub-section.

3.3.1 Actors

The first guiding question addresses what persons, projects and organisations are currently involved in the discourse on RRI. Each actor type is discussed below.

In total 536 different persons are involved in the discourse either as an author and/or as a project member at the time of the investigation. In terms of its size, the discourse only really took off in 2009, when the number of persons involved grew from 2 to 28 and has doubled almost every year since (see Table 3.2).

This is mirrored by the growth of the total number of documents published, which has risen from 5 to 182 in the last 4 years. Important causes for this increase are the start of multiple RRI projects funded by the EU and national funding bodies as of 2010, and the start of a journal specialising in RRI in 2014. Moreover, in 2014, the most productive year thus far, 45 book chapters and 47 journal articles have been published. The majority of reports on RRI (as opposed to journal articles and book chapters), however, have been published in 2011 to 2013, so before 2014. In terms of its volume, the academic discourse (journal articles plus book chapters) on RRI lags behind the publication of reports (which are mostly policy related). When reports being published is an indication of a new policy being put into place (four reports by the European Commission have been included in this analysis), policy on RRI may be a factor contributing to the rise of attention for RRI in academic publications. In addition, this development hints that the focus of RRI, over time, has shifted from policy makers to members of academia. A possible reason for this could be that policy was felt to be bearing fruit and therefore could do with less attention or alternatively, because priorities of policymakers have moved away from RRI.

Table 3.2 Persons involved in RRI discourse over time

Year	Growth	Total
2003	1	1
2004	0	1
2005	0	1
2006	1	2
2007	0	2
2008	0	2
2009	26	28
2010	15	43
2011	62	105
2012	35	140
2013	171	311
2014	225	536

Not surprisingly, RRI being both concerned with research and the focus of research, the RRI discourse is dominated by researchers (87%) and research support (an additional 2%). Persons working in policy (1%) and business (10%) are considerably less represented in the discourse (see Fig. 3.1).

This is paralleled by the organisations that persons are affiliated to: the majority are research related with 53% being universities. Only 9% of the organisations are business related. Despite its focus on economic growth (see for example Geoghegan-Quinn 2012) and creating marketable products and services (the innovation in RRI is understood to include reaching out to markets, see e.g. (Von Schomberg 2012; Owen et al. 2012; Zwart et al. 2014), at the moment RRI mainly involves (semi-) public, non-profit/not for profit organisations. Organisations that have experience or are engaged in innovation in terms of actually creating marketable products and services are underrepresented.

One of the six pillars of RRI proposed by the EU commission (see e.g. Geoghegan-Quinn 2012) is gender. On the junior and mid-level career-stages the involvement of both sexes in the discourse is well-balanced (see Fig. 3.2).

In the senior positions, however, this is rather different with two-thirds being men. This imbalance is accounted for entirely by the research community, as both in business and policy more women than men are represented. If gender is regarded as one of the cornerstones of RRI, having a balanced involvement of persons from both sexes might be called for at all career levels. At the moment this is not the case, especially in academia which by far is the largest contributor to the discourse.

The persons involved in the discourse are affiliated to 246 different organisations. These organisations are based in 89 countries in six contents across the globe (see Fig. 3.3).

The discourse is dominated, however, by the Western, especially European, countries, which account for over 80% of the organisations. So, although RRI has a global outreach, certainly in its intentions and aims, the discourse largely is a European affair. What is more, within Europe organisations are mainly based in EU member-states with the UK, the Netherlands, Germany and Italy, which account for 50% of the organisations.

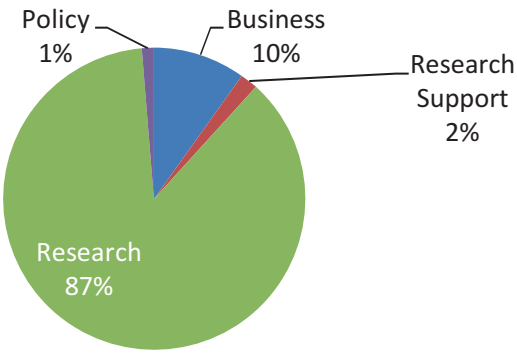

Fig. 3.1 Percentage of persons in RRI discourse per occupation category

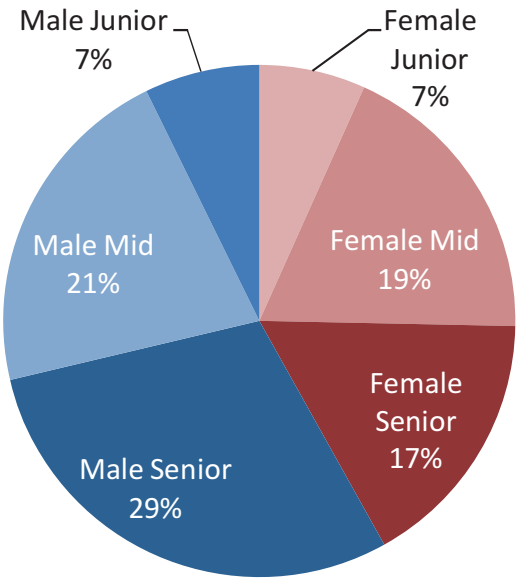

Fig. 3.2 Gender balance of persons involved in RRI discourse per career-level

Projects are a major contributor to the discourse. By means of 17 RRI projects, 260 persons are involved in the discourse affiliated to 26 different organisations. The size of projects ranges from small (3–15 persons) to mid-size (15–40) and large (beyond 40 persons). On average projects take three to 4 years, which means that for the duration of a project bonds are forged between the persons within it. The number of persons per organisation in a project typically ranges between one and three. Most projects involve multiple organisations and therefore are a means for connecting persons across different organisations. This becomes apparent in the networks of connections between persons on the one hand and organisations on the other that span the RRI discourse (see Sect. 3.3.2 below).

Pockets of mutually connected persons/organisations that come forward within the overall network can be accounted for by the involvement of actors in the RRI

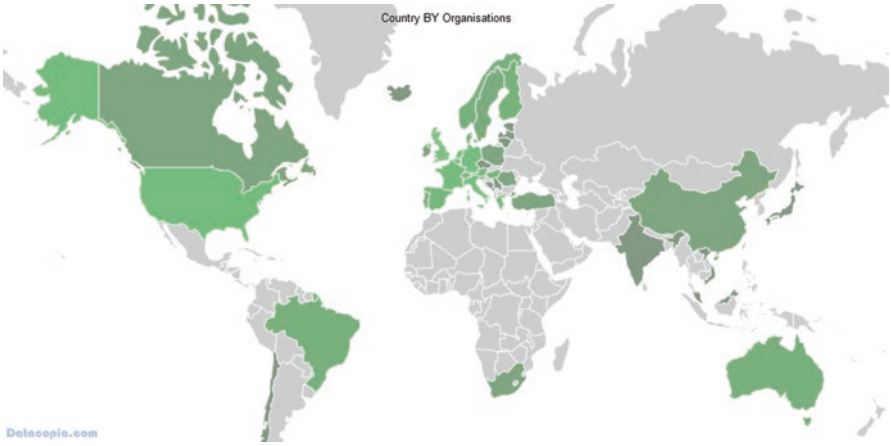

Fig. 3.3 Number of organisations per country (*lighter tone of green* = higher number of organisations)

projects. For instance, the SATORI project connects 28 persons from 15 different organisations, forming a grid of 27 interconnected persons and one of 15 interconnected organisations.

What is more, some persons are involved in multiple projects, either successively or simultaneously. These persons function as 'hubs' between the different projects, not only connecting the different projects but also connecting the different organisations and persons involved in those projects. For instance, after the project had ended, together the members of the ETICA project joined 9 of the successive RRI projects. This way Catherine Flick, for example, acts as a hub between the three RRI projects she was/is a member of, connecting 40 persons. Projects therefore are an important vehicle in organising discourse. They provide the connections which link the different pockets of inter-connected persons and the organisations they are affiliated to across the discourse.

The same applies to publications. Co-authoring an article or book-chapter also is regarded as forging a bond, or even more so than with projects, is understood as displaying an existing connection between persons and their organisations within the discourse. Of the 536 discourse-members, 312 (co-)authored one or more of the 182 documents that have been published on RRI and included in this analysis. With one or two exceptions the number of persons (and organisations) involved in creating a document is much smaller than that of projects, typically ranging between 1 and 6. If at all, documents, therefore, connect a much smaller number of persons than projects do, therefore creating clusters of interconnected persons in the actor networks that are much smaller than those created by the average project.

In terms of activity (measured by the number of project memberships and co-authored documents) 70% of the 246 organisations are on a par, having one or two persons involved in the discourse either as a project-member or co-authoring one document. Only the top 8% of most active organisations in the discourse are linked

to four or more projects/publications and/or have more than five persons involved. Being involved in 50% of all sources (projects/documents) and accounting for 42% of the persons involved in the discourse, these 19 organisations constitute the core of the RRI discourse.

Thirteen organisations in this group are universities, with the Delft University of Technology, the De Montfort University, and the universities of Exeter, Arizona State and Szeged being most prominent. Only four organisations are not universities, namely the Rathenau Institute, the Karlsruhe Institute of Technology, VTT and Fraunhofer ISI. These organisations, however, also are (semi-)public research institutes, which corroborates the finding that the RRI discourse mainly is a non-profit (non-business) affair.

For prominent and less prominent organisations alike, the number of persons affiliated to an organisation more or less matches the number of sources they are involved in. Exceptions to this are the De Montfort University and the University of Exeter and to a lesser extent the University of Namur and the Rathenau Institute, which account for significantly more sources than persons. This seems to indicate that these organisations harbour persons that are more active than the average contributor to the discourse.

Generally speaking, the majority of the most active persons in the discourse are affiliated to one of the 19 most active organisations. Of the 536 persons contributing to the RRI discourse, 85% either is involved in one project or has (co-)authored one document. Only 15% is associated with two or more projects and/or documents. What is more, the top 3% most active persons are associated with 50% of all sources. The centre of the discourse in term of activity is formed of Richard Owen, Bernd Stahl, Jack Stilgoe, Armin Grunwald, Jeroen van den Hoven, Phil Macnaghten, David Guston, Grace Eden, Neelke Doorn, Marina Jirotka and René von Schomberg. Except for Neelke Doorn and Grace Eden, the prominence of these persons is corroborated by the fact that they all are among the more influential authors on RRI.

When the number of cross-citations is regarded as a means to measure influence on the discourse, largely the same group of persons emerges as most influential, namely Owen, Macnaghten, Stilgoe, von Schomberg, Guston, Grunwald, and Stahl (see Table 3.3).

However, also persons that do not come to the fore in terms of their activity are singled out as influential in terms of citations, for example, Fisher, Goldberg, and Sutcliffe. So there is only a small group of persons that stands out in both activity and number of cross-citations, while a larger segment of persons only is distinguished in terms of one of these two measures.

3.3.2 Connections Between Actors

The second guiding question addresses what connections exist between the different actors that are part of the RRI landscape. Through co-occurrence of persons (and the organisations they are affiliated with) across projects as project-members and documents as co-authors, connections are established between persons, projects and

Table 3.3 Cross citation index compared to Google Scholar index per author

Author	Cross	Google
Owen, R.	87	314
Macnaghten, P.	53	218
Stilgoe, J.	52	188
Von Schomberg, R.	40	147
Fisher, E.	29	57
Guston, D. H.	27	59
Gorman, M.	21	40
Goldberg, N.	14	55
Grunwald, A.	11	20
Stahl, B.C.	8	64
Asante, K.	8	3
Williamson, G.	8	2
Maynard, T.	6	33
Baxter, D.	6	29
Depledge, M.	6	29
Rip, A.	5	13
van der Burg, S.	5	6
Sutcliffe, H.	4	26
van den Hoven, J.	3	22
Hellström, T.	1	56
Amanatidou, E.	1	29
Cagnin, C.	1	29
Keenan, M.	1	29

organisations. This way networks of connected actors are formed that encompass the majority of the actors involved in the discourse.

Of the 17 RRI projects, 6 projects are not connected to any of the other projects (see Fig. 3.4). For two of these projects, this can be explained by a lack of persons found in the sources that are included in this study, namely the PIER and FaRin projects. So no conclusions can be drawn concerning their connectedness to other actors.

Although all their project members have been included in the source materials, no connections with other projects emerge from the data for the KARIM, IRRESISTIBLE, ResAgora, RI Framework, and New Technologies as Social Experiments projects. The two projects mentioned last are the only two of the six projects that are non-EU funded, which may account for their lack of out-bound connections. However, so is FRRIICT, and that project is well connected to other projects. So, not being funded by the EU does not necessarily has to be a reason for a lack of outbound connections. Moreover, the Delft University of Technology, which is the sole organisation behind the 'New Technologies as Social Experiments' project, is involved in other projects on RRI. Furthermore, all of these six projects have a relatively small number of persons involved compared to the other 11 projects. Having fewer persons and organisations involved therefore seems to decrease the chance of one or more of these actors also being involved in other projects and thereby of the projects being connected.

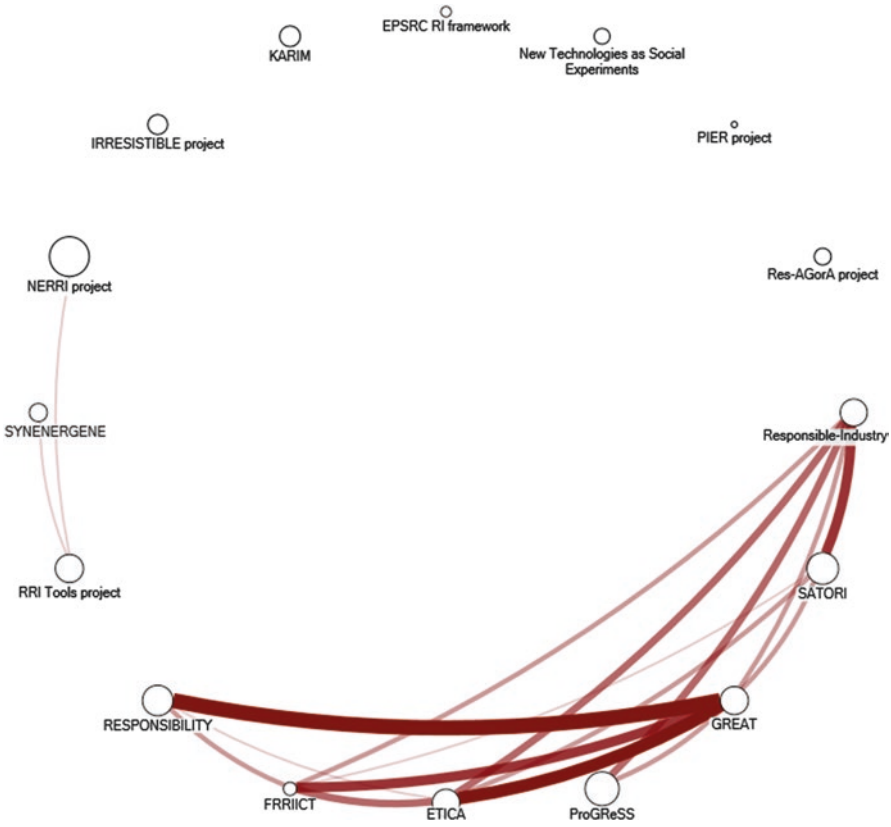

Fig. 3.4 Connections between RRI projects

Through their involvement in projects and documents, 220 of the 246 organisations that are active in the discourse are connected to one or more other organisations. Together these organisations form an intricate network of 199 organisations surrounded by 21 organisations that are less well connected (see Fig. 3.5).

Rather than depicting individual elements, the purpose of these network representations is to unveil the overall social structure of the discourse. The network shows a structure of 18 clusters of organisations that within their cluster all are connected to each other. In turn, these clusters are connected to other clusters by one or more of the organisations within that cluster being connected to organisations outside their cluster (the 'hub'-organisations). A total of 36 organisations acts as the main hubs that account for connecting the clusters of organisations they are connected to, to the overall network.

Most of these hubs are located in the Netherlands (7) and the UK (9), followed by Denmark and France (3 each). This is not surprising as these are all EU member states (so eligible for EU funding) and also emerged among the countries most active within the RRI discourse. Hub organisations can be divided into two types:

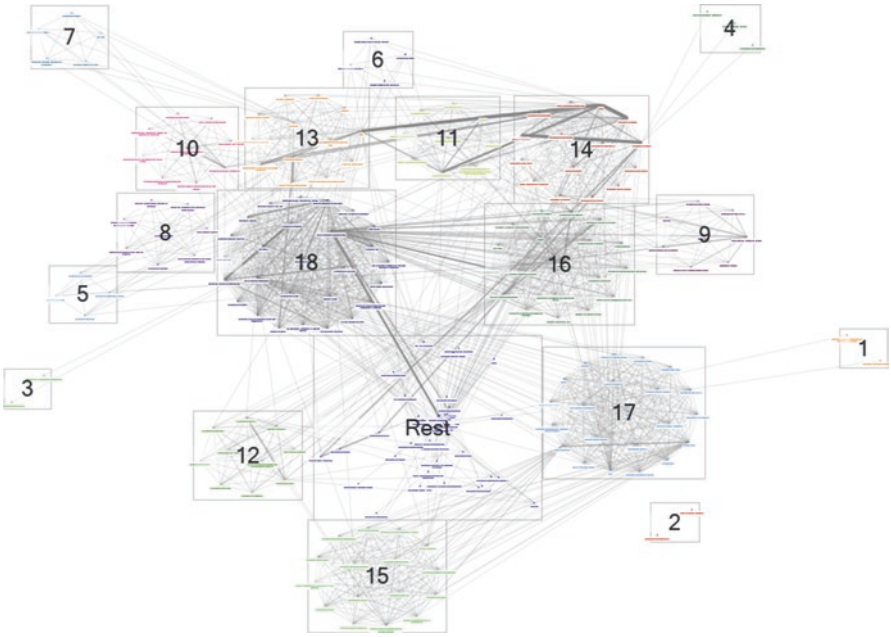

Fig. 3.5 Overview of clusters of organisations within RRI discourse

hubs that connect a great number of organisations (15 up to 72), and hubs that while connecting a smaller number of organisations (4–8) establish connections to clusters that otherwise would not be connected to the network. Both of these types of hub organisations are important to the discourse because they account for glueing the network together.

At the level of persons, similar network-patterns occur but in a more fine-grained way. Of the 536 persons that are active in the discourse, 501 are connected to at least one other person. The 25 non-connected persons are single authors of documents, which do not appear in any of the other sources (documents and projects) beyond that document. The 501 persons are grouped into 56 clusters of persons that all are connected to each other, based on their co-occurrence in documents and projects (see Fig. 3.6).

These clusters range in size from two up to 52 persons. Together these connected persons span a network of clusters without a real centre: there is no single person that is connected to all other persons. Instead, there are 28 persons that function as hubs connecting the different clusters of persons. Similar to hub organisations these persons either connect a wide range of other persons, or account for connecting a small number of persons uniquely.

When their activity, number and uniqueness of their connections, and their cross-citations are taken into account 31 persons emerge from the data as most influential in the discourse on RRI (Table 3.4).

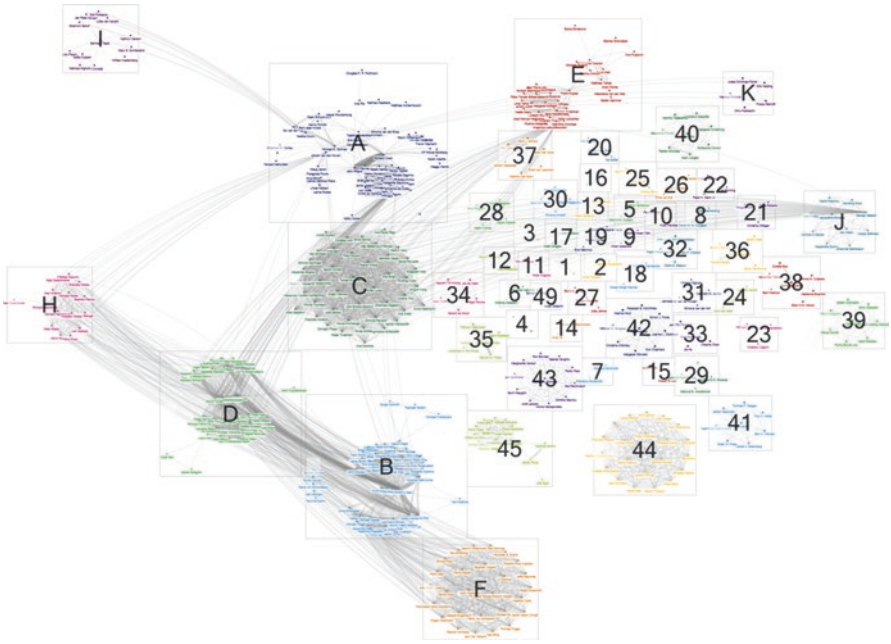

Fig. 3.6 Persons connected to persons in 56 clusters (*A–J* and *1–46*)

Similarly, taking activity, number and type of connections, and cross-citations into account 27 organisations emerge from the data as most influential in the discourse on RRI (Table 3.5).

The 31 key persons are affiliated to 22 different organisations. Of these organisations 13 also emerge from the data as one of the 27 key organisations. Again the EU member-states, and especially the Netherlands and Germany emerge from the data as most implicated host-countries of key organisations in the discourse.

3.3.3 Areas Implicated

The third guiding question addresses what areas of application are implicated by the actors in the RRI discourse. In total 14 different areas emerge inductively from analysis of the source material.[2] Of these, ICT, health, business, Nano, and Bio are addressed most often across all types of actors, followed by climate/environment,

[2] Establishing (exact) borders between areas is notoriously hard if not impossible. To assure the relevance of the current classification it was mapped onto the areas forwarded by the EU H2020 portal (ec.europe.eu/horizon2020/en/find-your-area. Accessed 7-12-2015). This showed that apart from business, finance, genetics and engineering, which were not mentioned at all by the EU, the areas found in this investigation map onto those on the portal.

Table 3.4 Top 28 key persons in the RRI discourse (Based on activity, connections & cross-citations)

	Person	Affiliated organisation
1	Bernd Stahl	De Montfort University
2	Richard Owen	University of Exeter
3	Phil Macnaghten	Durham University
4	Jack Stilgoe	UCL
5	Jeroen van den Hoven	Delft University of Technology
6	Armin Grunwald	Karlsruhe Institute of Technology
7	Michael Gorman	University of Virginia
8	Doris Schroeder	University of Central Lancashire
9	Marina Jirotka	University of Oxford
10	Philippe Goujon	University of Namur
11	Catherine Flick	De Montfort University
12	Veikko Ikonen	VTT
13	René von Schomberg	European Commission
14	David H. Guston	Arizona State University
15	Erik Fisher	Arizona State University
16	Simone van der Burg	Radboud University Nijmegen
17	Behnam Taebi	Delft University of Technology
18	Benjamin Schrempf	European Academy of Technology and Innovation
19	Christopher Coenen	Karlsruhe Institute of Technology
20	Frank Kupper	Free University Amsterdam
21	Ibo van de Poel	Delft University of Technology
22	Michael Obach	Fundación TECNALIA Research & Innovation
23	Miltos Ladikas	University of Central Lancashire
24	Neelke Doorn	Delft University of Technology
25	Petra Ahrweiler	EA European Academy
26	Sheena Laursen	Experimentarium
27	Tsjalling Swierstra	Maastricht University
28	Xavier Pavie	University of Paris Ouest

emerging technologies, and country/region. Least addressed are energy, education, finance, food, engineering, and genetics. Of the sources analysed 46% address one area of application, 10% two, and 3% more than two. The majority of projects and documents analysed either have no or one area of application. This means that most sources are either focussed on RRI in general without a specific area of application, or on one specific area of application such as ICT or Nano. In sources that implicate multiple areas, most frequently these are combinations of the most addressed areas: ICT with health, Nano with business, Nano with Bio, or Nano with health. Rather than broadening the area of application, these combinations seem to further narrow the area of application, thereby delimitating the discussion of RRI within a source.

Table 3.5 Top 27 key organisations in the RRI discourse (Based on activity, connections & cross-citations)

	Organisation	Country
1	Delft University of Technology	NL
2	VU University Amsterdam	NL
3	De Montfort University	UK
4	Fraunhofer ISI	GER
5	University of Padova	IT
6	University of Oxford	UK
7	VTT	FIN
8	University of Twente	NL
9	Lancaster University	UK
10	University of Exeter	UK
11	Tilburg University	NL
12	University of Central Lancashire	UK
13	University of Namur	BE
14	Radboud University Nijmegen	NL
15	Durham University	GER
16	Arizona State University	USA
17	Leiden University	NL
18	Ciência Viva	POR
19	Experimentarium	DEN
20	UCL	UK
21	University of Southern Denmark	DEN
22	Valahia University Targoviste	ROM
23	European Union/Commission	EU
24	London School of Economics and Political Science	UK
25	University of Szeged	HUN
26	Karlsruhe Institute of Technology	GER
27	Utrecht University	NL

All 17 projects have at least one area of application that they address. Two projects focus on more than four areas and therefore can be considered to having a broad scope, while the other projects focus on one or two areas and therefore have a more specific focus. The areas that receive most attention across projects are health and ICT. Besides finance, food and genetics, which that are not being addressed by projects, the other 12 areas are distributed evenly among the projects.

Of the 246 organisations, 146 address one or more areas of application. Of these, only 14 organisations target more than one area. These organisations either have a more specific focus than average when several areas are being addressed together in the same source or a broader involvement in the discourse when these areas stem from multiple sources. While all of the 14 areas that emerge from the investigation are targeted by at least three organisations, the areas ICT, bio, education, country/region, and health are targeted most frequently by organisations (see Table 3.6).

Of 536 persons involved in the discourse, 332 are associated with one or more areas of application of RRI. Again, the areas are evenly distributed among the persons, except for food and engineering that are associated with persons less than average, and health, Nano, Bio, and country/region and ICT that are implied more often. When the frequency by which an individual is associated with an area of application is considered as a measure of expertise, then ICT (6) and Nano (8) have the greatest number of experts in the discourse followed by finance and climate.

The total number of organisations associated with an area relates to the number of persons with the same area by a ratio of two persons to one organisation on average (see Table 3.6). Areas that receive most or least attention in the discourse thus do so both in terms of the number of persons and organisations. With regards to the relation between the number of documents and the number of persons, there are a few exceptions to this fixed ratio, namely health, emerging technologies and finance. Health is less frequently discussed in documents in relation to RRI than the total activity by actors in that area would predict, while emerging technologies and finance have received more attention than the number of persons and organisations would predict. A possible explanation for the divergence of health is that this area has received its attention in projects rather than in publications. As for finance and emerging technologies, it seems to be the case that the persons addressing these areas are more productive than the average person in the discourse.

Overall the table shows that in the discourse RRI often times is being applied to, or being discussed in relation to areas that are considered to be contested such as Nano, Bio, ICT and genetics. However, some of the areas that have been addressed less frequently by the discourse, such as energy, climate/environment and finance, also have been the topic of public attention over the last few years, and therefore must be considered societally relevant. For example, the area 'finance' has been the

Table 3.6 Total number of persons, organisations, projects, and documents per area of application

	Area	Persons	Organisations	Projects	Documents
1	Health	122	52	4	17
2	Nano	71	39	2	27
3	Bio	68	36	1	16
4	ICT	64	36	4	17
5	Business	63	32	2	16
6	Country/region	44	25	2	8
7	Climate/ environment	24	21	2	7
8	Education	23	18	2	1
9	Energy	23	11	1	1
10	Emerging technologies	15	8	0	9
11	Finance	13	6	0	6
12	Genetics	11	4	0	2
13	Engineering	7	3	1	2
14	Food	5	3	0	3

topic of debate related to the global credit crisis while energy, food and climate/environment have had public attention related to global warming.

Both the areas of ICT and health, which receive much attention in the discourse, have a strong tradition as a topic of (applied) ethics. Because ethics is considered to be one of the main contributors to the concept of RRI (see, for example, Grunwald 2011 and Stahl et al. 2013), and is the most addressed feature in the discourse (see Table 3.8), it comes as no surprise that these areas also have crossed over to the RRI discourse.

Besides in relation to a specific issue or technology, RRI also is being discussed in relation to a particular discipline or geographic area. Business, for example, which discusses RRI in relation to the market, such as what it means for entrepreneurship, corporations or SMEs, is among the most frequently addressed areas. Likewise, the area of engineering, rather than focussing on an issue or technology, discusses what RRI means for this discipline. Finally, RRI is also being discussed in relation to specific areas or regions, for instance, Brazil or Northern Vietnam. The fact that RRI is discussed in relation to a specific country or region not only shows that RRI gains traction in different geographic areas and contexts (see Fig. 3.3), but also that RRI may have different implications when it is implemented in different geographic contexts.

3.3.4 Terms of Addressing

The fourth guiding question discusses in what terms RRI is addressed by the actors involved in the discourse. Overall RRI is framed and perceived by contributors to the discourse in terms of 9 different requirements, 4 different issues and 8 different means (see Table 3.7). Requirements represent demands to the R&I process to render it 'responsible'. For example, 'anticipation' requires the R&I process to anticipate future (societal) consequences, while 'engagement' requires the engagement of stakeholders such as the general public into the R&I process. Issues represent normative or societal concerns that could or should be addressed when implementing RRI, for example, 'gender' or 'privacy'. Means represent different ways to implement or support implementing RRI, such as 'education', 'tool' and 'governance'. Each type of framing is discussed below.

Of the requirements, 'inclusion', 'engagement', 'anticipation' and 'participation' are addressed most frequently across all source types. These requirements thus come forward as most significant for rendering R&I processes responsible and hence may be understood as core requirements. This is not surprising as these aspects have been discussed as such by influential accounts of RRI, for instance, by Owen et al. (2013), von Schomberg (2011) and Sutcliffe (2011). However, also responsibility is regarded as a core concept in the theoretical discussions of RRI. Not only is it a component of the term RRI itself, it moreover has been suggested that its focus on responsibility is what distinguishes RRI from its predecessors (see Owen et al. 2012; Stahl et al. 2013; Grunwald 2011). In terms of numbers of persons and organisations addressing it, however, 'responsibility' does not receive as much attention from the sources as its importance to RRI would suggest. Of the four remaining requirements, 'interdisciplinarity' and 'deliberation' also emerge as sig-

Table 3.7 Type and Total number of persons, organisations, projects, documents and sources per term

Term	Type	Persons	Organisations	Projects	Documents	Sources
Sustainability	Issue	25	17	2	12	14
Privacy	Issue	19	16	0	7	7
Gender	Issue	1	1	0	4	4
Risk	Issue	8	6	0	3	3
Ethics	Means	127	65	5	46	51
Approach	Means	136	58	2	41	43
Governance	Means	116	59	2	34	36
Policy/ Regulation	Means	79	47	2	23	25
Framework/Account/Model	Means	78	50	7	17	24
Tool	Means	74	43	5	11	16
Education	Means	66	57	3	9	12
Knowledge transfer	Means	21	16	1	4	5
Anticipation	Requirement	79	36	2	24	26
Responsibility	Requirement	37	30	1	24	25
Engagement	Requirement	87	57	4	17	21
Participation	Requirement	77	48	5	10	15
Inclusion	Requirement	111	49	3	8	11
Deliberation	Requirement	22	16	3	8	11
Interdisciplinarity	Requirement	30	14	0	9	9
Capacity	Requirement	3	10	1	2	3
Openness	Requirement	3	2	0	3	3

nificant although they do not appear part of the core requirements, while ‘capacity’ and ‘openness’ have only received relatively little attention by the discourse.

Issues that are in need of addressing by RRI are much less frequently discussed among the sources than requirements. Topping the list, ‘sustainability’ for example, which has been brought forward by different accounts on RRI as an important motivation to engage with RRI (see Von Schomberg 2011; Sutcliffe 2011), only is addressed by a small segment of the discourse (25 of the 536 person affiliated to 17 of 246 possible organisations). Discussion of RRI by the discourse thus tends to be in generic terms rather than focussed at specific (types of) issues. Especially projects notably do not target any specific issues: except for two projects that address ‘sustainability’, no further issues have emerged in relation to any of the projects. Least addressed is the issue ‘gender’. It is only addressed by one organisation, the European commission, who, as is touched upon earlier, have included it as one of the pillars of RRI. Thus far, however, it has not been picked up by any of the other actors in the discourse. It might be of interest to investigate whether the male dominance in the senior (research-) positions in the discourse (see Fig. 3.2 above) is a factor in the neglecting of gender by the discourse.

Lastly, of the means to implement or support implementing RRI, ‘ethics’, ‘approach’ and ‘governance’ emerge as most significant. This is not unexpected as RRI is being framed as a novel governance approach (see e.g. Fisher and Rip 2013; Von Schomberg 2012), which incorporates existing approaches such Technology Assessment (TA), Privacy by Design, and Socio-Technical Integration Research (STIR) under its umbrella (see e.g. Grunwald 2011). What is more, ethics, as outlined above (Sect. 3.3.3) is considered to be one of the main contributors to RRI. This significance of ethics to RRI is corroborated by its emergence as an independent term in the analysis distinct from other approaches which have been grouped under ‘approach’. Also ‘Policy/regulation’ is frequently discussed by the discourse. This can be accounted for by its kinship to RRI as governance, but also signals the importance of R&I policy as a driving force behind the discourse.

In comparison with the other means, ‘framework/ account/ model’ (7 projects), ‘tool’ (5) and ‘education’ (3) are, relative to overall the number of sources addressing these terms, frequently addressed by projects. What these three terms have in common is that they represent means that are specifically geared towards supporting the implementation of RRI. ‘Framework/account/model’,[3] for example, encompasses the model of RRI offered by von Schomberg (2011), a European normative model for RRI (ProGress-project) and a framework for RRI in ICT (FRRIICT-project), while the term ‘tool’ among others includes an observatory and a governance framework.

[3]Among ‘Methods and Features’ a distinction is made between ‘Approach’ and ‘Framework/ account/ model’. In the analysis the code ‘Approach’ is used to represents methods or theories that are being discussed in the sources as a way to implement RRI, for instance, Technology Assessment (TA), Privacy by Design, and Socio-Technical Integration Research (STIR). In contrast the code ‘Framework/ account/ model’ is applied where a description of RRI itself or (a set of) features/ dimensions of RRI is given by a source. For example, the definition of RRI offered by von Schomberg (2011), a European normative model for RRI (ProGress-project) or a framework for RRI in ICT (FRRIICT-project).

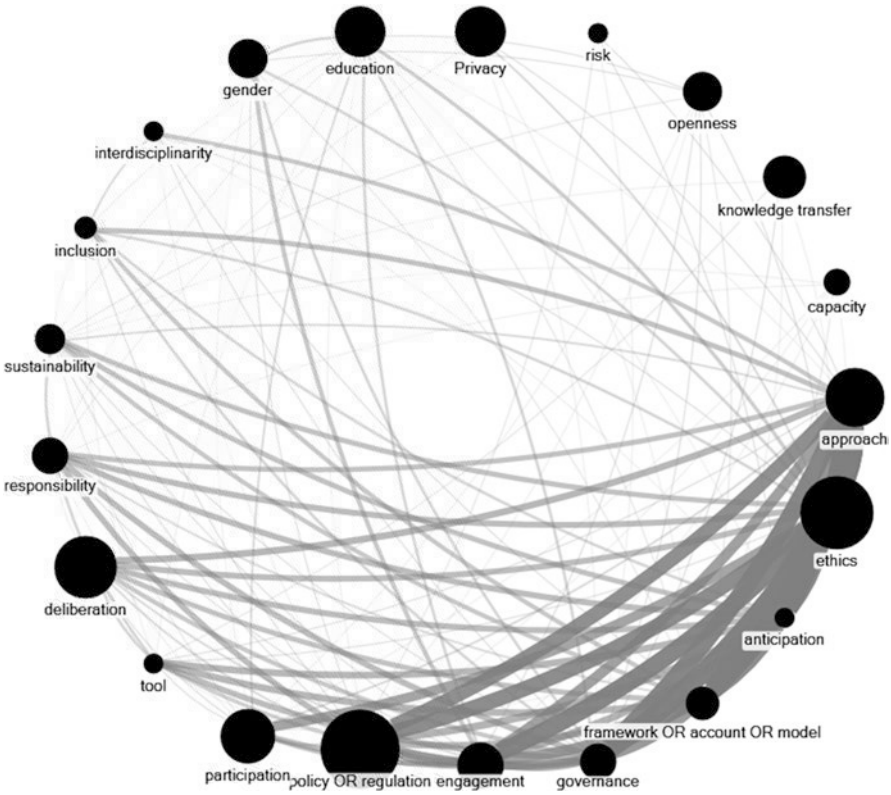

Fig. 3.7 Connections between terms of addressing RRI

So, compared to documents which predominantly discuss means of addressing RRI in broader, theoretical terms such as approaches and governance, many projects (also) have a more practical perspective on RRI providing hands-on support to actors implementing RRI.

In almost all the sources included in the analysis, the different terms do not appear by themselves but in concert with other terms. Fig. 3.7 is a graph that represents the different terms (edges), and the interconnections between them within the discourse (edges). The size of the edge represents the number of sources addressing the term (ranging from 5 to 61). The strength of the connection (the number of co-occurrences across sources, ranging from 2 to 38) between terms is represented by the width and opacity of edges.

Besides ‘capacity’ all terms are connected to one or more of the other terms. At the bottom of the graph, the terms are represented that have the strongest connections to other terms. Most of these also are among the most frequently addressed terms by the discourse (larger edge). Apart from with each other, based on the strength of their connections, the most addressed terms in the discourse ‘ethics’, ‘approach’ and ‘governance’ are most frequently discussed together with ‘framework/account/model’, ‘policy/regulation’, and ‘anticipation’.

Table 3.8 Organisations connected to 3 or more terms

Organisations	Number of terms
Delft University of Technology	9
University of Exeter	8
De Montfort University; European Commission	7
Arizona State University	6
Durham University	5
Karlsruhe Institute of Technology	4
University of Twente; University of Oxford	3

On average a member of the discourse focuses on one or two requirements, means and/or issues. The exceptions to this are the most active persons in the discourse who address more than two terms in their perception or framing of RRI. Being involved in multiple projects and/or authoring multiple documents, therefore, is a predictor of the depth/breadth of the angle of the involvement of persons in RRI.

Organisations that are found to be most active in the discourse also have the most diverse involvement with RRI in terms of features and methods. For example, Delft University of Technology, De Montfort University and the Universities of Exeter, Arizona and Durham stand out both in terms of activity and in being associated with features and methods (see Table 3.8). Here, breadth is therefore linked to depth and diversity.

3.4 Conclusion and Recommendations

To help to untangle the rapid expanding discourse a landscape study was undertaken mapping the state of the discourse on RRI as it stood at the end of 2014. Based on this study, 13 overarching characteristics of the discourse and related recommendations can be deduced:

Exponential Growth Since 2009
As of 2009, the number of persons has doubled every year, amounting to 526 persons in 2014 affiliated with 246 different organisations and residing in 89 different countries. Of these 168 persons are involved in the 18 projects on RRI that have been undertaking since 2009. Moreover, 312 persons are (co-)authors of documents. Starting in 2003, 16 reports, 75 academic journal articles, 83 book chapter in 6 edited volumes, and 2 monographs on RRI have been published.

Policy Push, Rather Than Market Pull
The majority of publications on RRI are academic papers rather than policy reports. However, most of the policy documents (and the papers by Von Schomberg who is EU policy officer) preceded the bulk of the academic work. Moreover, over 90% of RRI activity took place in publicly funded institutions, the majority of which resided within EU countries and countries that are known to fund RRI activities such as the

Netherlands and the UK. This gives rise to the assumption that RRI not so much is motivated by market-demand of R&I actors such as researcher and innovators, but rather has been pushed by policy-makers.

Western Affair
With over 90% of the organisations involved coming from the US, Australia but foremost Europe, RRI predominantly is a Western affair. Moreover, when regions in third-world countries are implicated as an area of application this is done by actors from first-world countries.

Conceptual Rather Than Practical
Despite calls for bridging of RRI theory and practice (Stilgoe et al. 2013; Stahl et al. 2013), RRI still is chiefly discussed conceptually in terms of frameworks and approaches rather than practically in terms of tools and knowledge transfer. The requirement to implement RRI, for example by researchers on the EU H2020 programme, warrants more attention for addressing RRI in practical terms.

Activity Dispersed Although Dominated by a Core
While the majority of the organisations and persons active in the discourse are implicated in only one or two sources, a core of 31 persons and 27 organisations accounts for over 50% of all activity in terms of publications, project membership and acting as hubs connecting the discourse. This core group dominates the discourse both in breadth, depth and diversity of terms used and areas of application.

Fragmentation Across Network, Cohesion Within Clusters
When interrelatedness of actors in the discourse is concerned, a two faced picture emerges: on the one hand the discourse encompasses clusters of interconnected actors (interconnected projects, organisations and persons), but on the other hand not all of these clusters of actors are connected to the other clusters. Clusters may consist of one to up to 50 actors, depending on the type of actor (there are only a limited number of projects, so clusters of projects contain fewer actors than large clusters of persons or organisations). As a result, the discourse displays cohesion within clusters and across a number of connected clusters whilst at the same time, it is fragmented, as there are clusters of actors that are not or only weakly connected.

Public, Non-profit Rather Than Business/Industry
Rather than involving them as a partner, business and industry are being addressed by the discourse as areas of application. The Responsible-Industry project, for example, addresses the application of RRI in industry but does not include industry partners although three partners are non-profit organisations representing industry. Developments since the landscape study took place, however, may help remedy this imbalance. In a new call for proposals in spring 2016 the Dutch MVI programme, for example, requires the involvement of industry partners in the research projects (NWO 2016).

Three EU RRI Keys Widely Covered, three Underrepresented
Of the 17 terms used to address RRI that emerged from the analysis, the three EU keys (Geoghegan-Quinn 2012) Ethics, Governance (and related terms Policy and Regulation) and Engagement (and related term Inclusion) are most widely covered.

At the same time the EU keys Education, Openness and especially Gender are underrepresented by the discourse. With RRI being an overarching theme of the EU research funding programme H2020 (European Commission 2013), further attention by the discourse for these three keys is recommended.

Gender Imbalance
Overall, 14% more men are involved in the discourse than women. However, while the discourse is well balanced for the junior and mid-career levels, at the senior level there is a male over-representation with a ratio of two to one. At this level, the discourse, therefore, does not live up to the EU RRI keys. Besides in EU documents, gender is not being addressed by the discourse. It is worth investigating to what extent the gender imbalance at the senior level offers a reason for this lack of addressing.

Anticipation and Sustainability as Additional EU RRI Keys
Of the most addressed terms, Anticipation and Sustainability currently are not recognised by the EU as an RRI key. As they emerged from the analysis as being of relevance it, therefore, is recommended these terms are considered as further keys.

Responsibility Underrepresented
The concept of responsibility is portrayed as a cornerstone of RRI, setting it apart from its predecessors and connecting the myriad of approaches and theories under the RRI umbrella (Stilgoe et al. 2013; Grunwald 2011). Although it did receive attention by the discourse, this is not as much as would be expected based on the most addressed terms such as ethics and governance. Further attention therefore is recommended.

Contested Rather Than Societal Relevant Areas
Of the 14 areas RRI is being applied to, (ethically) contested areas such as ICT, Bio and Nano are most implicated while other areas that have societal relevance based on their public attention such as finance and climate remain relatively underrepresented.

Diversification Needed
When the areas of application are compared to those suggested to be funded by the EU under the H2020 programme (see European Commission n.d.), some areas emerge as currently underrepresented such as Food and Energy, or are not being addressed at all such as Transport and Space. Since this study was undertaken, however, some of the less addressed areas already are being picked up. Food, for example, is the main area of application in a study undertaken by Wageningen University (Bruijnis et al. 2015).

The landscape study provides access to the discourse by presenting a comprehensive overview of the different experts and sources that are currently available in the discourse. Moreover, the analysis shows which actors act as hubs in the different (thematic/regional) segments that together make up the discourse. Their position and connections within the discourse make these hubs well equipped for unlocking segments of the discourse to (new) actors joining it as well as to others interested in

RRI. Furthermore, the study hints towards future directions of the discourse that are relevant to both policy-makers and RRI researchers. Well-represented areas and terms of addressing may be interpreted as important and therefore in need of further attention, while underrepresented areas represent opportunities for further research or justify further policy attention.

Acknowledgements The research for this chapter was first undertaken for a deliverable (D 6.4.b) of the EU funded FP7 Governance of REsponsible innovATion (GREAT) project.[4] The author would like to thank Bernd Stahl and Sara Wilford for their collaboration on the deliverable and fellow consortium members of the GREAT project for their support and feedback. The author also would like to thank the two anonymous reviewers for their valuable suggestions and comments to improve the quality of the chapter.

Appendix I: Coding Rules Used for Deductive Thematic Coding

Related questions	Code	Coding rules
(1) & (2)	Organisation	Institutes involved in projects as mentioned on project websites. Institutes that funded a project for example the EU or EPSRC. Institutes authors of book chapters, journal papers, reports or policy documents are affiliated with. Overall institute is coded, not the underlying centre or department. If in one source multiple institutes are associated with a person only the main affiliation of that person is coded.
(1) & (2)	Person	Authors of book chapters, books, journal papers, reports, or policy documents. Project members as mentioned on project websites. If project members are not mentioned on a project website, other available documents such as deliverables are included and investigated as a source in the data. When a person is deceased he or she is excluded from the codes, for example, Prof. Herbert Gottweis. Members of Advisory Boards of projects are not coded as these persons are not actively working on the project themselves.
(1) & (2)	Project	When the content of the source is about a project, for example the sources representing the websites of projects or journal papers that stem from a project.
(3)	Area of Application	Technology or field implicated in the paper, for example, ‘financial innovation’ or ‘geo engineering’.

[4] http://www.great-project.eu/

Related questions	Code	Coding rules
(4)	Term	When a certain requirement to the R&I process to render it 'responsible' is discussed in a source, for example, 'anticipation' and 'engagement' When certain issues that could or should be addressed when implementing RRI is discussed in a source, for example, 'gender' and 'privacy'. When a certain means to implement or support implementing of RRI is discussed in a source for example, 'education' and 'governance'.

References

Asante, K., and R. Owen. 2012. *Towards a framework for responsible financial innovation*. Available at: http://innovation-finance.altran.fr/files/LUX_2012_Towards_a_framework_for_responsible_financial_innovation.pdf. Accessed 6 Nov 2012.

Biernacki, P., and D. Waldorf. 1981. Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods and Research* 10 (2): 141–163.

Bruijnis, M.R.N., et al. 2015. Moral "Lock-In" in responsible innovation: The ethical and social aspects of killing day-old chicks and its alternatives. *Journal of Agricultural and Environmental Ethics* 28 (5): 939–960.

European Commission. n.d. *Horizon 2020 – The EU framework programme for research and innovation*. Available at: https://ec.europa.eu/programmes/horizon2020/. Accessed 7 Dec 2015.

———. 2013. Horizon 2020. Work programme 2014–2015. Science with and for society, Brussels. Available at: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-swfs_en.pdf.

Fisher, E., and A. Rip. 2013. Responsible Innovation: Multi-level dynamics and soft intervention practices. In *Responsible Innovation*, ed. R. Owen, M. Heintz, and J. Bessant. Chichester: Wiley.

Geoghegan-Quinn, M., 2012. *Responsible research and innovation.* Europe's ability to respond to societal challenges. European Union. Available at: http://ec.europa.eu/research/science-society.

Given, L.M. 2008. *The SAGE encyclopedia of qualitative research methods*. Los Angeles: SAGE.

Grunwald, A. 2011. Responsible innovation: bringing together technology assessment, applied ethics, and STS research. *Enterprise and Work Innovation Studies* 7: 9–31.

Macnaghten, P., and R. Owen. 2011. Environmental science: Good governance for geoengineering. *Nature* 479 (7373): 293–293.

NWO. 2016. NWO-MVI (Maatschappelijk verantwoord innoveren). NWO-MVI (Maatschappelijk verantwoord innoveren). Available at: www.nwo.nl/onderzoek-en-resultaten/programmas/maatschappelijk+verantwoord+innoveren. Accessed Feb 20 2016.

———. 2010. *Responsible innovation. Project summaries*, Netherlands Organisation for Scientific Research (NWO).

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., J. Stilgoe, P. Macnaghten, M. Gorman, E. Fisher, and D.H. Guston. 2013. A Framework for Responsible Innovation. *In Responsible Innovation*, ed. R. Owen, M. Heintz, and J. Bessant. Chichester: Wiley.

Oxford English Dictionary. n.d.. Available at: http://www.oed.com/. Accessed 23 Nov 2013.

Siune, K. et al. 2009. *Challenging futures of science in society. Emerging trends and cutting-edge issues*. Report of the MASIS Expert Group. European Commission. Available at: http://boa.unimib.it/handle/10281/9743. Accessed 8 Jan 2013.

Stahl, B.C., G. Eden, and M. Jirotka. 2013. Responsible Research and Innovation in Information and Communication Technology. Identifying and engaging with the ethical implications of ICTs. In *Responsible Innovation*, ed. R. Owen, M. Heintz, and J. Bessant. Chichester: Wiley.

Stilgoe, J., R. Owen, and P. Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580.

Sutcliffe, H. 2011. *A report on responsible research & innovation*. Available at: http://www.matterforall.org/pdf/RRI-Report2.pdf. Accessed 6 Nov 2012.

Von Schomberg, R. 2011. *Towards responsible research and innovation in the information and communication technologies and security technologies fields*. SSRN Electronic Journal.

Von Schomberg, R. 2012. Prospects for technology assessment in a framework of responsible research and Innovation. In *Technikfolgen abschätzen lehren*, ed. M. Dusseldorp and R. Beecroft, 39–61. VS Verlag für Sozialwissenschaften.

Zwart, H., L. Landeweerd, and A. Van Rooij. 2014. Adapt or perish? Assessing the recent shift in the European research funding arena from "ELSA"to "RRI". *Life Sciences, Society and Policy* 10 (1): 1–19.

Chapter 4
"Response-able Practices" or "New Bureaucracies of Virtue": The Challenges of Making RRI Work in Academic Environments

Ulrike Felt

Abstract In recent years, "Responsible Research and Innovation" (RRI) has become a new buzzword at the core of European science policy discourses and beyond. Using a narrative approach, this paper aims to explore how academic researchers can potentially make sense of RRI and turn it into an academic core value. Narratives on research and its relation to society drawn from different sources in the Austrian context will be used to reflect on how they contribute to creating shared meaning, participate in the constitution of a broader sense of direction and valuation, and enable or constrain researchers' actions. Using epistemic living spaces and narrative infrastructures as key-sensitizing concepts, the paper identifies and elaborates on three main narrative clusters that collectively frame the ways in which researchers can make sense of their work and engage with questions of RRI. In conclusion, this allows identifying the potential resistances RRI might encounter, the research still to be done in order to understand the dynamics at work and the work needed to support developing the concept's full potential.

4.1 Introduction

In recent years, "Responsible Research and Innovation" (RRI) has become a new buzzword at the core of European science policy discourses and beyond (de Saille 2015). The emergence of RRI can be seen as aligned with a wider trend according to which "the ideal of value-free curiosity-driven science"—although this concept is evidently a myth—"has been replaced by science responding to societal concerns" (Bos et al. 2014, 151). It is an expression of a situation in which a powerful role is attributed to innovation in terms of ensuring economic growth and well-being, with increasing demands placed on research to address specific societal needs.

U. Felt (✉)
Department of Science and Technology Studies, University of Vienna, Vienna, Austria
e-mail: ulrike.felt@univie.ac.at

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_4

Talk of a "new renaissance", in which Europe would build and be built on a new social contract fostering "shared responsibility between science, policy and society", would be another element supporting our observation of a shift (ERAB 2009). This way of positioning science and innovation in society promises to ensure both "socially beneficial action as well as freedom of thought" (ERAB 2009, 7), but it does not address how both could be achieved in one and the same move. Part and parcel of this rethinking exercise is a deeply rooted concern of policy-makers regarding a missing "innovation-friendly climate", expressed through the voicing of regular complaints about the absence of a continued and, if possible, unconditioned public support for research and innovation. This move in focus towards broader societal needs, values and challenges can also be traced—with different degrees of normativity—to their being embedded in new funding lines in a number of national contexts (e.g. Felt et al. 2016; Stilgoe and Guston 2017). We thus encounter novel ways of framing legitimate research, e.g., proposing or even prescribing tighter collaborations of scientists with researchers in social sciences and humanities or with societal actors engaged in the respective problem zone for which researchers attempt to develop solutions. Under the label of RRI (and similar denominations), we thus witness what Ribeiro et al. (2017, 81) describe as "concerted experimentation in many academic circles" with new forms of interventions into processes creating knowledge and innovations. Under the label of RRI we witness how resources and spaces could be mobilised for this new way of approaching research—despite varying to a large degree across different national and institutional contexts.

By definition, a buzzword is but "a word or phrase used by members of some in-group, having little or imprecise meaning but sounding impressive to outsiders".[1] Recently, Bensaude-Vincent (2014) convincingly argued how, for example, the term "public engagement with science" has become a buzzword and gained power exactly because it represents a positive value while remaining somewhat vaguely defined. In that way, it retains great interpretative flexibility while managing to successfully convey a shared and largely uncontested matter of concern. A buzzword is thus capable of assembling diverse sets of actors around it and of developing the power to set normative goals. We can thus interpret the implementation and spreading of RRI as an effort toward "steering with big words" (Bos et al. 2014).

"Responsible research and innovation" indeed fulfils all of the specifications of being a buzzword. It has been presented as the new governance model for research and innovation, "placed at the centre of the Europe 2020 strategy".[2] According to dominant EC discourse, RRI is—or promises to be—an all-embracing approach that simultaneously promotes excellent science, a competitive economy, and societal progress without compromising sustainability, ethical acceptability, and societal desirability in regard to research and innovation. RRI therefore stands for the

[1] http://www.collinsdictionary.com/dictionary/american/buzzword

[2] Regulation (EU) No. 1291/2013 of the European Parliament and the Council of 11 December 2013 Establishing Horizon 2020—the framework programme for research and innovation (2014–2020)—of 20 Dec. 2013, Official Journal of European Union, L347/104, http://ec.europa.eu/research/participants/data/ref/h2020/legal_basis/fp/h2020-eu-establact_en.pdf

idea that science and technological development should be conducted responsibly, which is, in principle, an uncontested and incontestable ideal. Indeed, discussions of responsibility of/in science have persisted for a number of decades; it is therefore both a long-standing concern and a concept that has always needed adaptation to new situations in which technoscientific and societal developments become entangled. Furthermore, RRI responds to a shared matter of concern: to find adequate new ways of creating a mutually supportive co-development of techno-science and society in democratic societies. RRI thus seems to function as a 'moral glue' that holds the abovementioned disjunct promises—economic, societal and scientific benefits—together, allowing the wider sociotechnical imaginary (Jasanoff and Kim 2015) of a European innovation society to fully unfold. It is this basic agreement of a wide set of different actors on the ideal captured by RRI that also explains the proliferation of narratives and initiatives around it.

The very idea of opening up research and innovation to a broader range of societal actors and values, also expressed through the idea of pluralising expertise, appears promising; however, if taken seriously, it would demand a radical rethinking of some of the very practices and values that are deeply entrenched in contemporary research cultures. This chapter aims to reflect on these issues by asking the question, "How can researchers potentially make sense of RRI and turn it into an academic core value in a research world that focused on rather narrowly defined ideals of excellence and relevance?" The question will be explored taking a narrative approach, in particular investigating dominant forms of narratives on academic research and its relation to society as these are key to RRI-related sense-making practices. Narratives are not only seen as a way of sharing meaning in practice, but also as participating in the constitution of a broader sense of direction, value and purpose of academic work, in the reconfiguring of individual and institutional identities, and in the enabling and constraining of researchers' actions. This, in turn, will allow identifying potential problems and hurdles to be overcome if we want to successfully make RRI a core value in research. Empirically I will draw on a broad set of materials—interviews with researchers and policy makers, policy documents, focus groups with researchers, participant observations mainly in the Austrian context—collected in the framework of five major research projects. While these projects did not explicitly addressed the question of RRI, they all investigated in great detail the changing lives in contemporary research in different fields, the shifting practices in attributing value to different parts of academic work as well as the new relations researchers are supposed to develop with societal actors.[3] Building

[3] My gratitude goes to the many researchers who took the time to participate in interviews and discussion groups in the following projects conducted between 2004 and 2014: "Let's talk about GOLD. Analysing the interactions between genome research(ers) and the public as a learning process", funded by GEN-AU as an ELSA project; "Knowing – Knowledge, Institutions and Gender. An East-West Comparative Study", funded by the European Commission, FP6. "Living Changes in the Life Sciences. Tracing the Ethical and Social within Scientific Practice and Work Culture", funded by GEN-AU as an ELSA project. "Making Futures Present. The Coproduction of Nano and Society in the Austrian Context, funded by FWF". "Transdisciplinarity as culture and practice", funded by BMWFW under the programme provision.

on experiences from these projects will allow pointing to both potential problem zones which could hinder making RRI a core value in academic research as well as areas where more research on RRI in practice would be needed.

In what follows, I will set the stage by briefly reflecting on the current efforts to define RRI and the potential it might hold. I will then argue my approach to the question of RRI in academic practice and present the key-sensitising concepts—*epistemic living spaces* and *narrative infrastructures*—that guide my analysis. Building on this, I will elaborate on the different clusters of key narratives that frame the ways in which researchers (can) make sense of their work and engage with questions of RRI. The potential effects of the co-presence of these narratives will then be a further perspective addressed. This will, in conclusion, allow identifying the potential resistances RRI might encounter and explicate the work that needs to be done to support developing the full potential the concept of RRI might possess for rearticulating the relation of science and society.

4.2 Hope and Promise: Efforts to Define RRI

RRI as a new concept must be understood as a further step in a long history of policy debates on the ways in which science, technology and society should engage with each other (Stilgoe and Guston 2017). Questions of how we should govern novel research and technological developments, how participation can help shape research agendas in a way that responds better to societal needs and concerns, and how broader societal values could be integrated into research and innovation processes have indeed long been on the agenda. Debates on ethical, legal and social aspects of research (ELSA) (e.g. Zwart et al. 2014), various forms of technology assessment (e.g. Guston and Sarewitz 2002), mid-stream modulation (e.g. Fisher et al. 2006), enhanced ethical integration in laboratories (e.g. van der Burg and Swierstra 2013) and the need to open up policy processes to other types of expertise (e.g. Stirling 2008) are but a few examples of the long ongoing debates. While these activities were all meant to realize the spirit of engaging research and society, we simultaneously witness a constant concern that "an effective program for influencing […] policy could easily be sacrificed in favor of a sham program that merely gives the impression of doing so" (Fisher 2005, 322).

Thus, well before RRI gained prominence as a concept through the most recent EU research framework program, Horizon 2020, we can identify numerous efforts on the European level to open up research to societal actors and values. RRI is therefore considered a further effort to rethink the complex relationship between techno-scientific and societal developments carrying important promises and hopes: it is tied to the strong idea that through a successful integration of research and innovations with societal needs and values, we can ultimately support the creation of more adequate responses to the "grand challenges" of our time and thus demonstrate the utility of research and innovation for societal development.

In the early phase of discussions around RRI, we witness diverse efforts to find a basic definition of the new label and to clarify the foundational principles. The aim was to leave the notion sufficiently vague to allow broad adherence while making it sufficiently concrete to turn it into a reasonably well-functioning device for policy-making and research practice. RRI should allow both avoiding unintended consequences before the development and implementation of innovations as well as moving issues of governance away from "reactive forms […] to proactive forms" (Ribeiro et al. 2017, 89). In an early move, we find rather broad definitions of RRI as being "a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to 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)" (von Schomberg 2011, 9). Key notions here are acceptability, desirability and proper societal embedding, which should all be cared for in the *process* of innovation and not once the ready-made innovation enters society. In a later writing, von Schomberg (2013, 54) specifies that this means that we should not leave choices solely to the market mechanism, as is often proposed, but should instead deliberate much more carefully on "the normative dimension of what counts as an 'improvement'". This is a highly relevant specification as it calls for shifting our attention from the value *of* innovations (mostly meant in terms of market value) to the values *in* innovations. The latter means posing the question of whose values and concerns are considered when developing knowledge and innovations and calling for a more careful deliberation of the direction, scale and speed of innovations (Felt et al. 2007).

This is explicitly addressed by Owen et al. (2013, 29), who clearly argue that RRI should support explicitly "looking to those prospective, forward-looking dimensions of responsibility, (notably *care* and *responsiveness*) which allow consideration of purposes and accommodate uncertainty, a defining feature of innovation". This is in line with earlier voices stressing how much the *direction* of innovation trajectories matter and asking "whose visions of the future should drive the direction of progress" (Felt et al. 2007, 35). Posing similar argumentations, Stilgoe et al. (2013, 1570) identify four dimensions that should characterise RRI-related engagements, i.e., "anticipation, reflexivity, inclusion and responsiveness," which should be fostered in research and innovation systems. Together, these dimensions should successfully "provide a framework for raising, discussing and responding" to the key questions with which any sociotechnical development of contemporary societies is confronted.

In regard to realising RRI in research practice, we can frequently identify a combination of the following demands: research should (1) engage with a diverse set of societal actors and build new relations with society in order to address uncertainties and potentially unintended consequences; (2) create spaces for addressing divergent values relevant to both research and society and thus implement from within the research and innovation systems processes of collective care; and (3) foster inter- and transdisciplinary collaboration to be able to deal with the complexities of societal challenges. The third point was explicated in a report of an expert group

exploring the options for strengthening RRI (EC 2013), stressing that "more inter- and trans-disciplinary research should nurture greater innovation and creativity, and make it more likely that research and innovation are directly targeted at solving societal challenges".

It is safe to say that the emerging research field around the notion of RRI has so far strongly focused on issues of the responsible governance of technological development and design.[4] Less attention has been given to mapping the many meanings of responsibility in research (Glerup and Horst 2014) and to better understanding the "subterranean dimensions of RRI" (Ribeiro et al. 2017, 82). A recent ESF policy briefing indicates that academic practices, routines and (e)valuation schemes in place could potentially limit the full implementation of RRI (Felt et al. 2013). Therefore, this paper aims to offer a reflective inroad to address these less present concerns and to reflect on the boundary conditions needed in order to make space for RRI in academic research practice.

4.3 RRI in Academic Practice: Conceptual Reflections

While it is highly relevant to develop a better conceptual grasp of what RRI should mean and to develop programmes to foster it, it seems essential first to better understand how individual researchers or research groups can navigate and cope with the complex realities of contemporary research environments and the new demands that are expressed through RRI. Doing so requires reflecting on the "responsibility conditions" researchers find and how they open up or close down engagement with RRI related issues. A more person-centred approach seems suited to address this perspective. In a study of contemporary knowledge cultures in academia, the concept of "epistemic living spaces" has been developed, drawing attention to researchers' individual or collective perceptions of "the multi-dimensional structures—symbolic, social, intellectual, temporal and material—which mould, guide and delimit in more or less subtle ways researchers' (inter)actions, what they aim to know, the degrees of agency they have and how they can produce knowledge" (Felt 2009, 19). This concept directs our focus to the room for manoeuvring researchers perceive that they have in performing research, following their ideas and reflecting on them, arranging the private and the professional realms and engaging with societal issues—all aspects that researchers often implicitly equate with their quality of life in research. At the same time, the concept also underlines the fact that researchers are not passive subjects who inevitably submit to structural change and shifts in values; rather, they aim to express their agency through the diverse kinds of work—e.g., actions they take, resistances they express, alternative stories they tell— they invest in shaping their epistemic living spaces from within in ways to make them worth inhabiting. The exact form of an epistemic living space surely differs

[4] For a more complete overview on the dimensions and definitions of RRI see Burget et al. (2017) and Ribeiro et al. (2017).

according to the moment in a researcher's career, the investment a researcher can make in resisting certain pressures or selectively embracing some changes, the epistemic subfield in which s/he is working, the institutional infrastructure and its elasticity, institutional and leadership cultures and more formal and informal networks the researcher can entertain in the work environment and beyond.

The concept of epistemic living spaces also raises our attention regarding the ways in which narratives matter as sense-making devices (Czarniawska 1998) — making sense of narratives circulating and developing narratives to make sense of one's own life in research. It sensitises us to how societal imaginaries as well as shifting institutional and policy framings are experienced by individual researchers and how these factors potentially affect their practices. Narratives have been found to be particularly important resources for sense-making when having to address uncertain, novel or unusual developments. They become the social coordinates which allow orientation in shifting environments. Looking at how researchers' room for manoeuvring is shaped in and through narratives and how researchers can thus potentially accommodate RRI perspectives in their work is thus at the core of this chapter. Narratives are therefor never mere stories, but important resources for analysts to understand change beyond formal structural shifts.

In investigating researchers' epistemic living spaces, I draw particular attention to different types of narratives, how they are produced, structured, circulated and taken up, what resources are used to build them and what meanings of research and responsibility and researchers' role in all this are articulated. This draws the analyst's attention to the fact that space (epistemic living space, in our case) never simply is given and forms the setting for the stories researchers tell but, more importantly, that "the production of space" can be observed "through the act of narration" (Donald 1997:183). In that sense, I understand narratives as windows to the social, epistemic, political, symbolic and many other facets of life in research, and it appears essential to reflect on the different plots that are developed to bring elements in meaningful wholes and make sense (Czarniawska 2004). Thus, it is essential to observe how diverse actors make use of characteristic, shared accounts that express wider imaginations about research and its relation to society, which values matter, how research and innovation (should) work, and the place and agency actors have in both. Following the acts of narrating research is thus, arguably, the closest that we can get to researchers' experiences under contemporary conditions. The many conversations (formal, such as interviews, or informal, such as over coffee) with researchers on which this paper is built are such moments where storytelling happens. So are policy reports and other moments when constructing coherent storylines it at stake. Narratives thus are seen as "reflect[ing] prevailing institutional structures, express[ing] values and reinforce[ing] collective aspirations" (Felt et al. 2007, 73). In the world of research and innovation, narratives "tacitly define the horizons of possible and acceptable action, project and impose classifications", define values and norms that guide us, "distinguish between relevant issues from non-issues, and central actors from non-actors" (Felt et al. 2007, 73).

These narratives constructed by academic researchers and policy makers, however, should not be considered a simple collection of different stories; rather, they

should be viewed as contributing to and simultaneously being nourished and stabilised by a wider *narrative infrastructure* around issues of academic research and innovation and its relation to society.[5] A narrative infrastructure is—in the context of this chapter—a network of temporally stabilised narratives through which meanings and values of academic knowledge/work and its relation to society can be articulated, circulated and exchanged across space and time. Such infrastructures also contain diverse sets of actors who create, adapt, multiply, support and entangle such dominant narratives. As argued below, these narratives can take different forms, including assessments, reconfigurations of past developments, future-oriented accounts voicing promises and improvements but also potential threats, and moral reflections of what is good science and innovation and how a good researcher should be. They also encode the hopes and expectations of individuals and institutions and thus become "the vehicles whereby [these hopes and expectations] are transmitted and made emotionally real" (Larkin 2013, 333). Finally, they can be more or less formal accounts—some being strategic and others being more procedural—all of them addressing specific audiences, such as members of a scientific community, diverse policy makers or different publics.

Such a narrative infrastructure not only enables and constrains possible (alternative) narratives but simultaneously participates in the definition of relevant actors and potential relations between them as well as of forms of agency. In that sense, it reflects and performs the material and social settings in which responsible research and innovation can occur. However, narrative infrastructures are always evolving, reconfiguring in ever-new constellations. They perform specific temporalities, identify preferred directions of development, and point at specific vulnerabilities. Narrative infrastructures thus undergird any major subsystem of modern societies—in the case of this chapter, the research and innovation system that should embrace RRI—to form the ambient discursive environment (Larkin 2013; Felt 2016).

Narrative infrastructures are therefore not a mere accumulation of stories and of human actors developing, exchanging and integrating them. They manage to create coherence, although many different actors and institutions are involved, and no single lead agent can be identified. In that sense, we should understand the narrative infrastructures in which RRI stories are embedded and simultaneously change as a form "tacit governance" (Felt and Fochler 2010) of research and innovation. This means that I see narrative infrastructures as part of a strongly mediated and less visible form of steering. Looking at the work narratives of RRI do and how they relate to other dominant narratives on research and innovation will offer us insights into the ways such interventions in the research system will intersect with "the respective practices, traditions, ideals and experiences that already exist in the respective research fields and institutions" and will (have to) find arrangements. This will make us aware of the potential fault lines and ruptures the introduction of RRI might create, leading to a rejection of or distancing from reflecting on responsibility, but also of the new spaces of possibility and negotiation where the changing relations of

[5] This notion was first used by Deuten and Rip (2000) to study design processes in an organisation.

research and society can find consideration and engagement in academic practice (Felt and Fochler 2010).

4.4 Narrative Infrastructures and How They Participate in Shaping Epistemic Living Spaces

What are the dominant clusters of narratives that collectively form the contemporary narrative infrastructures in the area of academic research and innovation? How do they participate in the shaping of researchers' epistemic living spaces and, through this, create conditions that are favourable or not to actively engage with issues of responsibility in research and innovation? In what follows, I argue that when contemporary researchers want to express their understanding of their epistemic living space and the space of RRI in it, they must relate to broader societal narratives on research and innovation, what is expected from science and what society values about research and innovation. This is visible through accounts in diverse media but also in the public and private environments in which researchers interact with members of society. The media is a dominant player here in defining what is a narrative of public interest and in making assessments about what is considered interesting and relevant research. Simultaneously, we can also trace wider policy narratives (often also repeated by the media), such as that on "the public" potentially not being sufficiently supportive of science or even technophobe, the complaint that universities often act as though they can remain in their famous ivory tower (i.e., disconnected from society) or the need to compete and produce innovations faster than ever before to avoid losing the place in the global economic race.

Although these are important elements in the narrative infrastructure of contemporary research and innovation, in what follows, I focus on those parts of the narrative infrastructure that are most prominently and explicitly visible within academic research institutions. I identify three intersecting narrative clusters (see Fig. 4.1) traceable in researchers' interviews as well as in policy discourses on different levels within and outside academic research institutions. They concretely address academic research, its organisation and the expectations of researchers, thus participating in shaping epistemic living spaces. It is this interaction between these different narratives, the arrangements they can form or the paradoxes and contradictions they create that matter. Together, they (often implicitly) create room for considering responsibility in research and of researchers in contemporary research—in short, they play an essential part in forming what I call the responsibility conditions of contemporary academic research.

The first cluster of narratives on contemporary research revolves around the substantial reorganisations in the research system—often subsumed under the label of *new public management (NPM).* This reorganisation of research is expected to increase output and efficiency and support the expectation of ever-faster innovation. It comes along with significant shifts in the "orders of worth" (Stark 2009) that

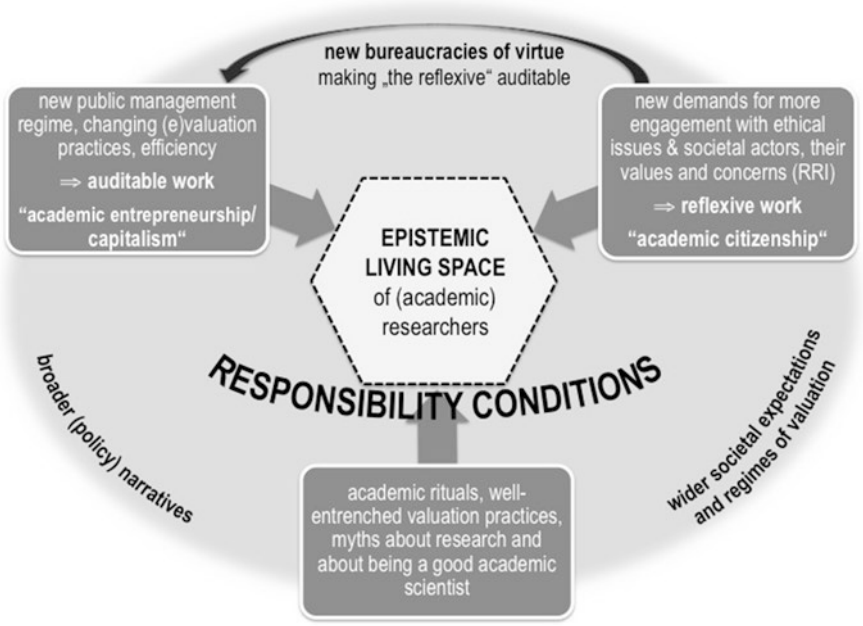

Fig. 4.1 Narrative infrastructure participating in the shaping of epistemic living spaces of researchers

guide academic work and the choices researchers make within it (e.g. Fochler et al. 2016). Auditing and ranking structures have been put in place to measure the scientific output and other features that are expected to allow a (constant) comparative assessment of people and institutions (Strathern 2000; Shore 2008). Thus, within institutions, increasingly, only those academic activities count that can be counted, whereas others are attributed less importance, need more work to make them visible or are even neglected. These seemingly objective forms of measurement have been shown to have significant governance effects on the practices of institutions and individuals (Espeland and Sauder 2007), particularly when they come hand in hand with a growing discourse on efficiency and productivity. Indeed, well-entrenched audit cultures are powerful in defining what constitutes legitimate and worthwhile inquiries (Shore and Wright 2015), and once researchers internalise auditing criteria and adapt their behaviour to "game" specific aspects of the system, they start to engage in self-auditing processes along these criteria. Evaluations, the criteria they make explicit and the narratives that gravitate around them, can therefore be seen as a form of assisted sense-making (Dahler-Larsen 2011). This does not mean that researchers cannot resist such imposed orders of worth, but it needs work, considerable emotional investment and certain forms of risk taking—all elements not equally affordable to everyone in the system.

This is part and parcel of a deep restructuring of the temporal logics guiding academic lives and epistemic work. An explicit sign of this change is a growing

"projectification" which has grasped many areas of our lives. For academic research this means that it is predominantly conducted through third-party-funded projects. Shorter work contracts in the early phases of academic life become typical as research time is becoming something to be bought and sold (Ylijoki 2015). This also participates in the restructuring of academic timescapes, thus changing the rhythms of work and multiplying the demands on clarifying the relations between diverse outputs and time spent on them (Garforth and Cervinková 2009; Felt 2016; Gibbs et al. 2015). Within a projectified academic world, we observe the spreading of the language of work packages, deliverables, roadmaps and person months. This allows the turning of time into an essential commodity in the research system, and the question of who owns whose time and can spend it on specific tasks is moving onto the agenda. Young researchers in particular are thus requested, more than ever before, to invest considerable energy to build a coherent academic life out of the temporally fragmented elements, both epistemic and social.

With denser accountability regimes, the idea of control and planning enters the scene. The future and how to ensure it come to play an ever-larger role, and working towards a specific future is to be staged as a key element in academic work. Here, we clearly see signs of what Appadurai (2013, 223–224) has aptly labelled trajectorism, i.e., "a deeper epistemological and ontological habit, which always assumes that there is a cumulative journey from here to there, more exactly from now to then, in human affairs […]. Trajectorism is the idea that time's arrow inevitably has a telos, and in that telos are to be found all the significant patterns of change, process and history". Planning and working towards a future is then also tied to the ideal that we need to focus on a specific direction and that researchers—earlier in the knowledge generation cycle than ever before—are asked to reflect on the relevance of their research and on potential applications.

NPM logics, the re-timing of research and a growing obsession with controlling the future thus collectively lead to a call for an intensification of auditable work and for researchers to become more of academic entrepreneurs (Shapin 2008) who are capable of navigating the complex funding and career landscape and behaving in accordance with the logic of academic capitalism (Fochler 2016).

The second cluster of narratives is related to a growing demand for researchers to perform what I summarise under the label of "reflexive work" (see the precursors of RRI mentioned above). Over the past two decades, researchers have, more than ever before, been expected to engage in reflection on ethical, social and legal issues related to their research (see the ELSA/ELSI programmes put in place in Europe and the US to accompany research, particularly in the life sciences but also in other domains; see Hilgartner et al. (2017)), which should go well beyond questions of obtaining ethical clearance for research. Increasingly, the topics of fraud and transgressions of good scientific practice have moved up on the agenda, leading to a stronger regulatory network and to the expectation that researchers more explicitly address the different issues related to what it means to conduct good research. Furthermore, researchers are expected to develop their capacities to reflect on the potential future impact of their work, i.e., to anticipate potential problems and address them proactively. Finally, integrating, engaging or at least communicating

with diverse societal actors relevant to problems to be solved has been a gradually growing duty of researchers (see for example Felt et al. 2013). In short, these demands that have appeared over the past two decades and have been translated into diverse programmes on the European level and in national settings cover much of what is now subsumed under the label of RRI. This shift is closely related to universities' and policy makers' increasing concerns about the societal support for research and thus has gained considerably importance over the past two decades.

Thus, the figure of researchers is being reimagined; they are increasingly conceptualised in a framework, Macfarlane (2007) aptly labelled "academic citizenship." This means that researchers are no longer solely expected to focus only on doing research and teaching but should also care for the infrastructure that supports academic life, fulfil their civic mission, care about service to communities, engage with citizens in diverse formats and much more. In some higher education institutions, this has been introduced under the label of academic citizenship or as the third mission and has, in some places, been used as additional criteria in hiring. Yet, as Macfarlane also points out, status is related to "the extent to which the activity is regarded as 'scholarly', whether the activity is internal or external to the university and the degree to which the activity is 'visible' to colleagues and rewarded in performance-related terms" (Macfarlane 2007, 264). Service to society is in that sense often seen as external to the core activities and therefore less valued, and, as studies have shown, is unequally distributed, with female researchers often taking a larger share of this type of care and articulation work (Kerr and Lorenz-Meyer 2009). The degrees to which researchers are committed to activities which would fall under the label of RRI, therefore will reflect "the micro politics of life within modern universities" (Macfarlane 2007, 267).

Although these two narrative clusters, which call for auditable and reflexive work respectively, obviously stand in tension and place different types of demands on researchers, they both must be seen on the backdrop of a third cluster of narratives that strongly refer to past conditions of research. These narratives contain reflections on past academic rituals, the values and valuation practices that have supposedly guided research successfully over long periods, myths about how good research should work as a knowledge-producing enterprise, what conditions creativity needs and what makes a good researcher. Researchers narrate a past where they were freer to choose the topics they found interesting and follow their intuition, where less time was spent on counting formal outputs and on "selling" their findings, when there was less time pressure and when academic careers were still attractive. What we can observe through the re-performance of these narratives is the continuous work that goes into what Hobsbawn and Ranger (1983, 1) call the "invention of tradition", an effort to create "continuity with a suitable historic past". This produces some feeling of stability in an environment that is geared towards innovation and fast change. Indeed, the sharing of academic myths never serves to solely "refer to an event alleged to have taken place long ago" but, above all, allows one to establish a specific pattern or habit as timeless, thus holding validity for the past, present and future (Lévi-Strauss 1963, 209). Supported by specific memory practices, these narratives on a "golden past" play an important role in positioning

the work of researchers. They become explicitly visible in backward-looking reflections, which Ylijoki (2005) aptly labels "academic nostalgia". She argues that this nostalgia should be understood as a sign of "current tensions and dilemmas in work" and researchers' efforts to find a balance between emerging entrepreneurial values and the norms and morals associated with traditional academic self-understanding.

4.5 Articulating the Three Narrative Clusters

As the quality of researchers' epistemic living spaces is tightly related to how researchers see themselves capable of carving out a space that is institutionally acceptable and sufficiently meaningful for them to develop their projects but also realise their aspirations, we need to better understand how they position themselves in the narrative infrastructures just described. What does the coexistence of these different narrative clusters mean for a successful implementation of RRI as the core value in academic environments? How do they articulate and thus form part of what I call the *responsibility conditions*? What agency do researchers have, and how much work do they have to do in developing their own sense-making narratives at the intersection of these three narrative clusters?

Approaching these questions, let us first look at the relation of the third cluster of narratives to the two others. To start with, it is important to realise that the narrative cluster revolving around ideals of new public management and the one related to increased reflexivity, in one way or another, distance themselves from and usually implicitly and explicitly critique past academic ways of functioning. They construct an academic past as insufficiently responsive to societal demands and concerns and not fully subscribing to the ideal of accountability, excellence, efficiency and attentiveness to societal values and concerns. Thus, academic pasts—so the narratives go—seemingly fostered neither academic entrepreneurship nor academic citizenship sufficiently. This construction of a deficient past enables individuals to perform a present that would be better adapted to realise the future imaginaries of progress through innovation.

However, also the reverse valuation can be traced in the narratives of those researchers who see currently changing research environments in contradiction with the classical academic norms and values. Embracing past academic traditions as a leading value often means voicing objection and resistance to what is labelled as the neoliberal reconfiguration of research under the name of new public management. This means critiquing the narrow vision of "counting only what can be counted" and seeing research at risk of narrowing too much towards strategic goals. However, these golden past narratives interestingly also—although to a lesser degree—object to an all-too-enthusiastic opening towards societal values and concerns. They would much rather like to see science advancing along its own logic, referring to past successes that were ensured by keeping science at a reasonable distance from society. In the long term, the argument would go, science has always proven to quasi-automatically create benefits for society that are not necessarily easily accountable

in the short term. Therefore, the cluster of traditional narratives would neither clearly support those that try to strongly foster reflexive work.

Finally, it is essential to reflect on the relation between the narrative clusters that focus on auditable work and reflexive work. Although there is broad agreement that researchers' adoption of the role of academic citizens is key for the place of the university in contemporary societies, numerous analysts have noted that "the drive to make higher education more 'efficient' through a more 'performative' reward structure and the casualisation of academic labour is undermining academic citizenship" (Macfarlane 2007, 271). This also means, so his argument continues, that the current academic system "rewards an individual rather than collective ethic," which in turn might hinder the putting in place of RRI as a shared value system. What happens when reflexive work, and thus RRI, is to be performed under the auspices of strong new public management ideals, is therefore a key question.

Two types of potential risks concerning the implementation of RRI can be identified from my observations in the field. First, reflexive work runs the risk of becoming "*form*-alised", i.e., translated into the activity of the filling out of specific *forms*, as in the case of ethical review processes. Here, we note the attraction of what Becker has called "the discrete charm of the form" (Becker 2007). A form seems attractive because it creates "tangible evidence", in our case, proving that broader societal values have been considered. At the same time, it is a tool of standardisation of what counts as the formal fulfilment of reflexive work and allows control. Informed consent forms, as introduced in the medical domain, are an excellent example, often triggering the critique that they serve the medical establishment more than they actually create informed choices for citizen-patients (Felt et al. 2009). A similar argument can be made about ethics forms in regard to project applications, which foster tick-boxing rather than wider reflection throughout the research processes. Second, given that we live in a strictly time- and efficiency-oriented research system, there is the risk that reflexive work might become reduced to a ritual act often performed at the end of a project and completed "by the book" i.e., not engaging in context/problem-specific models. Reflexive work might thus fall into what has been called in a recent report the "ritualization trap," i.e., forgetting why engagement with society has been called for in the first place and that it should continually adapt to ever-new situations. Or it might be completely outsourced by "delegating it solely to the social sciences and humanities" (Felt et al. 2013). While interdisciplinary collaboration between natural and social sciences proves essential in RRI, the critique here is that a separation of research and reflection through complete outsourcing carries the danger of not allowing discussions on societal values to become embedded in research and be admitted to the core of academic practice.

4.6 Concluding Thoughts

The chapter started by asking what might be the potential problems when trying to make RRI an academic core value. I took an approach centred around the researchers as actors in the field, pointing at how their "epistemic living spaces" are shaped by different forces captured in and expressed through three major narrative clusters. The focus on narratives was chosen in order to better see the possibilities and limitations in how researchers can attach meaning to and therefore make sense of RRI in their academic work environments. I highlighted that these narrative clusters stand in tension with each other while shaping researchers' room for manoeuvring in important way. However, it is also important to see that this does not necessarily mean that they do so in a deterministic manner. For researchers this means that when trying to develop narratives about their lives and their futures in research in a way that is satisfactory to them and allows them to unfold creatively, they have to create some personal arrangement between partly conflicting narrative clusters.

In conclusion, I, therefore, want to highlight four points flowing from the previous observations, which might need both closer consideration and more research in order to better understand the micro-dynamics at work.

First, I want to argue that introducing RRI as a buzzword, establishing a new cluster of narratives on responsible research and innovation and funding some related projects/programs will most probably not be enough to reach the claimed shifts in research culture and practice. To this end, we first have to gain a detailed understanding of how researchers can make sense of RRI in an environment in which their epistemic living space is strongly shaped by new public management ideals with a pinch of nostalgia for a time when these issues did not need to be considered. It is the responsibility conditions researchers navigate that require careful consideration when we expect RRI to be fully integrated into research practices. This also calls for considering the many meanings of responsibility in RRI and making the debate not only about the potential problems RRI might cause but about how to think about what new benefits this engagement could create for both science and society. In particular, the way in which research is valued, organised and timed needs close scrutiny because these elements shape the distribution of reflexive work among different researchers and within research processes. If we simply allow the NPM and RRI narratives to exist next to each other without a clear idea of how time and space can be made for reflexive work, we run the risk that introducing the RRI principles might not lead to the expected opening up of science towards societal challenges, concerns and values. Instead, we might witness the pragmatic creation of new forms to be filled out to prove that some form of RRI has been performed, the process of reflection might turn into an annex ritual to be performed in the beginning or the end of a project, or we could witness the complete outsourcing of reflexivity and anticipation to other researchers. This would ultimately lead to nothing more than the creation of "new bureaucracies of virtue."[6] It would thus support

[6] This notion was inspired by Jacob and Riles (2007), who coined it to study informed consent as

the research systems' 'account-ability' logic also in the field of RRI but without fostering researchers' willingness and capacity to explore values-sensitive responses to the complex questions that arise at the interfaces of science and society, i.e., making researchers 'response-able'.

Second, to make RRI work in a sustainable manner, it is essential to create institutional environments and foster narratives that allow *caring for society* to be fully integrated into the very core activities of academic research, well beyond any single project. RRI should thus be understood and implemented as a "technology of humility" (Jasanoff 2003), i.e., become an approach that acknowledges the complexities and uncertainties linked to research and innovation in contemporary societies and that "make[s] apparent the possibility of unforeseen consequences; make[s] explicit the normative that lurks within the [scientific and the] technical; and acknowledge[s] from the start the need for plural viewpoints and collective learning" (Jasanoff 2003, 240). Therefore, RRI would need to become integral part of the process of knowledge creation. This means developing, nurturing and valuing researchers' capacities of anticipation, reflexivity, inclusion and responsiveness. In the end, researchers should learn to not only ask questions relevant to their field but also develop sensitivities for detecting how societal problems to be solved are framed, how innovations might affect people in very different ways, how the benefits from knowledge and innovation are often unevenly distributed and where the limit is of what they know beyond the lab. This means that although it is important to integrate social scientists and other humanities scholars into the core of research and innovation processes, RRI should not solely be delegated to them and thus be kept out of the core research and innovation business. In short, a successful establishment of RRI also means that new knowledge relations have to emerge "between people engaged in different knowledge-generation practices and thus ways of seeing and explaining the world" (Felt 2014). This calls for opening up to cooperation not only across disciplinary boundaries but also between scientific and diverse societal actors and making space and time for this type of interdisciplinary work in the highly competitive climate that governs research as well as education. Education is essential here, as it not only trains the next generation of academic researchers but above all cultivates highly specialised citizens who will work at the interfaces of science and society and will, in a fast changing technoscientific world, increasingly be required to have these reflexive capacities for which RRI stands.

Third, the situatedness of researchers needs to be considered when they are supposed to engage in RRI work. Actually, when we hear the call for engagement, we might point with Suchman (2013, 157) at the ways in which "commerce and politics get both entangled and obscured in contemporary calls for 'user' relevance in all things." Indeed, when thinking of considering societal values in research we might encounter two very different meanings: "calls for value in the sense of utility, and a recognition of values as inextricable from the conduct of research." Therefore when

---

one expression of such new bureaucracies of virtue. In this paper, the notion is developed in a slightly different direction and is not specifically tied to the core questions of ethics in biomedicine.

implementing RRI, it is essential to differentiate "between normative research enlisted in the service of agendas—public or private—in which the frame is not itself open to question, and research that affiliates with efforts to question the frames within which politics, markets or any other entities are disciplined." It is the latter which was at the core of many actors who wanted to see RRI implemented in academic practice. In that sense, and the fear remains as for earlier programmes (e.g., ELSI/ELSA programmes mentioned earlier), that serious RRI engagement could also be sacrificed in favour of a more pragmatic and less time-consuming window-dressing program that merely gives the impression of doing RRI. It could follow a pragmatic framing aiming at rendering research more robust towards societal questioning but not to open to develop new framings of innovation and new innovation pathways. This perspective needs careful reflection in the funding processes of RRI-related activities, which should not normatively define what type of activities to foster and should not try to develop ready-made toolboxes that are then distributed in form of one-size-fits-all formats; rather, they should invite more open context- and institution-specific forms of engagement.

Finally, to grasp the complexity of the process of integrating RRI into contemporary research and innovation, it might be helpful to think of RRI in terms of a "techno-moral regime"[7] that should be put to work within the wider research and innovation system. Choosing the metaphor of "regime" is helpful for several reasons. It makes us alert to the institutions of research and innovation and the people who govern them, to their ideologies and the myths they cherish (see the abovementioned narrative infrastructures), and to the outputs they aim to achieve but also to the ways in which power is exercised (e.g., through evaluation practices, preferred narratives, career scripts, and many more). The notion of the regime also draws our attention to the many explicit and implicit prescriptions, ranging from science policies regulating the types of research that can be performed, over definitions of good academic practice to how we think of and try to foster societal progress as tightly coupled to science and technological development. Simultaneously, any specific regime is never uncontested. As I argue, any techno-moral regime building on RRI related values is contested by ideals of efficiency, accountability rules, tight time frames and much more; in short, the techno-morality of new public management is a powerful and often implicit opponent to the one performed through RRI related values. The techno-moral regime of RRI must therefore grapple with opposition and varying forms of dissent or resistance from both outside and within the institutions of research and innovation. Analysing RRI as a regime invites us to simultaneously look at the different levels, spaces and moments in which responsible research is negotiated and makes us aware of the strong commitment by many of the involved actors—researchers, societal actors, institutional leadership, policy makers and funding bodies—that is needed to allow RRI to remain open-ended and process-oriented in an academic world that prefers clear directions and promised deliverables.

[7] This notion has been inspired by Hecht's (2001) use of technopolitical regime.

References

Appadurai, Arjun. 2013. *The future as cultural fact*. London: Verso.

Becker, Peter. 2007. Le charme discrete du formulaire. In *Politiques et usages de la langue en europe*, ed. Michael Werner, 217–241. Paris: Editions de la maison des sciences de l'homme.

Bensaude Vincent, Bernadette. 2014. The politics of buzzwords at the interface of technoscience, market and society: The case of 'public engagement in science'. *Public Understanding of Science* 23: 238–253. doi:10.1177/0963662513515371.

Bos, Colette, Bart Walhout, Alexander Peine, and Harro van Lente. 2014. Steering with big words: articulating ideographs in research programs. *Journal of Responsible Innovation* 1: 151–170. doi:10.1080/23299460.2014.922732.

Burget, Mirjam, Emanuele Bardone, and Margus Pedaste. 2017. Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and Engineering Ethics* 23: 1–19. doi:10.1007/s11948-016-9782-1.

Czarniawska, Barbara. 1998. *A narrative approach to organization studies*. Thousand Oaks: SAGE.

———. 2004. *Narratives in social science research*. London: Sage Publications.

Dahler-Larsen, Peter. 2011. *The evaluation society*. Stanford: Stanford University Press.

de Saille, Stevienna. 2015. Innovating innovation policy: The emergence of 'Responsible Research and Innovation'. *Journal of Responsible Innovation* 2: 152–168. doi:10.1080/23299460.2015.1045280.

Deuten, Jasper J., and Arie Rip. 2000. Narrative infrastructure in product creation processes. *Organization* 7: 69–93.

Donald, James. 1997. "This, here, now: Imagining the modern city." In Imagining Cities: Scripts, Signs, Memory, (ed.) Sallie Westwood and John Williams, 181–201. London: Routledge.

EC. 2013. *Options for strengthening responsible research and innovation*. Luxemburg: Publications Office of the European Union.

ERAB. 2009. *Preparing europe for a new renaissance. A strategic view of the european research area*. Luxembourg: Office for Official Publications of the European Communities.

Espeland, Wendy Nelson, and Michael Sauder. 2007. Rankings and reactivity: How public measures recreate social worlds. *American Journal of Sociology* 113: 1–40. doi:10.1086/517897.

Felt, Ulrike, ed. 2009. *Knowing and living in academic research. Convergence and heterogeneity in research cultures in the european context*. Prague: Academy of Sciences of the Czech Republic.

———. 2014. Within, across and beyond: Reconsidering the role of social sciences and humanities in europe. *Science as Culture* 23: 384–396. doi:10.1080/09505431.2014.926146.

———. 2016. Of time-scapes and knowledge-scapes: Re-timing research and higher education. In *New landscapes and languages in higher education*, ed. Peter Scott, Jim Gallacher, and Gareth Parry, 129–148. Oxford: Oxford University Press.

Felt, Ulrike, and Maximilian Fochler. 2010. Riskante Verwicklungen des Epistemischen, Strukturellen und Biographischen: Governance-Strukturen und deren mikropolitische Implikationen für das akademische Leben. In *Steuerung von Wissenschaft? Die Governance des österreichischen Innovationssystems. Innovationsmuster in der österreichischen Wirtschaftsgeschichte, Band 7*, ed. Peter Biegelbauer, 297–328. Innsbruck: StudienVerlag.

Felt, Ulrike, Brian Wynne, Michel Callon, Maria Eduarda Gonçalves, Sheila Jasanoff, Maria Jepsen, Pierre-Benoît Joly, Zdenek Konopasek, Stefan May, Claudia Neubauer, Arie Rip, Karen Siune, Andy Stirling, and Mariachiara Tallacchini. 2007. *Taking European knowledge society seriously*. Luxembourg: Office for Official Publications of the European Communities.

Felt, Ulrike, Milena Bister, Michael Strassnig, and Ursula Wagner. 2009. Refusing the information paradigm: Informed consent, medical research, and patient participation. *Health* 13: 87–106.

Felt, Ulrike, Daniel Barben, Alan Irwin, Pierre-Benoît Joly, Arie Rip, Andy Stirling, and Tereza Stöckelová. 2013. *Science in society: Caring for our futures in turbulent times, policy briefing 50*. Strasbourg: ESF.

Felt, Ulrike, Judith Igelsböck, Andrea Schikowitz, and Thomas Völker. 2016. Transdisciplinary sustainability research in practice: Between imaginaries of collective experimentation and entrenched academic value orders. *Science, Technology & Human Values* 41: 732–761. doi:10.1177/0162243915626989.

Fisher, Erik. 2005. Lessons learned from the ethical, legal and social implications program (ELSI): Planning societal implications research for the national nanotechnology program. *Technology in Society* 27: 321–328.

Fisher, Erik, Roop L. Mahajan, and Carl Mitcham. 2006. Midstream modulation of technology: Governance from within. *Bulletin of Science, Technology & Society* 26: 485–496. doi:10.1177/0270467606295402.

Fochler, Maximilian. 2016. Variants of epistemic capitalism: Knowledge production and the accumulation of worth in commercial biotechnology and the academic life sciences. *Science, Technology & Human Values* 41 (5): 922–948.

Fochler, Maximilian, Ulrike Felt, and Ruth Müller. 2016. Unsustainable growth, hyper-competition, and worth in life science research: Narrowing evaluative repertoires in doctoral and postdoctoral scientists’ work and lives. *Minerva* 54: 175–200.

Garforth, Lisa, and Alice Cervinková. 2009. Times and trajectories in academic knowledge production. In *Knowing and living in academic research. convergence and heterogeneity in research cultures in the european context*, ed. Ulrike Felt. Prague: Institute of Sociology of the Academy of Sciences of the Czech Republic.

Gibbs, Paul, Oili-Helena Ylijoki, Carolina Guzmán-Valenzuela, and Roland Barnett, eds. 2015. *Universitites in the flux of time: An exploration of time and temporality in university life*. London: Routledge.

Glerup, Cecilie, and Maja Horst. 2014. Mapping ‘social responsibility’ in science. *Journal of Responsible Innovation* 1: 31–50. doi:10.1080/23299460.2014.882077.

Guston, David H., and Daniel Sarewitz. 2002. Real-time technology assessment. *Technology in Society* 24: 93–109.

Hecht, Gabrielle. 2001. Technology, politics, and national identity in France. In *Technologies of power: Essays in honor of Thomas Parke Hughes and Agatha Chipley Hughes*, ed. Michael Thad Allen and Gabrielle Hecht, 253–294. Cambridge, MA: The MIT Press.

Hilgartner, Stephen, Barbara Prainsack, and J. Benjamin Hurlbut. 2017. Ethics as governance in genomics and beyond. In *The handbook of science and technology studies*, ed. Ulrike Felt, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr, 823–851. Cambridge, MA: MIT Press.

Hobsbawm, Eric, and Terence Ranger, eds. 1983. *The invention of tradition*. Cambridge: Cambridge University Press.

Jacob, Marie Andrée, and Annelise Riles. 2007. The new bureaucracies of virtue: Introduction. *PoLAR: Political and Legal Anthropology Review* 30: 181–191. doi:10.1525/pol.2007.30.2.181.

Jasanoff, Sheila. 2003. Technologies of humility: Citizen participation in governing science. *Minerva* 41: 223–244.

Jasanoff, Sheila, and Sang-Hyun Kim, eds. 2015. *Dreamscapes of modernity. Sociotechnical imaginaries and the fabrication of power*. Chicago: Chicago University Press.

Kerr, Anne, and Dagmar Lorenz-Meyer. 2009. Working together apart. In *Knowing and living in academic research. Convergence and heterogeneity in research cultures in the european context*, ed. Ulrike Felt, 127–168. Prague: Academy of Sciences of the Czech Republic.

Larkin, Brian. 2013. The politics and poetics of infrastructure. *Annual Review of Anthropology* 42: 327–343. doi:10.1146/annurev-anthro-092412-155522.

Lévi-Strauss, Claude. 1963. *Structural anthropology*. New York: Basic Books.

Macfarlane, Bruce. 2007. Defining and rewarding academic citizenship: The implications for university promotions policy. *Journal of Higher Education Policy and Management* 29: 261–273. doi:10.1080/13600800701457863.

Owen, Richard, John Bessant, and Maggy Heintz, eds. 2013. *Responsible innovation. managing the responsible emergence of science and innovation in society*. Chichester: Wiley.

Ribeiro, Barbara E., Robert D. Smith, and Kate Millar. 2017. A mobilising concept? Unpacking academic representations of responsible research and innovation. *Science and Engineering Ethics* 23: 81–103. doi:10.1007/s11948-016-9761-6.

Shapin, Steven. 2008. *The scientific life: A moral history of a late modern vocation*. Chicago: The University of Chicago Press.

Shore, Cris. 2008. Audit culture and Illiberal governance. *Anthropological Theory* 8: 278–298. doi:10.1177/1463499608093815.

Shore, Cris, and Susan Wright. 2015. Governing by numbers: Audit culture, rankings and the new world order. *Social Anthropology* 23: 22–28. doi:10.1111/1469-8676.12098.

Stark, David. 2009. *The sense of dissonance: Accounts of worth in economic life*. Princeton: Princeton University Press.

Stilgoe, Jack, and David H. Guston. 2017. Resposible research and innovation. In *Handbook of science and technology studies*, ed. Ulrike Felt, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr, 853–880. Cambridge, MA: MIT Press.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42: 1568–1580. doi:10.1016/j.respol.2013.05.008.

Stirling, Andy. 2008. "Opening up" and "closing down": Power, participation, and pluralism in the social appraisal of technology. *Science, Technology & Human Values* 33: 262–294.

Strathern, Marilyn. 2000. The tyranny of transparency. *British Educational Research Journal* 26: 309–321.

Suchman, Lucy. 2013. Consuming anthropology. In *Interdisciaplinarity: Reconfigurations of the social and natural sciences*, ed. Andrew Barry and Georgina Born, 141–160. London/New York: Routledge.

van der Burg, Simone, and Tsjalling Swierstra. 2013. *Ethics on the laboratory floor*. Basingstoke: Palgrave Macmillan.

von Schomberg, René. 2011. Prospects for technology assessment in a framework of responsible research and innovation. In *Technikfolgenabschätzen lehren: Bildungspotenziale transdisziplinärer Methoden*, ed. M. Dusseldorp and R. Beecroft. Wiesbaden: VS Verlag.

———. 2013. A vision of responsible research and innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. Richard Owen, John Bessant, and Maggy Heintz, 51–73. Chichester: Wiley.

Ylijoki, Olili-Helena. 2005. Academic nostalgia: A narrative approach to academic work. *Human Relations* 58: 555–576. doi:10.1177/0018726705055963.

———. 2015. Conquered by project time? Conflicting temporalities in university research. In *Universities in the flux of time. An exploration of time and temporality in university life*, ed. Paul Gibbs, Oili-Helena Ylijoki, Carolina Guzmán-Valenzuela, and Ronald Barnett, 94–107. London/New York: Routledge.

Zwart, Hub, Laurens Landeweerd, and Arjan van Rooij. 2014. Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'. *Life Sciences, Society and Policy* 10: 1–19.

Chapter 5
The Conceptualization of RRI: An Iterative Approach

Pim Klaassen, Frank Kupper, Sara Vermeulen, Michelle Rijnen, Eugen Popa, and Jacqueline Broerse

Abstract To stimulate research and innovation (R&I), to contribute to the solution of societal challenges and to align R&I with societal values, the European Commission has launched the governance framework of *Responsible Research and Innovation* (RRI). RRI figures in many high-level EU policies as a means to promote smart growth, and a growing community of R&I practitioners from both the public and private sectors appears committed to it. Although debates on what RRI precisely entails have not reached closure yet, RRI provides an interesting avenue to explore ways of making R&I more societally germane. While recognizing the usefulness of keeping critical reflection on RRI's meaning alive, we suggest that to make the step from theorizing to implementation, RRI could benefit from a clearer conceptualization. This chapter presents the iterative trajectory in conceptualizing RRI followed as part of RRI Tools, one of a number of EC-funded research projects and support acts aimed at fleshing out what RRI can and should be, and the conceptualization of RRI that this led to. It suggests that RRI is best captured if in R&I governance attention is paid to the five p's of *Purpose*, *Products, Processes, Preconditions* and *People*, and that further elaborations on the meaning of RRI should happen in dialogue with attempts at practicing RRI.

Frank Kupper, Sara Vermeulen and Michelle Rijnen contributed equally to this work.

P. Klaassen (✉) • F. Kupper • E. Popa • J. Broerse
Athena Institute, Vrije Universiteit Amsterdam, Amsterdam, North Holland, The Netherlands
e-mail: p.klaassen@vu.nl

S. Vermeulen
Athena Institute, Vrije Universiteit Amsterdam, Amsterdam, North Holland, The Netherlands

Radboud Universiteit Nijmegen, Nijmegen, The Netherlands

M. Rijnen
Athena Institute, Vrije Universiteit Amsterdam, Amsterdam, North Holland, The Netherlands

Dutch Cancer Society, Amsterdam, The Netherlands

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_5

5.1 Introduction

We are faced with global crises in the spheres of climate, finance and food and with trends including ageing populations, environmental degradation and rising disparities in income and wealth (World Economic Forum 2016). All of these pose a challenge to the resilience of the organizational and governance arrangements of our societies and economies. Efforts are being undertaken to deal with these crises and work is being done in response to today's risks to our planet and its inhabitants. This is for instance illustrated by the recent UN agreement on sustainable development goals signed in September 2015 and the Paris Agreement under the United Nations Framework Convention on Climate Change that went into effect on 4 November 2016.

Arguably, all such challenges can only be tackled through concerted action by actors at societal levels from business to policy and from civil society to research and innovation (R&I). In this chapter, we will focus on how R&I can contribute to solving today's complex problems and respond to today's risks. One of the issues this brings us to, is that although R&I's role with regard to, for instance, the fight against infectious diseases, malnutrition or climate change, might be crystal clear to some, it is also debated. Thus, R&I might for instance contribute to better and more affordable healthcare, to more efficient and different resource use, to the transition to a bio-based economy and so on, but R&I also sparks controversies—for instance over UMTS, carbon capture and storage, use of genetically modified organisms for fuels or foodstuffs, or geoengineering.

In this light, it is of no small importance that the European Commission has identified seven Societal Challenges to be dealt with in its research funding programme Horizon 2020.[1] Moreover, in the EU we have seen that during the last 6 years both at the EC-level as well as through actions by research funding organizations and academic researchers, efforts have been put into developing and implementing a governance framework aimed at directing R&I efforts to more responsible ways of working: *Responsible Research and Innovation* (RRI). RRI has been proposed as a unifying framework that aspires to integrate ethical reflection, stakeholder engagement and responsive change into research and innovation (R&I) practices (Stilgoe et al. 2013).

In parallel with this narrative that presents RRI in relation to complexities of the world we inhabit today, RRI's emergence can also be explained with reference to (not-independent) developments in philosophical and sociological studies of R&I, R&I policy, Technology Assessment in all its well-known versions, and so on (Owen et al. 2012). Overall, what the past two decades in these fields of study show, is an increasing focus on all possible forms of interaction between R&I and society. The articulation of this theme has been recorded and discussed, inter alia, by (Nowotny et al. 2001; Etzkowitz and Leydesdorff 2000; Regeer and Bunders 2009; Callon

[1] These seven can all be found here: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/societal-challenges

et al. 2009). What these views have in common, despite all sorts of differences in emphasis, is the recognition that R&I processes are not assessed solely internally (by scientists themselves) and disciplinarily (by using domain-specific criteria) but also externally (by society) and inter-, multi- and trans-disciplinarily. In addition, they acknowledge that the purpose for which knowledge is produced goes beyond the mere quenching of the scientific thirst for knowledge, so as to include solving real-life problems.

Although RRI gains popularity, closure has not yet been reached with regard to the concept's meaning. For instance, Oftedal notes that "the more specific content of RRI is largely left open" (Oftedal 2014, p. 1) while Zwart et al. describe RRI as a buzzword whose conceptualization is "open-ended" (Zwart et al. 2014, p. 3) and the source of "confusion". Wickson and Carew also subscribe to the idea that "without concrete elaboration and conceptual development, the interpretive flexibility of RRI will be so broad as to render the concept meaningless" (2014, p. 256). And even scholars whose names almost immediately pop up when RRI is discussed, have expressed concerns regarding the vagueness surrounding the very idea of responsibility in research and innovation. Thus, Owen, Macnaghten and Stilgoe note that the notion suffers from "ambiguity as to motivation, theoretical conceptualisation and translation into practice" (Owen et al. 2012, p. 751).

This brings us before a quandary. On the one hand, the flexibility in the notion of RRI is expedient since it provides a conceptual space for assimilating and comparing diverse approaches that have been developed in the past before the notion of RRI entered the scene. Those who had already been working on specific aspects of responsibility in research and innovation (e.g., making science and innovation responsive to societal needs) will find in RRI a useful mainstay and an opportunity for reflection. If RRI is to work as a guiding concept (De Jong et al. 2016), RRI must allow for at least some interpretation and thus variation. On the other hand, the flexibility of the notion can also be detrimental to its application. We must not lose sight of the fact that the scholarship on the notion of RRI is also an instance of research and innovation. Thus, being true to form, we should appraise it based on the same standards that we use to observe others in their research and innovation practices. In short, if RRI is to be more than a sweet-sounding buzzword, it should eventually be crystalized into a policy instrument that achieves what it claims to achieve.

In this chapter, we want to present our way out of this quandary. We will show that, despite what common sense might suggest, an increase in analytical clarity does not necessarily imply a decrease in interpretive flexibility. Quite the contrary, if an abstract concept such as RRI is ever to become a sustainable force in shaping R&I practices, then we should not shy away from rejecting the old distinction between ideals (dreams) and practices (reality). Moreover, we will not only present what we found at the end of our road, but also that road itself. With respect to this we can say that we must seek conceptualization methods that make the most of both our ability to dream the ideal-thus-unspecific and our ability to observe and learn from the concrete-thus-specific.

On the whole, the route we took led us to a better understanding of RRI, an understanding we are now ready to flesh out and reflect upon. What we have found

is not an unyielding answer to the question 'What is RRI?'. Rather, we have reached what we see as a sensible approximation of this solution, one that is capable of reconciling the need for abundant dreaming and concrete governance actions—and perhaps, even one that inspires both such dreaming and such actions.

5.2 Laying the Path While Walking It: Outline of Our Iterative Exploration of What RRI Means

The ideas presented in this chapter are largely developed in the context of EC-funded FP7 support action *RRI Tools*. The project's aim was to foster RRI through the development of a toolkit tailored to the use in implementation of RRI by users from different R&I stakeholder groups and through training and advocacy activities. A multidisciplinary consortium consisting of 26 partners operating in 30 European countries collaborated on this.[2]

One of our roles in this project was the conceptualization of RRI that would be central to the different project tasks. What we share here, however, is not the academic version of an official project deliverable, but rather an essay that provides insight in the process of informal iterative concept development that we have engaged in throughout the project, and into the preliminary conclusions regarding RRI that based on that process we have managed to draw. Some such conclusions can in a different, abbreviated form be found in deliverables that are available on the *RRI Tools* website (Klaassen et al. 2014). One reason for that is while formally the conceptualization of RRI was a task that belonged to Work package 1 and that was finished in 2015, our process of constantly re-imaging RRI continued with all the different (other) tasks we engaged in in the context of this project. What we present is in fact something like a rational reconstruction of our iterative conceptualization process throughout the project in light of what these have led us to conclude as regards the concept of RRI.

Six different project activities contributed to our understanding of the RRI concept: (1) literature review, (2) expert consultation, (3) stakeholder workshops, (4) identification and classification of promising practices, (5) specification and refinement, and (6) case-studies. Each of these contributed in a specific way to the resulting image of RRI. Vice versa, each of these six processes were informed by a certain ('raw') image of RRI, the image that we had at that specific moment when the concept was still in the making. This two-way relationship between the model and the six steps in gathering data and information is represented in Fig. 5.5. Although these steps will now be discussed in the indicated order, it is important to note that most of the six overlapped in time and were thus informed by one another. In this

[2] We feel indebted to all our colleagues in the RRI Tools project and would like to express our gratitude to them. Amongst other things, the partners included research funding foundations, universities, science centres and museums. For a complete list, see here: http://www.rri-tools.eu/who-we-are

way, we managed to compare, early on during the conceptualization process, the various conceptions of RRI that arose from each source. This lead to what we see as a very fruitful blend between what RRI *is* to various stakeholders and what RRI *should be* according to the same stakeholders.

The *literature research* included both academic and policy literature. The ground covered mainly concerned literature explicitly addressing RRI, but we also built on the plethora of conceptual, theoretical and empirical resources that fuel RRI—from constructive Technology Assessment to public engagement, from Gender Studies to research ethics, from STS to science communication. Early 2014 a very first working definition of RRI was developed, for use in the first stages of the RRI Tools project. According to this definition, RRI is a dynamic, iterative process by which all stakeholders involved in the R&I practice become mutually responsive and share responsibility regarding both the outcomes and process requirements.

During the *expert consultations,* the first ideas on the delineation and operationalization of RRI were elaborately discussed with experts from a wide range of fields pertinent to RRI. To wit, we discussed our preliminary conceptualization of RRI with the Advisory Board members of the RRI Tools project as well as with other experts within the RRI Tools group. The Advisory Board members were selected based on their expertise on the different RRI "keys" as identified by the EC: Ethics, Gender, Equality, Governance, Open Access, Public Engagement and Science Education; each key being represented by two experts.[3] The feedback we received found its way into the project's first deliverable, a Policy Brief on RRI (Klaassen et al. 2014). Gradually, a highly specialized community of experts has arisen, as a result of the RRI Tool project's aim and effort to build an RRI community of practice. Scholars from fields like Science and Technology Studies, philosophy of science, science communication, Technology Assessment, research ethics and research policy studies have interacted with one another, emerging as experts on RRI. However, these experts agreed that RRI should not be an idea that can only be grasped by a small intellectual elite. All actors that have an interest in research and innovation should translate this central idea within their own domains and this translation should lead straightforwardly to implementation.

With this in mind, *stakeholder workshops* were organized during the fall and winter of 2014. A total of 27 stakeholder consultation workshops were organized with stakeholders representing the following five domains: research, policy, business/industry, civil society and education. During the workshops, stakeholders were acquainted with the concept of RRI, invited to discuss RRI and to help the RRI consortium of RRI Tools to identify the opportunities, obstacles and needs they experience as regards putting RRI into practice. Workshops were held in 22 different countries, and 411 participants took part in them. The workshops provided us with valuable insights regarding the opportunities, obstacles and needs experienced by various groups whose work can be improved by a new research and innovation framework. Since these groups are driven by different social, economic and moral interests, the consultation workshops were also employed as an opportunity for the

[3] The experts of the Advisory Board are listed here: http://www.rri-tools.eu/en_GB/who-we-are

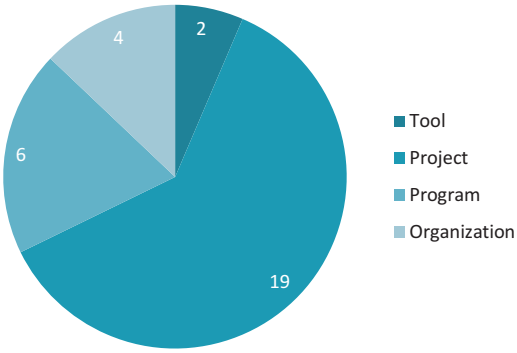

Fig. 5.1 Selected practices – 31 in total

stakeholders to hear each other's viewpoints on research and innovation. The discussions, which often revealed surprising differences and equally surprising similarities in worldview, were an eye-opening moment during the conceptualization process.

The workshops constituted a point of departure for developing a *catalogue of good practices*, in addition supplying a much-needed input regarding opportunities, obstacles and needs (Kupper et al. 2015b). All 411 participants in the stakeholder consultation workshops were invited to share one or more examples of research innovation practices that instantiate RRI to a greater or lesser extent. These cases could be research and innovation projects, but also funding programs and organisations related to research and innovation (see Fig. 5.1).

The assumption underlying the request to workshop participants to bring examples of RRI practices, is that concepts – as sets – can best be described by combining an intentional definition in which the criteria for set-membership are spelled out in general terms (viz., the working definition) with an extensional definition in which members of the set are enumerated (viz., the catalogue of RRI practices). Having collected these practices, a first selection of so-called 'promising practices' was made, leaving those out that did not meet any of the process requirements and/or outcomes of the RRI working definition. Hereafter, a database of additional promising practices was developed by making use of an online questionnaire. Together with the first selection, the body of good practices was now studied and assessed. From all these suggestions 31 practices ended up in an RRI catalogue of good practices. Some descriptive statistics concerning these practices can be found in Figs. 5.1, 5.2, 5.3 and 5.4. As these figures indicate, the practices included in the catalogue all dealt with one or more of the so-called policy agendas of Public Engagement, Science Education, Governance, Ethics, Open Access or Gender (Fig. 5.2), were all rather *inclusive* in terms of the amount and types of stakeholders they managed to assemble together (Fig. 5.3), and all contributed to one or more of the EU-defined Grand Challenges (Fig. 5.4).

Through an examination of the good practices and by revisiting the literature reviewed in the first step, we formulated a set of *criteria and indicators* for RRI (Kupper et al. 2015a). In various stages of this development, we applied the formu-

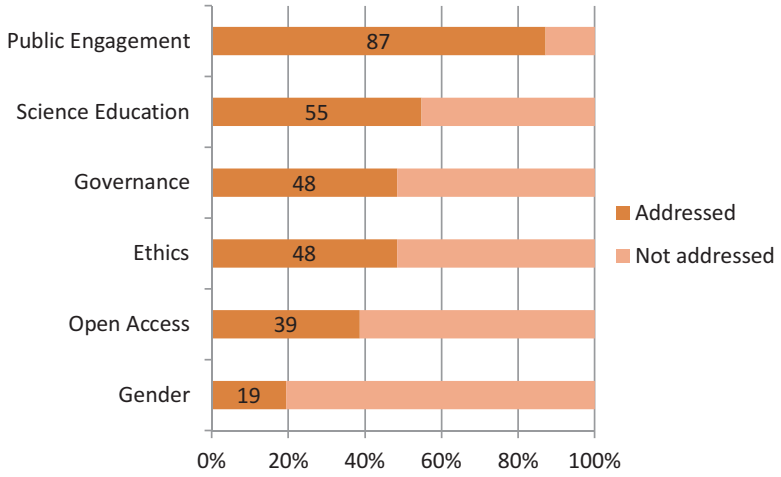

Fig. 5.2 policy agendas addressed in percentages (out of 31 practices)

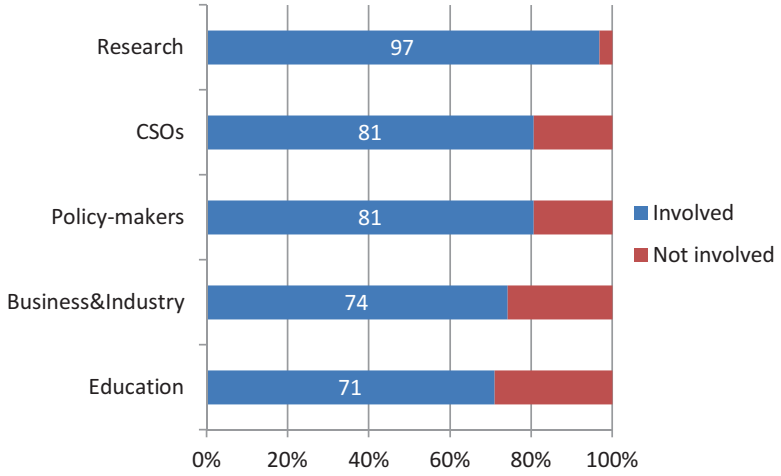

Fig. 5.3 stakeholders involved in practices in percentages

lated criteria to the selected promising practices mentioned, changing the final formulation so as to encompass as many of these practices as possible. At the same time, we kept an eye on the systematicity of the resulting set of criteria and indicators. We organized, merged and split some of these indicators in order to obtain an analytical instrument that is at the same time expedient (minimal overlap) and thorough (maximal applicability). In this way, i.e., by going back and forth between theoretical formulation and empirical application, we have sought to maintain the flexibility of the concept of RRI while increasing its clarity.

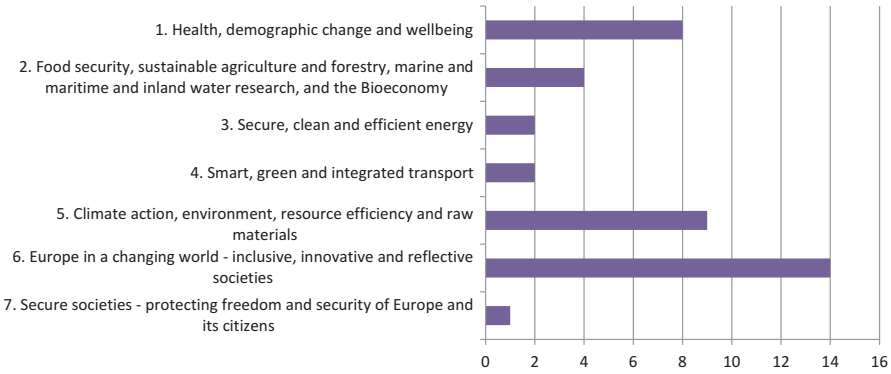

Fig. 5.4 Grand challenges addressed by the 31 selected practices

While the previous steps were successful in (abstractly) clarifying the nature of RRI, little in-depth suggestions were provided regarding the factual implementation of RRI. This is why we continued by selecting eight *showcases* and analysing them thoroughly. The analysis was based on semi-structured interviews with experts on the particular cases (mostly project or programme managers). This resulted in a series of eight elaborate narratives regarding responsibility in research and innovation, each delivering important lessons to be learned about the contemporary constraints and opportunities for applying RRI.[4] Table 5.1 briefly describes all eight and presents one distinctively illustrative lesson learned from each showcase.[5]

From the spring of 2016 onwards, these showcases have been used in training events on RRI throughout Europe, along with an abundance of other materials, facilitated by RRI Tools consortium members and affiliates. During such training sessions, again, feedback on the proper conceptualization of RRI was collected (Fig. 5.5).

5.3 Five Components of RRI

These iterative processes have brought us in a better position to tell a more refined story of RRI. Although in what follows we will tell this story with the conviction that it is the right story to tell, the one that most naturally follows from our iterative approach, we do not wish to suggest that the version here presented is the definitive one. Presenting work in progress might perhaps be at odds with current academic conventions. For conventionally, publishing and defending one's conclusions is something that takes place *after* the discovery has taken place. We cannot but reject such linearity. In our case, the process of discovering RRI through a continuous and multifarious interaction with various stakeholders *was* the process of building a case

[4] These can be found here: www.rri-tools.eu/training/resources

[5] See https://www.epsrc.ac.uk/research/framework/area/

Table 5.1 RRI showcases other iterative steps can of course complement the six represented above. It is not our intention to suggest that these six categories are in some sense sufficient or carved out in stone—they are simply the key steps we took on our journey. Keeping in mind the preliminary character of any result reached through this type of iterative conceptual modelling, it is both useful and personally rewarding to pause the modelling process and take a look at the results. In the next section, we will give a brief description of what we have so far learned about RRI

Showcase name (country)	Brief description	Lessons learned
Vinnova's funding programme *Challenge Driven Innovation* (Sweden)	By funding research and innovation in consortia of partners that come from different societal organizations, cross-disciplinary, cross-sectorial and challenge-oriented research and innovation are promoted. The focus is on (i) Future healthcare, (ii) Sustainable attractive cities, (iii) Competitive industries, and (iv) Information society. The programme features a three-stage process, allowing large numbers of projects access to funding in the stage of idea development and testing, with smaller numbers of projects moving to subsequent stages in the research and innovation trajectory.	Successfully implementing a challenge-driven research programme requires strong leadership, the courage to change ingrained structures and working methods, commitment of agency staff and an openness to organizational learning through processes of trial and error.
Framework for responsible innovation by the engineering and physical sciences research council (UK)	The AREA "code of conduct" of anticipate, reflect and engage is at the heart of the EPSRC's framework for responsible innovation. It requires of researchers that they not only have good ideas, but also consider what potential consequences their research might have. This Framework aspires to convey that the two are not separate matters, but rather are part of the same package deal.	To implement RRI in (academic) research, rules, regulations or specific grant conditions might not be the most pertinent facilitators. Rather, for the research community to embrace RRI, a framework such as this, with a proven track-record in helping researchers deal with societally challenging issues such as geo-engineering, appears more promising.
Stakeholder engagement in Fishery benchmarking research at The Portuguese Sea and Atmosphere Institute (Portugal)	Industrial, policy, research and societal stakeholders, all with a role in marine conservation, interacted in all stages of this research project designed to address a critical aspect in the sustainable use of marine resources—from agenda setting to follow-up.	The active engagement of all types of stakeholders contributes to the pertinence of the research, helps its products become accessible and facilitates implementing actions. However, different types of stakeholders — policy makers, researchers, representatives of industries and of Civil Society Organizations — all have their own specific roles in facilitating and building co-creation partnerships, and sometimes difficult conflicts can emerge.

(continued)

Table 5.1 (continued)

Showcase name (country)	Brief description	Lessons learned
Hao2 (UK)	"Social company" Hao2 develops and sells 3D virtual environments, with the specific aim not only to make money, but also to increase opportunities for people with autism and other complex needs. Its own workforce consists of some 80% of people with disabilities like autism.	Embracing gender and disability issues as integral part of one's business activities can lead to opportunities and growth. Openness, diversity and inclusion can be drivers of success rather than obstacles, as they can help companies become responsive and adapt to changing needs.
Knowledge for climate (The Netherlands)	To transform the Netherlands' vulnerability to climate change into opportunities, this programme aimed to increase knowledge about climate adaptation and improve the Dutch export position in climate and delta technology. It did so through co-creative projects, in which research, solutions and results resulted from dialogues between practicing professionals, policy-makers and scientists.	Boundary workers with the right knowledge and skills and sufficient time to promote mutual trust and project continuity are vital to successfully engage in co-creative research that involves researchers, policy makers and industry professionals.
Novo Nordisk's *blueprint for change* (Denmark)	Under the name of *Blueprint for Change*, pharmaceutical company Novo Nordisk developed a series of business cases aimed at identifying drivers of shared value creation, the measurement of societal and company benefits, and the sharing of information with stakeholders. Collaborations with research and local and national societal partners have been undertaken, in order for projects to serve both societal, environmental and economic success.	By taking decisions in ways that are financially, socially as well as environmentally responsible, the private sector can be a valuable partner in solving societal issues. Doing so requires that investments are made towards long term partnerships that cross sectoral borders.
Social innovation factory (Belgium)	A networking organization that searches for possible partners who can help strengthen concepts for social innovation, and that promotes, guides and supports businesses and their stakeholders in doing so.	Building networks requires a skill-set of its own, and is an important requirement towards realizing creative social innovations.
Xplore health (Spain)	An educational programme aspiring to bridge the gap between research and secondary Science, Technology, Engineering and Mathematics (STEAM) education, with an innovative educational approach that includes acquainting students with decision-making on science and innovation and incorporating insight on real-life challenges therein, as well as ethical, legal and social issues.	Implementing RRI in STEAM education enriches students' perspective on science and innovation and contributes to their empowerment with respect to finding solutions for societal challenges.

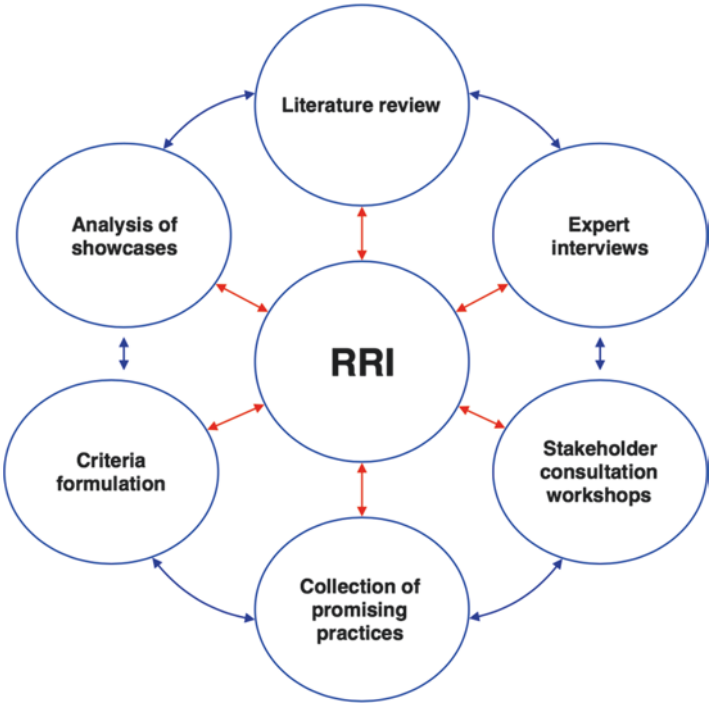

Fig. 5.5 The six activities involved in Iterative concept development

for a certain version of RRI. Vice versa, the process of concretizing all those hundreds of hours of data analysis into a model we defend as 'correct' is as much defending a stance as it is discovering it. The uncanny feeling of deciding to follow a certain ideal (RRI) while looking for that ideal can best be compared with the uncanny feeling one typically has while looking at a *mise en abîme* (which in The Netherlands we know as 'the Droste effect'). The uncanniness of it stems from the fact that an idea is employed in a discussion in which that very idea is at issue.

In what follows we want to distinguish between five components of the concept of RRI. We suspect that this five-fold structure is typical of concepts representing ideals we pursuit in other settings, ideals such as 'justice' and 'reasonableness', yet for the present purposes we will assume it to be an expedient way of crystalizing the idea of RRI. The five components are: *Purpose*, *Product*, *Process*, *Preconditions* and *People*. We refer to these informally as 'the 5P structure'. Each of these five components represents a specific vantage point for understanding RRI story. Each is thus essential for obtaining a full-fledged image RRI but also for distinguishing the kind of research that is further needed for giving this image more depth and perspective.

We will discuss these five components in the order given above, as this order represents what we have found to be a natural way of asking questions about RRI. The first question that comes to mind is: "What is the *purpose* of changing

current R&I environments – in any direction, not just towards an ideal of shared responsibility?" Having established a certain purpose, one can turn to questions regarding the more concrete outcomes that together would realize the designated purpose. The second question is thus: "What kind of *products* need to be obtained in order to eventually realize the designated purpose?" We assume there is always more than one way to obtain these products, so the third question concerns the manner in which the needed products are obtained: "Through what *processes* will these products be obtained?" Answering this third question will bring researchers in a better position to specify the kind of institutional setting in which these processes are to take place. To draw a parallel: deciding what music (process) you want to play on a certain instrument, say a guitar, will bring you in a better position to specify the needed characteristics of that instrument – whether you need an electric guitar or an acoustic one, a jazz guitar or a classical one. Our fourth question will thus concern the conditions under which the desired processes are to be created: "What institutional *preconditions* are necessary for hosting the development of the desired processes?" One might perhaps stop the questioning process at this fourth component, given that the entire setting is specified, top-down from an abstract description of purpose to the concrete settings in which this purpose is to be pursued. Yet the more we interacted with colleagues on the theme of RRI, the more we acknowledged the importance of the individuals' psychological predispositions and competences. The fifth and final question is thus: "What kind of individuals function well and efficiently in the designated institutional preconditions?" It holds for all stakeholder groups that fostering RRI from the perspective of that group is a very specific mission and that this mission requires a specific set of competences.

The burgeoning field of RRI can be seen as the systematic attempt to find an answer to these five questions at the present time all these five questions have been addressed in some form or another. However, some have inevitably received more attention than others. In what follows we will offer a brief overview of these five components in the way they result from our iterative conceptual modelling.

5.3.1 Purpose

The European Commission has identified seven societal challenges with which the European (and possibly international) society is nowadays confronted. These challenges, also known as the "grand challenges" are broad, long-term purposes that have been set through a simultaneous look at the past (European Environment Agency 2002, 2013) and at the future (Boden et al. 2010). The seven grand challenges range from health and wellbeing to sustainable energy and secure societies.[6] These seven challenges demand a contribution of research and innovation. At the same time, however, research and innovation themselves are contested in the pub-

[6] For more details regarding each challenge, see https://ec.europa.eu/programmes/horizon2020/en/h2020-section/societal-challenges

lic sphere. Issues that are raised for example relate to a lack of sensitivity to societal needs and concerns, the distribution of (new and unforeseen) risks and benefits, and emerging ethical controversies.

These ethical and societal aspects have long been described in the technology assessment and ethics literature (Rip et al. 1995; Schot and Rip 1997; Schomberg 2007), and were expressed in the consultation workshops in this project too. RRI aspires to deal with both issues at the same time. One of the major shifts in the RRI framework therefore is a primary focus on the question of purpose: what is the contribution that research and innovation can make? Rather than the effort to 'do things right', i.e. carefully investigate in order to mitigate potential negative impacts alongside the development of science and technology, RRI aspires to 'do the right things'. In the words of Owen et al. (2012), RRI seeks to move beyond what we don't want R&I to do towards what we do want R&I to do. To establish this shift, and create a productive environment to ask the question of purpose, RRI aspires to democratically open up research and innovation to processes of inclusive deliberation involving a variety of actors, tightly coupled to action and policy-making aimed to steer research and innovation towards desirable and acceptable ends. The different dimensions of these processes are discussed in 3.3.

5.3.2 Products

The grand challenges formulated by the European Commission constitute long-term purposes for research and innovation. Realizing such purposes will not be the result of any specific research and innovation process. Even large-scale research and innovation projects cannot, in one stroke, solve such complex issues as the sustainability of our economic processes and the security of our society. Furthermore, it would be highly unrealistic to hope that stakeholders involved in research and innovation would reshape their worldview overnight and reorganizing their professions around these seven grand challenges. It is thus necessary to distinguish between the purposes that make up the *raison d'être* of RRI and the short-term products that bring us closer to achieving these purposes.

Focusing therefore on the short-term adjustments, we have found that, in both literature and stakeholders' views, there is a natural inclination to make a distinction between two kinds of products resulting from research and innovation. On the one hand, there are products that constitute (proposed) solutions to research and innovation questions. We refer to these as 'R&I products'. On the other hand, there are products that, while not directly solving any research and innovation problem, create the proper social and cultural environment in which the research and innovation can take place. We refer to these as 'learning products'.

Following the definition suggested by Von Schomberg (2011), we started our conceptualization process from the assumption that responsible R&I products are ethically acceptable, sustainable and societally desirable. One of the main questions here is: when and how are processes and products ethically acceptable? Part of the

answer can be derived from accounts such as the Treaty on European Union (art. 2) that lists the values (supposed to be) shared in European societies like respect for human dignity, freedom, and equality. Other examples of shared values are wellbeing privacy, autonomy and security (Van den Hoven 2013). However, as we live in a pluralistic society, the interpretation of these moral values may differ between different cultural regions in Europe, but also between different people and groups. We argue that defining ethical acceptability in light of RRI implies an exploration of presumably common values and principles (to understand their situated meaning) and ethical assessments that go beyond protecting the rights, interests and desires of moral subjects (in line with Keulartz et al. 2004). Which values and norms contribute to a specific case of responsible research and innovation should be discovered through a process of reflective inquiry and deliberation between the stakeholders involved. With respect to sustainability, approximately the same argument can be made. Sustainable development is explained as meeting the needs of present generations without jeopardizing the ability of generations to come to meet their own needs (The Council of the European Union, 2006). In specific research and innovation contexts, however, the contribution to sustainability has to be a matter of inquiry and deliberation amongst the actors involved. With respect to societal desirability, an important observation is that science and society are continuously evolving together, subject to the same evolutionary trends. Boundaries are increasingly transgressed and new collaborative modes of knowledge production emerge (Gibbons et al. 1994; Nowotny et al. 2001). Solutions are found in opening up science through continuous meaningful deliberation with societal actors (Broerse et al. 2009). By incorporating such activities in the R&I process, science as a whole is thought to become more responsive to real-felt social needs, concerns, ambitions and interests (Haywood and Besley 2014). If public concerns and needs are understood, the likelihood of R&I processes and new technologies being successful increases, i.e. innovations and the design thereof will be consistent with needs of society. It is by now evident that RRI involves a shift in thinking from product to process. An important characteristic of this process is mutual learning of the actors involved. It is therefore good to also distinguish the learning products of RRI processes.

Learning products contribute to RRI because they create a kind of purposeful change in which responsibility is more easily, more often and more naturally achieved. Responsible research and innovation processes are fore example meant to lead to a wide range of empowered, responsible and reflexive stakeholders (researchers, policymakers, NGOs, educators, businesses etc.). It follows that research and innovation should not only lead to a certain form of specialized knowledge, whether incorporated in a technological product or expressed explicitly in written works, it should also lead to a closer relationship between science and society. One important aspect of that relationship is engagement of the general public, not only to develop a robust understanding of scientific work but also understanding socio-scientific issues and to become involved in deliberation and decision-making processes. Although such learning products were seen by many stakeholders as ‘by-products’ of science, these products’ importance in fostering responsibility has been widely

recognized. Further, as some stakeholders have noticed during the consultation workshops, undertaking to deliver both R&I products and learning products might change the actors' propensities and interests, leading them towards research questions and puzzles that are conducive to contributing to the solution of, amongst other things, the seven grand challenges.

5.3.3 Processes

The first two components provide a reference point for designing the processes through which the aforementioned long-term aims (purposes) and short-term aims (products) are to be achieved. We now want to focus on the path towards these aims. In doing so we distinguish RRI processes as the ones satisfying the following criteria (or 'process requirements'): (1) Diversity and Inclusion, (2) Anticipation and Reflection, (3) Openness and Transparency and (4) Responsiveness and Adaptive Change.

Diversity and inclusion refers to the early involvement of a wide range of stakeholders and publics in the deliberation and decision-making episodes that occur within research and innovation processes. This is accomplished through the timely and on-going involvement of a wide range of stakeholders and publics in deliberation and decision-making processes (Owen et al. 2012).[7] In different scholarly traditions, such as public engagement (Irwin et al. 2012; Wilsdon et al. 2005) and technology assessment (Palm and Hansson 2006), a participatory-deliberative turn has been argued for because of both normative democratic, instrumental and substantial reasons (Abelson et al. 2003; Wilsdon and Willis 2004).

A central issue in these and other studies concerns the right timing for engaging stakeholders. It is notoriously difficult to say with precision what 'the right timing' is. Indeed, the dilemma coined by Collingridge (1980) points precisely to the tension between the advantages brought by acting early (the ability to steer the research and innovation process in the desired direction) and the equally important advantages of acting late (knowledge regarding opportunities and limitations of the chosen direction). Still, to strive towards socially desirable (ethically acceptable, sustainable, and marketable) outcomes and to prevent misjudgements regarding each other's interests, it is vital to have stakeholders articulate their standpoints relatively early in the research and innovation process (Schot and Rip 1997). Subsequently, discussions about ideas and values should be carried out continuously as a groups values and interests may change during the R&I process (Abma and Broerse 2010).

In addition to the question of timing, the question arises what it means, in practice, to engage stakeholders within the research and innovation process. A genuine

[7] Diversity is understood here in relation with demographic variables such as age, gender and education level as well as cultural variables such as values, interests, religion and worldview (Von Schomberg 2011; Wilsdon and Willis 2004).

dialogue fosters mutual learning processes, in which actors in the dialogue listen to each other, learn about and understand each other's perspectives, and add new experiences to their repertoire (Stilgoe et al. 2013). Active participation of actors in dialogue, the willingness to share power, being respectful and open to others, and the ability to change one's own perspective, are all important conditions for constructing a genuine dialogue (Abma and Broerse 2010; Abma and Widdershoven 2006; Chilvers 2012). How these conditions are facilitated ultimately depends on the situation at hand and the relevant actors involved. The wishes and needs of actors vary between practices and need to be taken into account not only concerning the topic at hand, but also in constructing the dialogue itself.

Anticipation and reflection refers to understanding how the dynamics of R&I shape the future; envisioning the impacts of dominant and alternative R&I futures; reflecting on (alternative) problem definitions, preferred solutions and underlying values, assumption and beliefs. Research and innovation are unequivocally future-oriented activities, with the power of shaping and transforming our future immensely (Borup et al. 2006; Owen et al. 2012; Grinbaum and Groves 2013). This requires anticipation: looking forward in time by imagining the variety of possible impacts of research and innovation practices and reflecting on our values and roles in these practices (Schomberg 2011; Owen et al. 2012; Stilgoe et al. 2013). Anticipation can be explained as 'describing and analysing those intended and potentially unintended impacts that might arise, be these economic, social, environmental or otherwise', which is not the same as predicting the future (Owen et al. 2012, p. 38). In the past many anticipatory methodologies for science and technology have been developed, such as scenario development (Fisher et al. 2008), vision assessment (Grin and Grunwald 2000), ethical technology assessment (Swierstra 1997), constructive technology assessment (Schot and Rip 1997), and anticipatory governance approaches (Guston and Sarewitz 2002). These methods not only support actors in articulating their expectations, but provide means to explore alternative outcomes and implications that would otherwise be forgotten and help avoid reinforcing certain visions and making them into preordained roadmaps or trajectories (Owen et al. 2012). For anticipation to make sense we should be aware of how present dynamics and values influence the progression of science and innovation. This means that we should not only anticipate uncertain products of science and think about plausible, intended and unintended consequences, but that we need to reflect on underlying purposes, motivations, and actor roles as well (Owen et al. 2012).

Acknowledging that irresponsibility in science and innovation is a manifestation of the innovation ecosystem, implies that not only reflection on value systems of individual actors or institutions should take place, but that these actors and institutions also help build the collective reflexive capacity within the practice of science and innovation. A collective and institutional reflexive capacity lies at the heart of any learning process, and for research and innovation to progress – both in process as in outcomes – learning is a prerequisite. Reflexivity, or rather reflexive learning, requires both 'insight into the assumptions which tacitly shape our own understandings and interactions' by which the value of other sources of knowledge and perspectives will increase (Chilvers 2012).

Room for these reflective processes should be built into RRI practices to accomplish learning at different levels: first-, second- and third-order learning. The description of different levels of learning is found in the work of different scholars across management science, learning science and philosophy, with amongst the most influential the authors Argyris and Schon (1974). First-order learning refers to learning on the level of problem definition, possible desired solutions and routines. Convincingly argued that in case of new and complex issues, second-order learning is required, i.e. learning at the level of values and assumptions of actors involved, which means holding a mirror up to one's own activities, commitments and assumptions, being aware of the limits of knowledge and being mindful that a particular framing of an issue may not be universally held. We speak of third-order learning when a practice of research and innovation starts to transform itself and the way it is organized, connecting the process requirement of reflexivity to the dimension of change.

Openness and transparency refers to the honest and clear representation of research and innovation processes in society. By this we refer not only to correct and equal access to the products of research and innovation but also to a certain willingness to being open for and listen to input of people other than those directly involved in the research and innovation process. This willingness is a condition for responsiveness and adaptive change and should lead to a mutual understanding and trust. Transparency implies being open and clear about decision-making processes, for instance on issues such as who is included when, what is done with inputs (materials) and results in research and innovation processes (Abma and Broerse 2010; Rowe and Frewer 2004). By communicating decisions made in science and innovation policy, these processes become legitimate and both institutions and individuals can be held accountable. Moreover, open discussions about roles and responsibilities of stakeholders are indispensable, because through the evolvement of the concept of RRI new responsibilities emerge or responsibilities change and shift (Owen et al. 2012). Such discussions create awareness of roles and responsibilities and create clarity about ownership, which will ultimately lead to increased agency.

Open access to research information is argued to advance science, as it will promote and accelerate the constructive generation of new knowledge and prevent unnecessary duplication of research. Open access not only improves the quality of scientific work, but also benefit industry and government. For the wider community, it is argued that open access can benefit the 'informed citizen' or 'informed consumer', thereby improving knowledge and use of services (Houghton and Sheehan 2006; European Environment Agency 2013). Being open does not necessarily mean that raw data should be published and data sets become available without being edited. Openness should be meaningful; it needs to be understandable and usable for potential stakeholders and publics involved (Chilvers 2012). In practice, this might imply that the amount and level of openness depends on the context, situation and topic of the specific research or innovation practices.

Responsiveness and adaptive change refers to the development of a capacity to change existing routines of thought and behaviour, as well as overarching organizational structures and systems in response to changing circumstances, emerging

knowledge and value perspectives, views and concerns. This fourth cluster of process requirements is vital to RRI insofar as the capacity for change ultimately determines whether the effects of the previously described process requirements can manifest themselves. RRI requires that the direction people, organizations and practices take changes in response to (possibly changing) circumstances, values, ideas and needs of both stakeholders and the public to give true meaning to the requirements of inclusion and diversity. Second, openness and transparency are valuable from a democratic point of view, but become more significant through this fourth cluster of process requirements. It requires practices to respond to emerging knowledge, even if it is generated elsewhere, so a collective learning process can be build and R&I can be brought to a higher level. Something similar applies to anticipation and reflection. One can anticipate possible futures and reflect on one's role and actions in R&I, but without responding to changing understandings or newly emerging insights, R&I outcomes in the form of learning or desirable futures will most probably not arise. Our systems of science and innovation should thus be open to and enable transformative change by way of responsiveness. Several approaches have already been developed for increasing responsiveness in R&I processes. These include constructive technology assessment (Rip et al. 1995), real time technology assessment (Guston and Sarewitz 2002), midstream modulation (Fisher et al. 2006) and anticipatory governance (Barben et al. 2008). Responsiveness should however not be limited to a capacity for change at the level of individual researchers and or project groups, as actions of individuals are often steered by the rigidity of the systems of which they are part (Cavallo 2000). Responsiveness of R&I processes should extend beyond the responsiveness of individual researchers, and institutionally embed the capacity to adapt to changing circumstances and newly emerging knowledge in such a way that inclusive deliberation is tightly coupled to policy-making, action and change (Owen et al. 2012).

5.3.4 Preconditions

Now that RRI has been analysed in terms of purposes, products and processes, one might ask, what role do the key dimensions to RRI as identified by the European Commission play: Gender, Ethics, Open Access, Public Engagement and Science Education? In our conceptualisation, the interaction between processes, products and purposes is what makes an R&I practice RRI. However, the keys as formulated by the EC give us something like a normative baseline, a way of stating preconditions that have to be met on a *systemic* level, an *organizational* level and a *project* level in order for R&I to be able to take the shape of RRI.

To elaborate on this, we can say that for R&I to become truly RRI it is requisite that it takes place in the right environment. For this, governance repertoires need to be installed on all distinguished levels so the proper *preconditions* for making R&I responsible are created—and here is our fourth P. Focusing on the core processes distinctive of R&I projects, these can be said to be responsible if they entail open

and transparent cycles of inclusion, anticipation, reflection and responsiveness that lead to the variety of outcomes and impacts pictured above: engaged publics and stakeholder learning, responsible institutions, ethically acceptable, socially desirable and sustainable R&I outcomes, targeting the Grand Challenges. The variety of aforementioned agendas, on this view, form a subset of a number of conducive preconditions for such cycles to take place—preconditions that, in true RRI, are themselves open to change in response to the variety of types of outcomes RRI aims at.

We picture RRI to blossom optimally in organizational and systemic environments that are governed with an eye to the variety of preconditions that are conducive to RRI, ranging from the promotion of research integrity to banning exclusionary practices in both human resource management as well as research agenda setting. This means, for instance, that for research projects to become responsible, involved research institutes should have policies in place or develop them along the way of research projects taking off concerning everything from gender equality and gender in research, communicating and disseminating research results, engaging stakeholders in agenda-setting and decision-making, research integrity, open access, Intellectual Property issues, and risks and safety. On a systemic level, such preconditions include for instance incentives for academic researchers that do not exclusively promote publishing in peer-reviewed journals, but at least as much steer towards contributing to the solution of complex societal issues. For commercial R&D this would for instance require that existing guidelines and regulations for Corporate Social Responsibility (CSR) explicate what CSR means for the design and execution of R&D trajectories.

The main reason for introducing this multi-layered conceptualization of RRI, in which for instance issues relating to diversity and ethical reflection emerge both as aspects of responsible R&I cycles as well as in the form of conducive conditions, is that these conditions are not sufficient for R&I to be conceived of as *responsible R&I*, even if they might be necessary for putting RRI into practice. This can be illustrated with reference to ethics, for instance. Thus, for research in the health and life sciences, for example, it is vital to have directives in place concerning the use of laboratory animals—the three Rs of Replacement, Reduction and Refinement come to mind (Festing and Wilkinson 2007). Important as this may be, this in itself does not take one a long way on the inclusive, anticipatory, reflective and responsive path of RRI. Rather, the variety of governance arrangements hinted at here "must [collectively] aim for [the effective transformation of] present day practices of R&I towards 'responsibilisation', i.e. a process by which the involved actors internalise the issues of concern" (Kuhlman et al. 2016, p. 10).

5.3.5 People

More as a rule than as exception, putting RRI into practice will imply changing both what one does and how one does it. Put in the terminology of organizational management, RRI entices research organizations to amend their *missions and visions*

such that research is no longer a goal in itself but rather a means to accomplish independently identifiable goals best articulated through reference to societal needs and values. And this in turn requires that how R&I trajectories are shaped changes—along the lines sketched above.

As studies of change management (Worren et al. 2016) and sustainability transitions (Voß et al. 2009) have convincingly shown, the types of changes required by such soft-governance approach as RRI is—relying on dispersed actors *taking* responsibility rather than on a framework of rules and regulations *directing* actions—never come cheap. They take time and require cultural, attitudinal and behavioural changes by many on multiple levels—from governmental or non-governmental funding agencies to academic researcher institutions, innovative businesses and industries and civil society actors such as CSOs and citizens.

To group together this plethora of changes, we introduce our fifth and final P: The P for People, as those who travel through and connect all the different levels at which changes are requisite. People, moreover, who best pull of the transition that RRI aspires to contribute to, if they have an open mind and are responsive to change—as described under P number three. And people who, to achieve this, in many cases have to get attuned to new operational logistics, given that for instance including anticipation, reflection and responsiveness in work practices requires not only additional training that allows them to develop new knowledge and skills distinctive of all those process dimensions involved in practicing RRI—which in turn depends on preconditions being met such as time and a commitment by management. To briefly illustrate the latter, we can refer to the multitude of instances in which during our stakeholder consultation workshops we heard people say that soft skills requisite to successful engagement activities, time for undertaking these, and commitment from managerial layers to change (research) processes to become more inclusive were often lacking, both in research organizations, businesses, policy institutions and CSOs.

Arguably, then, the People we refer to are the *obligatory point of passage* (Callon 1984) that simultaneously cannot be avoided when trying to give meaning to RRI and to implement it *and* that remains almost invisible as target of action in itself, as so much of our attention is easily drawn to fleshing out any of the other conventionally referred to elements of RRI. Thus, we direct attention explicitly and specifically to People in a similar spirit as that in which, in the context of discussions revolving around the emerging technology of synthetic biology, *human practices* has become a term of reference (Rabinow and Bennett 2007). For any stakeholder in research and innovation to thoroughly grasp what it takes to make research and innovation more responsible, requires not only that they interact with people from diverse backgrounds and with different (societal) roles to play, but also that they find ways to truly learn from and about each other, their work and their commitments.

5.4 Looking Forward

As part of the *RRI Tools* project we developed a conceptualization of RRI that indicates how the processes of R&I should anticipate on and interact with its foreseen products, enabling the people involved to strive for alignment of the purposes of R&I with the values and needs of society. In the meantime, RRI has steadily continued finding its way into the science policy discourse and attempts are made to implement it in practice. Nevertheless, various ambiguities and differences in interpretation can still be found in the ways experts and stakeholders make sense of RRI principles, actions and results. Is this problematic? Although we recognize the risk of RRI becoming an empty buzzword, we doubt that only more theoretical work will necessarily lead to the desired changes in R&I practices. The meaning and implications of making RRI work should emerge from the interactions between various actors involved and organized around particular issues in specific contexts. In general, we can say that it is important to involve a relevant variety of stakeholders from start to finish in R&I trajectories, but what that means in terms of which stakeholders are engaged and what role they play ultimately depends on the context of application, the timeframe and the perspectives of the actors involved. For instance, in the context of commercial R&D, where issues involving intellectual property rights are at play, the engagement of stakeholders is likely to take different shapes than in the context of applied medical science, which again will be different from basic, curiosity driven science.

In our contribution to the collective attempt at figuring out what RRI can be, we have aimed for a middle road between leaving the criteria that distinguish RRI from R&I open to the context of application and making them specific and clear. Moreover, with our elaboration on RRI's five Ps we assume to carry a message that for a diversity of R&I stakeholders speaks to their motivations and interests and relates to their level of policy influence. And while pulling off a balancing act of presenting a conceptualization of RRI that could arguably be described as partly normative, partly descriptive, partly a critical analysis and partly an instance of public relations, we have also aimed at presenting a narrative that, in different ways, resonates with various audiences.

Thus, we trust that the Purposes of RRI are sufficiently tightly embraced by a sufficiently large number of R&I stakeholders from both commercial and public research institutes as well as R&I policy makers for RRI to really catch on; we assume that the deliverance of true RRI Products will not only contribute to reaching those Purposes, but accordingly will help strengthen RRI's reputation among researchers as valuable R&I enterprise, and work as a binding force that helps connect researchers' interests to those of policy makers, civil society organizations, citizens and society at large; and while recognizing that hurdles are on the way to realizing them, we see empirical evidence accumulating that suggests that the process requirements outlined here can be developed into productive guidelines to co-create RRI practices; furthermore, we urge R&I stakeholders to recognize that meeting the Preconditions for RRI requires a concerted effort on various levels of

R&I governance, and that although this does not come cheap, it will be worth its while; and we cheer for all the People who have so far contributed to realizing RRI, be it either from a policy perspective, the perspective of R&I practitioners, or that of R&I- or RRI-policy researchers, and we invite the latter to further investigate this important aspect of implementing RRI and the policy makers to acknowledge it, and treat it accordingly.

More than anything, however, we stress the importance of continuing the conceptual analysis mainly in connection with practical experiments in RRI. RRI is about a transformation of the research and innovation system. This involves new ways of thinking, doing and organizing research and innovation. Following the seminal work of Argyris and Schon (1974), we believe that researchers, innovators and their organisations learn from experience, gradually adjusting their assumptions and trying out new behaviour. This applies to their learning of RRI as much as it applies to anything. Offering more basic theory will not help them much in acquiring new repertoires for action. How to open up R&I processes to the ideas and concerns of a wider range of involved actors, how to respond adaptively to conversations, controversies, challenges and opportunities that arise, how to anticipate technological futures and reflect on their underlying values and our implicit or explicit concerns: if it is to contribute to the embedment and institutionalization of RRI in various contexts of research and innovation, this should all be acquired through experimenting and reflecting in practice. In line with Wickson and Carew (2014), we encourage researchers, innovators, funders, societal stakeholders and others to engage in analytic-deliberative processes to experiment with existing RRI frameworks like the one put forward here, but also develop their own evaluative criteria and standards to bring about the changes in their practice that they desire. If these experiences will be shared and used to build new experiments, RRI may indeed become the collective experiment in democracy that it can be.

References

Abelson, Julia, Pierre-Gerlier Forest, John Eyles, Patricia Smith, Elisabeth Martin, and Francois-Pierre Gauvin. 2003. Deliberations about deliberative methods: Issues in the design and evaluation of public participation processes. *Social Science and Medicine* 57 (2): 239–251.

Abma, Tineke, and Jacqueline Broerse. 2010. Patient participation as dialogue: setting research agendas. *Health Expectations* 13 (2): 160–173.

Abma, Tineke, and Guy Widdershoven. 2006. Moral deliberation in psychiatric nursing practice. *Nursing Ethics* 13 (5): 546–557.

Argyris, Chris, and Donald Schon. 1974. *Theory in practice: Increasing professional effectiveness*. Oxford: Jossey-Bass.

Barben, Daniel, Erik Fisher, Cynthia Selin, and David Guston. 2008. Anticipatory governance of nanotechnology: Foresight, engagement, and integration. In *The handbook of science and technology studies*, ed. Edward Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, 979–1000. Cambridge, MA: The MIT Press.

Boden, Mark, Christiano Cagnin, Vicente Caribias, Totti Könnölla, and Karel Haegemann. 2010. Facing the future: time for the EU to meet global challenges. Available at: http://ftp.jrc.es/EURdoc/JRC55981.pdf.

Borup, Mads, Nik Brown, Kornelia Konrad, and Harro van Lente. 2006. The sociology of expectations in science and technology. *Technology Analysis & Strategic Management* 18 (3–4): 285–298.

Broerse, Jacqueline E.W., Tjard de Cock Buning, Anneloes Roelofsen, and Joske F.G. Bunders. 2009. Evaluating interactive policy making on biotechnology: The case of the Dutch ministry of health, welfare and sport. *Bulletin of Science, Technology & Society* 29 (6): 447–463.

Callon, M. 1984. Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay. *The Sociological Review* 32 (S1): 196–233.

Callon, Michel, Pierre Lascoumes, and Yannick Barthe. 2009. *Acting in an uncertain world: An essay on technical democracy*. Cambridge, MA: The MIT.

Cavallo, D. 2000. Emergent design and learning environments: Building on indigenous knowledge. *IBM Systems Journal* 39(3.4): 768–781.

Chilvers, Jason. 2012. Reflexive engagement? Actors, learning, and reflexivity in public dialogue on science and technology. *Science Communication* 35 (3): 283–310.

Collingridge, David. 1980. *The social control of technology*. London: Francis Pinter.

Etzkowitz, Henry, and Loet Leydesdorff. 2000. The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. *Research Policy* 29 (2): 109–123.

European Environment Agency. 2002. *Late lessons from early warnings: The precautionary principle 1896–2000*. www.eea.europa.eu/publications/environmental_issue_report_2001_22. Accessed 26 Jan 2017.

———. 2013. *Late lessons from early warnings: Science, precaution, innovation*. www.eea.europa.eu/publications/late-lessons-2. Accessed 26 Jan 2017.

Festing, Simon, and Robin Wilkinson. 2007. The ethics of animal research. *EMBO Reports* 8 (6): 517–610.

Fisher, Erik, Roop Mahajan, and Carl Mitcham. 2006. Midstream modulation of technology: Governance from within. *Bulletin of Science, Technology & Society* 26 (6): 485–496.

Fisher, Erik, Cynthia Selin, and Jameson Wetmore, eds. 2008. *Presenting futures: The yearbook of nanotechnology in society*. Dordrecht: Springer Science.

Gibbons, Michael, Camille Limoges, Helga Nowotny, Simon Schwartzman, Peter Scott, and Martin Tow. 1994. *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: SAGE.

Grin, John, and Armin Grunwald, eds. 2000. *Vision assessment: shaping technology in 21st century society: Towards a repertoire for technology assesment*. Berlin: Springer Verlag.

Grinbaum, Alexei, and Christopher Groves. 2013. What Is "Responsible" about responsible innovation? Understanding the ethical issues. In *Responsible Innovation*, ed. Richard Owen and John Bessant, 119–142. Chichester: Wiley.

Guston, David, and Daniel Sarewitz. 2002. Real-time technology assessment. *Technology in Society* 24 (1–2): 93–109.

Haywood, Benjamin K., and John C. Besley. 2014. Education, outreach, and inclusive engagement: Towards integrated indicators of successful program outcomes in participatory science. *Public Understanding of Science* 23(1): 92–106.

Houghton, John, and Peter, Sheehan. 2006. The economic impact of enhanced access to research findings. vuir.vu.edu.au/472/. Accessed 26 Jan 2017.

Irwin, Alan, Torben Jensen, and Kevin Jones. 2012. The good, the bad and the perfect: Criticizing engagement practice. *Social Studies of Science* 43 (1): 118–135.

Jong, De, Frank Kupper Marije, Marlous Arentshorst, and Jacqueline Broerse. 2016. Responsible reporting: Neuroimaging news in the age of responsible research and innovation. *Science and Engineering Ethics* 22 (4): 1107–1130.

Klaassen, Pim, Frank Kupper, Michelle Rijnen, Sara Vermeulen, and Jacqueline Broerse. 2014. *D1.1: Policy brief. RRI tools project*. Amsterdam: Athena Institute, VU University Amsterdam.

Kuhlman, Stefan, Ralf Lindner, Sally Randles, Bjørn Bested, Guido Gorgoni, Erich Griessler, Allison Loconto, and Niels Mejlgaard. 2016. *Navigating towards shared responsibility*. doc.utwente.nl/102432/1/RES_AGorA_ebook.pdf. Accessed 26 Jan 2017.

———. 2015a. *D1.3: Report on the quality criteria of Good Practice Standards in RRI. RRI tools project*. Amsterdam: Athena Institute, VU University Amsterdam.

Kupper, Frank, Pim Klaassen, Michelle Rijnen, Sara Vermeulen, Remco Woertman, and Jacqueline Broerse. 2015b. *D1.4: A catalogue of good RRI practices. RRI tools project*. Amsterdam: Athena Institute, VU University Amsterdam.

Nowotny, Helga, Peter Scott, and Michael Gibbons. 2001. *Re-thinking science: Knowledge and the public in an age of uncertainty*. Cambridge, MA: Polity.

Oftedal, Gry. 2014. The role of philosophy of science in Responsible Research and Innovation (RRI): The case of nanomedicine. *Life Sciences, Society and Policy* 10 (5).

Owen, Richard, Phil Macnaghten, and Jack Stilgoe. 2012. Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39 (6): 751–760.

Palm, Elin, and Sven Ove Hansson. 2006. The case for ethical technology assessment (eTA). *Technological Forecasting and Social Change* 73(5): 543–558.

Rabinow, P., and G. Bennett. 2007. From bioethics to human practices, or assembling contemporary equipment. In *Tactical biopolitics art, activism, and technoscience*. Cambridge, MA: MIT Press.

Regeer, Barbara, and Joske Bunders. 2009. *Knowledge co-creation: Interaction between science and society*. Den Haag: DeltaHage.

Rip, Arie, Thomas Misa, and Johan Schot. 1995. *Managing technology in society: The approach of constructive technology assessment*. London: Pinter.

Rowe, Gene, and Lynn Frewer. 2004. Evaluating public-participation exercises: A research agenda. *Science, Technology & Human Values* 29 (4): 512–556.

Schomberg, von René. 2007. *From the ethics of technology towards an ethics of knowledge policy & knowledge assessment*. https://ssrn.com/abstract=2436380. Accessed 26 Jan 2017.

———. 2011. Prospects for technology assessment in a framework of responsible research and innovation. In *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden*, ed. M. Dusseldorp and R. Beecroft. VS Verlag: Wiesbaden.

Schot, Johan, and Arie Rip. 1997. The past and future of constructive technology assessment. *Technological Forecasting and Social Change* 268: 251–268.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580.

Swierstra, Tsjalling. 1997. From critique to responsibility; The ethical turn in the technology debate. *Techné: Research in Philosophy and Technology* 3 (1): 45–48.

Van den Hoven, Jeroen (ed.). 2013. *Options for strengthening responsible research and innovation*. http://ec.europa.eu/research/science-society/document_library/pdf_06/options-for-strengthening_en.pdf. Accessed 26 Jan 2017.

Voß, Jan-Peter, Adrian Smith, and John Grin. 2009. Designing long-term policy: Rethinking transition management. *Policy Sciences* 42 (4): 275–302.

Wickson, Fern, and Anna Carew. 2014. Quality criteria and indicators for responsible research & innovation: Learning from transdisciplinarity. *Journal of Responsible Innovation* 1 (3): 254–273.

Wilsdon, J., and R. Willis. 2004. *See-through science: Why public engagement needs to move upstream*, 1–69. London: Demos.

Wilsdon, J., B. Wynne, and J. Stilgoe. 2005. *The public value of science: Or how to ensure that science really matters*. London: Demos.

World Economic Forum. 2016. *The global risks report 2016*. https://www.weforum.org/reports/the-global-risks-report-2016. Accessed 26 Jan 2017.

Worren, Nicolay, Keith Ruddle, and Karl Moore. 2016. From organizational development to change management: The emergence of a new profession. *The Journal of Applied Behavioral Science* 35 (3): 273–286.

Zwart, Hub, Laurens Landeweerd, and Arjan van Rooij. 2014. Adapt or perish? Assessing the recent shift in the European research funding arena from "ELSA" to "RRI.". *Life Sciences, Society and Policy* 10 (1): 1–19.

Chapter 6
Responsible Innovation in Developing Countries: An Enlarged Agenda

Federico Vasen

Abstract The Responsible Research and Innovation framework emerged from the reflection on a socially desirable form of development of emerging technologies in Europe and the United States. In this chapter, I discuss how to further elaborate the framework in order to effectively engage in a dialogue with science, technology and innovation (STI) policy in the developing world, particularly in Latin America. In order to take on this task, I describe first the discussion about uncritical processes of STI policy transfer. Then I analyze the dominant framework of science, technology and innovation policies in the region. Finally, I propose topics that I think should be included in the RRI agenda; themes that will allow the framework to be more responsive to issues related to other geographical contexts. The proposed topics include: (a) expansion of its focus beyond emerging technologies, (b) inclusion of resistance to technologies and contentious politics, (c) global perspective on the production of innovations, (d) building of theoretical links with inclusive innovation frameworks and (d) the development of sensitivity towards intercultural dialogue.

6.1 Introduction

Throughout the 2000s, a moderate paradigm change in the theoretical frameworks that underpin Science, Technology and Innovation (STI) policy was recognizable. The hegemony of National Innovation Systems and evolutionary economics was questioned and new ideas intended to complement them with concepts originated in other fields of the social sciences. New perspectives nuance the centrality of competitiveness and economic growth as the key objectives of STI policies. It is argued that these generic economic objectives do not necessarily lead to the social benefits that "trickle-down" perspectives promise. Accordingly, the new generation of STI

F. Vasen (✉)
Instituto de Investigaciones Sociales, Universidad Nacional Autónoma de México, Mexico City, Mexico
e-mail: federico.vasen@sociales.unam.mx

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_6

policy intends to address social objectives in a more immediate way, although without denying the importance of economic growth or competitiveness. Elsewhere I have called this change a "post-competitive turn" (Vasen 2016a).

The development of the aforementioned new theoretical frameworks express a disenchantment with traditional approaches; those of which have not met the expectations of policy makers and analysts in terms of the social benefits that science and technology was expected to provide for society at large. This is particularly controversial in Latin America and other regions within the developing world, where traditional innovation policies are still hegemonic. Science and technology are still being considered mainly tools to increase greater socioeconomic development and quality of life and the analysis of their risks and negative externalities is usually not considered along with their potential benefits.

Frameworks such as the orientation towards 'grand challenges' (Kallerud et al. 2013), 'responsible research and innovation' (RRI)[1] (Owen et al. 2013) or 'inclusive innovation' (Bortagaray and Gras 2014) can be regarded as part of the 'post-competitive' turn. Out of these three, only inclusive innovation has been taken up and discussed intensively by Latin American scholars. This may not be a surprise, since this perspective focuses on the issues of inequality and poverty that are a main concern in the developing world. In contrast, "responsible innovation" or "grand challenges" are hardly mentioned in the STI policy discourse in Latin America. Do these frameworks address issues that are irrelevant to the local context? Are there ideological preconceptions that prevent a proper discussion on the frameworks*?* In this paper, I analyze specifically the discourse on RRI. I argue that the concerns that motivate the framework are indeed relevant to the Latin American context and are not visible in mainstream discourse. However, the tools that the proponents of RRI suggest should be used to deal with these concerns are expressed in terms that are very alien to local STI governance and political cultures.

In the next section, I discuss the conflicts linked to uncritical processes of STI policy transfer. Then I describe the origin of the multiple perspectives within the RRI framework. I intend to identify what they have in common and I relate these ideas to the discourse on STI policy in Latin America. Finally, I discuss why the main tenet of the RRI discourse is relevant to Latin America, but the tools and cases proposed do not fit with the main local concerns. I propose then five issues that should be included in the agenda of a RRI framework that can effectively make a contribution to local STI policy studies. Moreover, "a view from the periphery" might also be useful to enrich the European perspective on RRI with new perspectives and ideas.

[1] Although the concept advanced by most scholars focuses on innovation, the EU perspective also discusses responsible scientific research. This has been crucial for the current widespread use of the acronym "RRI" instead of "RI". In this paper, although I use the RRI acronym, I concentrate on innovation rather than on scientific research.

6.2 The Problem of Importing Theoretical Frameworks in STI Policy

Regarding the problem of importing ideas from other geographical contexts, the history of STI policy in Latin America has been marked by the uncritical transfer to the region of theoretical frameworks originated in developed countries. This practice has been criticized since the 1960s. At that time, Latin American authors had been highlighting the fact that the level of development of local science and technology systems and the peripheral position of the region in global academic communities required a specific treatment. This standpoint originated as a response to supply-driven science policies based on the linear model that international organizations proposed for the region during the postwar period (Oteiza 1992; Finnemore 1993; Feld 2015). Latin American thinkers noted that the emphasis on capacity building as a prerequisite to progress on the path of technological and social development neglected research agendas linked to issues with greater local social impact (Herrera 1972; Varsavsky 1969; Vasen 2016b).

This line of thought has been embraced by the majority of the local academics involved in science policy studies. Arocena and Sutz (2000) noted that the concept of National Innovation Systems -that had been created as an analytical concept to describe how the most dynamic national configurations on innovation had developed- was proposed in Latin America as a prescription, as a kind of "model" that emerging countries should emulate in order to achieve the long-awaited development. The example of the "Asian tigers", particularly South Korea, was repeated *ad nauseam* as a proof of the viability of this path, without considering the many differences between the Latin American and Asian contexts. According to Dagnino and Thomas (2001) these processes are not restricted to a simple translation of policy frameworks in which a signifier is translated into a different language but attempts to preserve the underlying meaning. What actually is taking place is what the authors call a *transduction* process in which a signifier is inserted into a new context and this creates new meanings (functions, dysfunctions, unwanted effects, etc). This situation, which Dagnino and Thomas described for the case of Latin American STI policy has been described in broader terms by Delvenne and Thoreau (2012) for all non-OECD countries.

6.3 What Is Responsible (Research and) Innovation?

Since the 1980s, a conceptual framework based on evolutionary economics and the idea of innovation began to take shape (Elzinga and Jamison 1995). Its main objective was to strengthen "innovation systems" that group together all the different stakeholders who are part of the processes that lead to the creation of new technologies. It is assumed that the existence of strong and dynamic systems of innovation contribute to central economic objectives, such as growth and competitiveness. The

generation of innovations that yield economic benefits becomes the privileged means through which science and technology improve a countries' economy and the quality of life of its inhabitants (Lundvall 1992; Nelson 1993).

The "responsible innovation" framework is not clearly represented by this scheme. Its central focus is not on economic competitiveness but in the governance of emerging technologies, emphasizing the inclusion and participation of a variety of actors (Owen et al. 2013; Guston 2014). The framework originates from the concept of "responsible development" in the context of the National Nanotechnology Initiative in the United States in the early 2000s. (Roco et al. 2011). This approach is heir to previous research done on the potential unintended consequences of scientific research and technological development. These issues, which initially were addressed by the philosophy of science, bioethics and technology assessment, received a new impetus in the early 1990s with the ELSI component (Ethical, Legal and Social Implications) of the Human Genome Project. In 1994, it was proposed to change the acronym ELSI with ELSA, replacing the idea of ethical, legal and social "implications" with "aspects". The objective was to abandon a linear view regarding the impacts. The change was aimed at including the discussion of particular issues which had potentially more diffuse implications and required a greater degree of public participation. The idea of "aspects" could not be easily restricted to questions of risk, safety and health, which could be framed and analyzed only with expert knowledge (Thoreau 2013; Zwart et al. 2014).

Shortly after the research in the field of ELSI / ELSA consolidated in the 2000s, the concept of Responsible Research and Innovation (RRI) arose. It was proposed as a more comprehensive concept, intended to encompass work on ELSA but also to give it a wider dimension and integrate it with broader areas of science policy. To understand the genealogy of this approach, it is also useful to refer back to the debate on biotechnology, particularly on genetically modified organisms (GMOs). A potentially very critical view regarding emerging technologies emerged in the European context that created discomfort in the industrial elites. Years later, it was feared that nanotechnology could be subject to the same critique. Rip (2006) employed the term "nanophobia-phobia" (a phobia within the scientific and industrial community about the emergence of a nanophobia within the public) to refer to these concerns. Later, the concept of RRI broadened its scope beyond the specific field of nanotechnology and today occupies a central place in the European Union's STI funding programme, *Horizon 2020*, under which is one of the cross-cutting concepts of the component "Science with and for Society", whose total budget for 2014 and 2015 was 91 million euros (European Union 2014a; Galiay 2014).

The responsible innovation framework is best characterized by the multiple elements attached to it rather than by a single definition. According to Zwart et al. (2014), this is a consequence of the top-down manner in which the concept –the signifier "RRI"- was proposed by the policy elite before any meaning was fixed. However, it is possible to distinguish some characteristic features that appear in the writing of most authors who use the term "RRI" which connotes that there is a more or less defined framework. Stilgoe et al. (2013) point out that what characterizes RRI is the intention to create spaces for discussion on aspects of innovations that are

matters of public interest or concern. Thus, with the participation of all affected stakeholders (government, academia, industry, civil society), it is expected that the innovations produced reflect more appropriately the values and interests of a wider set of actors and not only of those who promote the technology. These authors argue that responsible innovation has four integrated dimensions (anticipation, reflexivity, inclusion and responsiveness) that should guide processes of institutional reflection and assessment of scientific and technological initiatives.

Moreover, the European Union, which has been one of the main promoters of the framework has not offered an official definition of Responsible Research and Innovation, but refers to a characterization offered at one Competitiveness Council meeting, in which it was described as "a process for better aligning research and innovation with the values, needs and expectations of society. It implies close cooperation between all stakeholders in various strands comprising: science education, definition of research agendas, access to research results and the application of new knowledge in full compliance with gender and ethics considerations" (European Union 2014b).[2]

Thirdly, the research group led by Jeroen van den Hoven (2013) based at Delft Technical University has also advanced a definition with a somewhat different focus. The vision of these authors arises in the context of the discussion about technology and values. Based on the notion that technology is not neutral, they argue that it is possible to influence the development of emerging technologies in order to incorporate values that are considered desirable. Thus, it is expected that new technologies can solve moral dilemmas by creating technical designs that expand the possibilities of action faced by actors trying to solve ethical problems. This proposal is based on the idea of Value-Sensitive Design, a perspective that emerged at Stanford during the 1970s. Originally dealing with information technology, it was later extended to other emerging technologies. It should be noted that the Dutch research council (NWO) has created a specific funding program for responsible innovation inspired by these ideas. Another feature of this approach is its focus on technology and not science, so it is rather a vision of Responsible Innovation and not of Responsible Research and Innovation.

Critical perspectives on the RRI approach have also emerged recently. They see a pro-industry bias in the concept and argue that by institutionalizing and promoting public participation in a top-down manner RRI proponents could limit the potential for criticism; a possibility previously enabled by participatory mechanisms developed since the 1990s (Thoreau 2013). The upstream total stakeholder participation that RRI proposes would essentially prevent the emergence of a radical opposition and resistance movement and not necessarily lead to a democratic opening. In this sense, it is argued that RRI would be an ally of traditional technological development, allowing its proponents to "stop and think a minute" but without truly putting into question the technological pathways and overall benefits (Rip 2014). Finally, it has also been noted that this framework may increase the bureaucratization of

[2] This definition is based in the six point agenda that appeared in previous EU documents: engagement, gender equality, science education, open access, ethics and governance.

scientific research and innovation through RRI-compliance assessments, which may become required points of passage; although they would be prerequisites they would not necessarily foster reflection of the actors involved.

The starting point of RRI is a much friendlier vision of market dynamics and innovation that more radical previous reflections have linked to public participation in technology assessment. In this sense, responsible innovation is much more reformist than a revolutionary proposal. Its fundamental difference from other more conventional approaches is its claim that the dissemination and social appropriation of emerging technologies is not a phenomenon to be analyzed in purely economic terms (as can be seen in works like Bozeman et al. 2008). The importance it gives to consensus building among all sectors of society is also remarkable. The RRI framework in all its versions has both an analytical ("the adoption of a technology must be understood from a multidimensional perspective") and a normative component ("the participation of more stakeholders in early stages of development should be facilitated") that advocates against economic reductionism and the monopoly of expert knowledge.

6.4 Responsible Innovation and Socio-economic Development in Developing Countries

As previously noted, the agenda of STI policy in developing countries is strongly linked to the paradigm of the economics of innovation and the pursuit of competitiveness. While this is also true in the developed world, in the Latin American scene these frames are virtually hegemonic and do not leave much space for more reflective or alternative paths. The prevalence of a still strongly enlightened vision on the potential of technology (Macnaghten and Guivant 2011) can be thought of as a reason for this.

The central question that should be asked then is, "How useful or relevant is the framework of responsible innovation to the situation of developing countries, particularly in Latin America?" In this sense, I think it is important to distinguish what we might call the "core ideas" of the RRI perspective from the tools and enforcement mechanisms that have been developed in order to put the ideas into practice. As noted by Delvenne et al. (2011), the tools, derived mostly from the European tradition of public participation in science and technology and technology assessment, historically have not been echoed in the Latin American scene. However, this does not mean that the issues that are the root of the RRI approach are unrelated or irrelevant in these regions. Although public participation mechanisms in science and technology did not develop in the European way, numerous examples of critical reaction against technological developments can be identified in Latin America, particularly those related to environmental issues. This critique was conducted in most cases not institutionally but primarily as a phenomenon of resistance and mobilization. This situation indicates that conflicts over the "alignment of research and innovation with the values, needs and expectations of society" also occur in

developing countries. The mechanisms employed however cannot be captured with the tools available in frameworks such as "responsible innovation".

At this point, there are two conceptual alternatives available. The first would involve noting the inadequacy of the concept of RRI to characterize conflicts linked to the social acceptability of technologies in Latin America and criticize the 'Eurocentric' content of the notion, in a post-colonialist fashion. On the other hand, a second option would recognize the importance of establishing a dialogue between the conceptual frameworks developed in Europe and the Latin American agenda, and intend to contribute to the construction of notions that may simultaneously have global reach and address the specific local contexts. This second option presupposes the distinction between a shared ideological core and specific mechanisms and applications that are dependent on the context in which they are applied. I assume then that the concept of RRI involves concerns that are shared by stakeholders globally and are potentially applicable to all regions. However, the way the concept has been deployed in Europe does not match the main topics in the Latin American agenda in STI policy. Even if there are shared problems, the concepts and notions mobilized are mostly foreign to the local context. I do not want to suggest that this is a misconception of the developers of the RRI framework, as it was explicitly constructed for the European context. The main task is then to emphasize which topics are particularly relevant to the discussion of anticipation and social acceptability of emerging technologies in developing countries and to include them in the RRI agenda.

6.5 A 'Responsible' Agenda for Innovation in Developing Countries

One of the main precautions to be taken when discussing STI policies in Latin American is that, unlike in the context of the developed world, both science and technology are addressed primarily as tools to achieve a higher level of socioeconomic development. In this sense, there is a dominant view that STI are a means towards further development. The critical perspective that involves considering science and technology not only as tools for improving the quality of life but as threats and risk carriers (Beck 1992) is highlighted by some social actors but is absent in STI policy discourse. Unlike other approaches towards the risks of emerging technologies, responsible innovation is more optimistic about the possibility of controlling these risks and integrating their assessment into the innovation process (Zwart et al. 2014). However, even if RRI represents a more pro-innovation perspective, the idea of controlling the risks of technology as a whole is only hardly acknowledged by Latin American STI policies as a priority. All attempts to promote policies that engage with the risks and unintended consequences of technology must then challenge what I call the "luxury argument": the idea that all discussion of negative aspects of technology is an intellectual luxury for developing countries. The discussion of risks in an anticipatory and integrated manner, as outlined in the approach of

responsible innovation, could be then referred to as a luxury that cannot be afforded by developing countries. In this view, resources should be invested in the advancement of science and technology and only then, if necessary, the question of risks should be addressed. Discussing risks beforehand could slow down a countries' economic development and be detrimental in the long run. However, this perspective involves ignoring the consequences of not treating the risks of technology with a precautionary approach, as shown by many environmental disasters or forms of intensive exploitation of natural resources.

To confront this "luxury argument", I therefore believe that it is necessary to pose a *general principle* to address the issue of the risk of technological development in developing countries[3] stating that "*the question of social acceptability of new technology should always be analyzed in conjunction with the question of socioeconomic development.*" This is neither a genius idea nor is it a novelty, but keeping this idea in mind can pave the way when raising the issue of the risks of technology and social acceptability to a set of policy-makers who see STI policy only under the lens of innovation and competitiveness, and who characterize all other reflection as superfluous. This is particularly the case when resources are scarce and they subscribe to a trickle-down approach and state that competitiveness-oriented policies will eventually shed satisfactory results for all.

In the next sections, I will present five discussion points linked to the dynamics of conflicts related to the social acceptance of technologies and the "alignment" between research and innovation and societal expectations, needs and values. My intention is to introduce an issue agenda for a concept of responsible innovation more appropriate within the context of developing countries. This does not mean that the framework of RRI is fundamentally wrong. It is just a reminder that in order to meet its goal of aligning technological developments with social expectations in a different geographical context, new aspects have to be highlighted and new tools need to be developed.

6.5.1 Expand the Scope of Responsible Innovation Beyond Emerging Technologies

Although the focus of the RRI framework has been on emerging technologies, the broader problem being addressed is that of social acceptability and the alignment of social expectations with technological development. In this sense, I argue here that a vision of responsible innovation that takes into account that the particularities of the developing world should not be limited to new and disruptive technologies. While it is true that emerging technologies pose new challenges that require the

[3] This principle is naturally not exclusive for developing countries. It just becomes more important in that context since the issue of socioeconomic development has a much more important place on their political agenda.

creation of spaces for reflection, acceptability challenges are still an issue with more mature technologies. "Emerging technologies" usually refers to high-tech developments such as nanotechnology, robotics and biotechnology.

In contrast, in the developing world there are many cases where social acceptability issues related to technology are not necessarily linked with high-tech and emerging technologies. In the case of Latin America, there are conflicts linked to new technologies such as biotechnology as in the case of GM maize in Mexico and Colombia (Fitting 2014) or soy in Argentina (Arancibia 2013) or new technologies regarding the extraction of natural resources such as open pit mining. But among the most contentious cases are also situations related to older technologies such as Belo Monte hydroelectric dams in Brazil (Hall and Branford 2012) or the installation of pulp plants in Argentina and Uruguay (Vara 2007; Baya-Lafitte 2016). All these cases involve questions regarding the social acceptance of technologies and the alignment of innovation with social expectations. Responsible Innovation as a framework should not necessarily be restricted to the cases in which emerging technologies are involved and tools have to be developed to deal with issues of acceptability of mature technologies.

6.5.2 Contentious Politics as Part of the Process of Social Acceptability of Technologies

Literature related to the concept of responsible innovation, ELSI/ELSA or technology assessment, deals mostly with cases of public participation in science and technology within the context of developed countries. The methodologies implemented, such as consensus conferences, citizen juries or scenario workshops, presuppose a context in which dialogue in formal contexts is still possible and stakeholders engage in a rational debate.

There is however a more acute dimension of controversies linked to the acceptance of technologies in which the conflict has escalated to levels at which the application of these methodologies is not possible. In these circumstances, social participation is expressed not institutionally but routed through radical resistance (Pestre 2003; Bauer 1995). In these cases the social actors who oppose a particular technology adopt the disruptive techniques of contentious politics (demonstrations, blockades, boycotts) in order to add more power to their claim.Clearly it cannot be said that these forms of protest are specific to the developing world, as we can find many cases also in the United States, for example, controversies around fracking in the Midwest or the anti-nuclear movement (see Nelkin 1984 for one of the first analyzes on the subject). In the developing world however, the large differences in power between the promoters of technology and those who resist it, added to a diminished presence of the state, distrust in institutions and lower levels of education creates a context in which radical resistance is seen by protest groups as a more viable way to gain public visibility and influence political decisions.

Multiple cases that show this trend can be found in Latin America. In the previous section I mentioned the protest movement against the installation of a pulp-mill plant on the border between Argentina and Uruguay. In that situation, many factors came together: the rejection of potential water, air and visual pollution that the plant would produce, the discussion of whether the Finnish company, *Botnia*, was bringing outdated technology to the region no longer accepted in Europe and the diplomatic conflict generated by its location on an international boundary (Vara 2007). In this regard, many considered this protest as a case of NIMBY conflict because the discussion was mainly about localization and not about the technology itself. The bridge connecting Argentina and Uruguay at that point was closed for almost four years by the protest. The Argentine government took environmentalists' claim as a national cause and even brought the case to the International Court of Justice in The Hague.

Not in all cases such a strong commitment from the government can be found. In the resistance to mega open-pit mining projects in the Andes, the state appears as a partner of the mining companies rather than a supporter of local inhabitants. Local organizations partner with opposition politicians and other international environmental organizations through transnational advocacy networks (Keck and Sikkink 1998). In most cases, protesters have set up roadblocks and triggered repressive action by the authorities (Origlia 2015). In the Patagonian town of Esquel a referendum took place in 2003 where people voiced their opposition against mega-mining projects. In these cases it was the protest and mobilization of the people who generated the withdrawal of projects. It is hardly possible that in a climate of open conflict institutional mechanisms for technology assessment could have been useful to generate a proposal satisfactory for all stakeholders.

In conclusion, I would like to emphasize that if the objective of the responsible innovation framework is to facilitate the social acceptance of technologies, it cannot be restricted to institutional mechanisms that can only work within certain contexts. It is necessary to think how to address conflict situations in which social participation appears only in the most original form of resistance. It may be argued that the framework of responsible innovation should not develop tools to deal with these cases, since what these situations show is what happens when a preventive approach like responsible innovation is not followed through on. The solution would be then to implement the preventive measures recommended in the framework. My position is that, given the characteristics of the governance of STI in the region, the escalation of conflicts and the use of disruptive techniques of social protest is unavoidable in many cases. I am not maintaining the view that the intention of preventing the escalation of conflict through enlarged institutionalized public participation should be abandoned. But since power differences and lack of trust in institutions make resistance inevitable in many cases, it is necessary to find tools that enable responsible action in situations in which radical political action cannot be avoided.

6.5.3 A Global View of Innovation Value-Chains

In today's world, it is increasingly common that the technology development process takes place in multiple locations. While R&D, design and prototyping are often concentrated near the companies' headquarters in developed countries, manufacturing and other polluting and energy-intensive industrial production processes are often carried out –outsourced- in countries in the developing world.

Developing countries are often the source of the raw materials needed for the production of innovations, such as coltan in Sub-Saharan Africa for electronics or lithium in Bolivia and northwestern Argentina for batteries. In addition, these countries may also be recipients of technological waste, i.e. products at the end of their life cycle that are sent to developing countries as donations but are actually unusable products.

When reflecting on what it means to develop an innovation responsibly, it is necessary to include not only the undesirable impacts on its potential users but also the vision of all those who will be affected globally by the process of developing this technology and will participate in the new value chain. *Fairphone*, a mobile phone that comes with minerals obtained outside conflict zones and by paying decent wages to workers is an interesting example that addresses this dimension (Wernink and Strahl 2015).

These situations can be described as negative consequences of the globalization of the economy. To address these conducts, organizations such as the OECD (2011) have issued guidelines urging multinational companies to behave responsibly in the various countries in which they operate. A company's reputation may be adversely affected by these actions, which can damage their own Corporate Social Responsibility strategy. However, neither potential failure to comply with the guidelines nor potential reputational damage can provide a framework in which an irresponsible attitude in this regard can be adequately punished. Civil society can also play an important role making the irresponsible and criminal practices of companies visible.

While this point is not directly linked to the original focus of responsible innovation in emerging technologies, I think it is a central issue for a discussion agenda concerning the responsible development of technologies in a globalized world. The tools for risk analysis and public participation that RRI proposes should include the views of all those who will be involved in the global technological system created by this new technology and the positive and negative consequences they may face. This also includes the populations that will be affected by the globally distributed processes of resource extraction and manufacturing and not only consumers. Internalizing the greater number of externalities, both positive and negative, can be considered a responsible course of action.

6.5.4 Strengthen the Theoretical Link Between Responsible and Inclusive Innovation

In recent times, we have witnessed a growing interest in linking science, technology and innovation policy more directly with the problems of inequality and social inclusion. This trend is verified both by international organizations like the World Bank or the OECD (Paunov 2013; Dahlman et al. 2014) and academic researchers (Cozzens and Sutz 2012; Casas et al. 2014; Thomas 2012; Knorringa et al. 2012). Developing countries are the main beneficiaries of these policies because it is there where the most urgent needs of social inclusion are found.

The approaches currently available have different nuances and not all are necessarily aligned with the ideas of responsible innovation. While in general the need for specific STI policies to address the challenges of inequality and poverty are acknowledged, the approaches taken by international organizations focus on the need to include the "base of the pyramid" as consumers, and regard inclusive innovations as a means for expanding markets to a previously unreachable population (see for example Prahalad and Mashelkar 2010). In contrast, most academic visions regard beneficiaries not only as passive consumers but as citizens and innovators who can transform available technologies to suit their own needs and interests.

This brief comparison between existing frameworks on the subject reveals that it would be a mistake to think that inclusive innovation is necessarily responsible. Frameworks such as the "base of the pyramid" do not address the processes of including actors that is the central tenet of responsible innovation.[4] Moreover, neither the inverse statement is true. Not all responsible innovation is necessarily inclusive in the sense of "inclusive innovation". The literature on responsible innovation is primarily concerned about the inclusion of all stakeholders in the design of emerging technologies. It could be the case of a technology developed along with the participation of all stakeholders but whose objective is not linked to the problem of social inclusion, in the sense that it is understood in the literature on inclusive innovation.

In short, what is important is to create a dialogue between the two theoretical reflections -inclusive and responsible innovation. It could be possible then to move forward in the development of a more integrated framework in order to ensure that all inclusive innovation is also responsible and every responsible innovation is inclusive or at least does not contribute to social segregation.

[4] As Pansera (2014) noted, the main objective of this framework is to include the poor in the market economy. It puts no emphasis on social justice or transformation or other broader dimensions that could be related to responsible innovation.

6.5.5 *The Importance of Establishing Mechanisms for Intercultural Dialogue*

Finally, one last point to which I want to draw attention is related to the origin of the values that the framework of responsible innovation tends to promote.[5] Even if responsible innovation frameworks aim to include all potential actors in the innovation process, the challenge is even greater in situations of intercultural dialogue with indigenous peoples and other traditional knowledge holders. It is true that the problem of cultural acceptance is not specific to only cases of the developing world. Cultural differences regarding the appropriation of technology occur also within Western cultures, as shown for example by Jasanoff (2005) in her study about perceptions of risk in biotechnology. However, the potential incommensurability when Non-Western cultures are involved is even greater. It could even be considered that modern technology, closely related to the project of Baconian-Galilean modern science, is a Western project in itself.

To avoid criticism of responsible innovation as a Eurocentric or colonial framework, it is necessary to incorporate sensitivity within RRI towards the issue of traditional knowledge. The cases of biopiracy in which the intellectual property of indigenous groups was not recognized could be an example to consider and not to follow (Shiva 1997). However, this does not imply that any use of traditional knowledge automatically qualifies as responsible. The 'science-wars' in South Africa regarding AIDS therapy shows that 'traditional knowledge' is a contested and conflictive notion when used to address current issues (Green 2012).

On a broader level, traditional knowledge should not be limited to specific applications but also considered in regard to the different conceptions of human welfare. Bolivia and Ecuador have adopted the idea of *sumak kawsay* or good living from aboriginal people, which conveys an idea of human development very different from the standard in the Western world (Radcliffe 2012). In short, it is important that a concept of responsible innovation incorporates criteria and tools for addressing problems of the social acceptance of technologies in intercultural contexts that conform to notions of justice, diversity and democracy (Olivé 2010). This will give the responsible innovation framework a broader scope and a greater capacity for understanding the dynamics associated with technology in the developing world. Moreover, sensitivity to these issues is also currently strongly needed in the Western world where societies have become more and more multicultural.

[5] It is worth noting that some definitions of RRI include an explicit mention to European values as an intrinsic part of RRI (von Schomberg 2013).

6.6 Conclusion

In this paper, my aim was to discuss responsible innovation within a broader geographical context than usual studies. In this sense, I questioned whether the concept is relevant in order to analyze cases outside the region where it originated and whether it can be useful as a framework for public policy in those settings. I argued that the overall objective that inspires responsible innovation (the alignment research and innovation with the values, expectations and needs of society) is laudable and is not necessarily restricted to the European level. However, the tools that the framework RRI usually discusses are mostly linked to the European context and the tradition of technology assessment and public participation that has not been successfully developed in other regions of the world. In turn, the overwhelming majority of empirical studies linked to the RRI framework are focused on Europe or North America.

In the fifth section, I pointed out five issues that should be considered in order to make the RRI framework more suitable within the context of developing countries, particularly Latin America. The issues mentioned have greater relevance in the context of developing countries. However, this does not mean they are only applicable to them. They can also serve to put more emphasis on issues that may have become blind spots in the European context, and a foreigner's perspective can help to visualize them. Problems of the acceptance of mature technologies, the challenge of intercultural dialogue and the relationship between technology and social inclusion can occur in all geographical contexts.

Finally, this text does not seek to "sell" a concept or impose a buzzword in regions that have not shown much interest in it. My aim is to discuss what Latin American science and innovation policy studies can learn from developments taking place in other regions, as well as what Europeans can learn from an outsider's critique of the concepts they have created. In this sense, and in the case of Latin America, the framework of responsible innovation can be useful to incorporate into local science policy, still hegemonically dominated by the approach of innovation used for competitiveness, a broader view point that emphasizes a more critical approach to technological development and enhances the participation of more stakeholders. Achieving this could be an important step towards the creation of a more multidimensional science and technology policy; one that analyzes the contribution of innovation in terms that exceed economic outputs and that are closer to meeting the needs and expectations of society as a whole.

References

Arancibia, Florencia. 2013. Challenging the bioeconomy: The dynamics of collective action in Argentina. *Technology in Society* 35: 79–92.

Arocena, Rodrigo, and Judith Sutz. 2000. Looking at national systems of innovation from the south. *Industry and Innovation* 7: 55–75.

Bauer, Martin, ed. 1995. *Resistance to new technology. Nuclear power, information technology and biotechnology*. Cambridge: Cambridge University Press.

Baya-Lafitte, Nicolas. 2016. Black-boxing sustainable development. Environmental risk assessment on the river Uruguay. In *Knowing governance: The epistemic construction of political order*, ed. Richard Freeman and Jan-Peter Voß, 237–256. London: Palgrave Macmillan.

Beck, Ulrich. 1992. *Risk society. Towards a new modernity*. London: Sage.

Bortagaray, Isabel, and Natalia Gras. 2014. Science, technology and innovation policies for inclusive development: Shifting trends in South America. In *Science, technology and innovation policies for development*, ed. Gustavo Crespi and Gabriela Dutrénit, 255–285. Berlin: Springer.

Bozeman, Barry, John Hardin, and Albert Link. 2008. Barriers to the diffusion of nanotechnology. *Economics of Innovation and New Technology* 17: 749–761.

Casas, Rosalba, Juan M. Corona, and Roxana Rivera. 2014. Políticas de Ciencia, Tecnología e Innovación en América Latina: entre la competitividad y la inclusión social. In *Perspectivas Latinoamericanas en el Estudios Social de la Ciencia, la Tecnología y el Conocimiento*, ed. Pablo Kreimer, Antonio Arellano Hebe Vessuri, and Léa Velho. México: Siglo XXI.

Cozzens, Susan, Judith Sutz. 2012. *Innovation in informal settings: a research agenda.* Discussion paper for GRIID Network, International Development Research Centre, Ottawa, 2012. http://www.idrc.ca/EN/Lists/Publications/Attachments/1130/IID%20Framework%20July%2029.pdf. Accessed 4 Mar 2016.

Dagnino, Renato, and Hernán Thomas. 2001. Planejamento e políticas públicas de inovação: em direção a um marco de referencia latino-americano. *Planejamento e políticas públicas – ppp* 23: 205–231.

Dahlman, Carl, Esperanza Lasagabaster, Kurt Larsen, and Karthryn Hoffman. 2014. *Inclusive innovation. Harnessing creativity to enhance the economic opportunities and welfare of the poor*. Washington, DC: World Bank Group/Trade and Competitiveness.

Delvenne, Pierre, and François Thoreau. 2012. Beyond the 'charmed circle' of OECD: New directions for studies of national innovation systems. *Minerva* 50: 205–219.

Delvenne, Pierre, Martin Erpicum, Pierre Hupet, and Federico Vasen. 2011. Modernités multiples et critique sociale des technologies en Europe et en Amérique latine. *Cahiers de science politique de l'Université de Liège* 19.

Elzinga, Aant, and Andrew Jamison. 1995. Changing policy agendas in science and technology. In *Handbook of science and technology studies*, ed. Sheila Jasanoff, 572–592. London: Sage.

European Union. 2014a. *Horizon 2020 Work Programme 2014–2015 16. Science with and for Society*. Avaliable at: http://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/main/h2020-wp1415-swfs_en.pdf.

European Union. 2014b. Press release. 3353rd Council meeting Competitiveness (Internal Market, Industry, Research and Space) Brussels, 4 and 5 December 2014, http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/intm/146048.pdf. Accessed 20 July 2017.

Feld, Adriana. 2015. *Ciencia y política(s) en Argentina (1943–1983)*. Bernal: Universidad Nacional de Quilmes.

Finnemore, Martha. 1993. International organizations as teachers of norms. The United Nations Educational, Scientific,and Cutural Organization and Science Policy. *International Organization* 47: 565–597.

Fitting, Elizabeth. 2014. "Cultures of corn" and anti-GM activism in Mexico and Colombia. In *Food activism. Agency, democracy and economy*, ed. Carole Counihan and Valeria Siniscalchi. London: Bloomsbury.

Galiay, Philippe. 2014. *Responsible research and innovation: A cross cutting issue in Horizon 2020*. http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=4146. Accessed 22 July 2017.

Green, Lesley. 2012. Beyond South Africa's 'Indigenous knowledge – Science' wars. *South African Journal of Science* 108: 44–54.

Guston, David. 2014. Understanding 'anticipatory governance'. *Social Studies of Science* 44: 218–242.

Hall, Anthony, and Sue Branford. 2012. Development, dams and Dilma: The sage of Belo Monte. *Critical Sociology* 38: 851–862.

Herrera, Amílcar. 1972. *Ciencia y política en América Latina*. México: Siglo XXI.
Jasanoff, Sheila. 2005. *Designs on nature. Science and democracy in Europe and the United States*. Princeton: Princeton University Press.
Kallerud, Egil, Effie Amanatidou, Paul Upham, Mika Nieminen, Antje Klitkou, Dorothy Sutherland, Maria Lima Toivanen, Juha Oksanen, and Lisa Scordato. 2013. *Dimensions of research and innovation policies to address grand and global challenges*. Oslo: Nordisk institutt for studier av innovasjon, forskning og utdanning.
Keck, Margaret, and Kathryn Sikkink. 1998. *Activists beyond borders: Advocacy networks in international politics*. Ithaca: Cornell University Press.
Knorringa, Peter, Cees van Beers, and André Leliveld. 2012. Frugal innovation in Africa. Tracking Unilever's washing-powder sachets. In *Transforming innovations in Africa: Explorative studies on appropriation in African societies*, 59–78. Leiden: Brill.
Lundvall, Bengt-Ake. 1992. *National systems of innovation: An analytical framework*. London: Pinter.
Macnaghten, Phil, and Julia Guivant. 2011. Converging citizens? Nanotechnology and the political imaginery of public engagement in Brazil and the United Kingdom. *Public Understanding of Science* 20: 207–220.
Nelkin, Dorothy, ed. 1984. *Controversy. Politics of technical decisions*. Thousand Oaks: Sage.
Nelson, Richard, ed. 1993. *National systems of innovation: A comparative analysis*. Oxford: OUP.
Olivé, León. 2010. Conocimientos tradicionales e innovación. In *Observaciones filosóficas sobre la transdisciplinariedad*, ed. Álvaro Peláez and Rodolfo Suárez, 107–129. Barcelona: Anthropos.
Organization for Economic Cooperation and Delopment. 2011. Guidelines for multinational enterprises. 2011 Edition. OECD Publishing. http://dx.doi.org/10.1787/9789264115415-en. Accessed 20 June 2016.
Origlia, Gabriela. 2015. Famatina: la historia de un pueblo en Argentina que en nueve años expulsó a cuatro mineras. In: BBC Mundo, 11 Nov 2015. http://www.bbc.com/mundo/noticias/2015/11/151111_argentina_famatina_expulsion_mineras_bd . Accessed 4 Mar 2016.
Oteiza, Enrique. 1992. El complejo científico y tecnológico argentino en la segunda mitad del siglo XX: la transferencia de modelos institucionales. In *La política de investigación científica y tecnológica argentina - historia y perspectivas.* Buenos Aires: CEAL.
Owen, Richard, John Bessant, and Maggie Heintz, eds. 2013. *Responsible innovation: managing the responsible emergence of science and innovation in society*. Chichester: Wiley.
Pansera, Mario. 2014. *Discourses of innovation and development: Insights from ethnographic case studies in Bangladesh and India.* PhD dissertation, University of Exeter. http://hdl.handle.net/10871/18523. Accessed 27 June 2016.
Paunov, Caroline. 2013. *Innovation and inclusive development: A discussion of the main policy issues*. OECD Science, Technology and Industry Working Papers, 2013/01. Paris: OECD Publishing.
Pestre, Dominique. 2003. *Science, argent et politique. Un essai d'interprétation*. Paris: INRA.
Prahalad, Coimbratore K and Ragunath Mashelkar. 2010. Innovation's Holy Grail. *Harvard Business Review,* July–August 2010.
Radcliffe, Sarah. 2012. Development for a postneoliberal era? *Sumak kawsay,* living well and the limits of decolonization in Ecuador. *Geoforum* 43: 240–249.
Rip, Arie. 2006. Folk theories of nanotechnologists. *Science as Culture* 15: 349–365.
———. 2014. The past and present of RRI. *Life Sciences, Society and Policy* 10: 17.
Roco, Mihail, Barbara Harthorn, David Guston, and Philip Shapira. 2011. Innovative and responsible governance of nanotechnology for societal development. In *Nanotechnology research directions for societal needs in 2020*, ed. Mihail Roco, Chad Mirkin, and Mark Hersam, 561–617. Berlin: Springer.
Shiva, Vandana. 1997. *Biopiracy. The plunder of nature and knowledge*. Cambridge: South End Press.
Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42: 1568–1580.

Thomas, Hernán. 2012. Tecnologías para la inclusión social en América Latina: de las tecnologías apropiadas a los sistemas tecnológicos sociales. Problemas conceptuales y soluciones estratégicas. In *Tecnología, Desarrollo y Democracia*, ed. Hernán Thomas, Guillermo Santos, and Mariano Fressoli. Buenos Aires: MINCyT-UNQ.

Thoreau, François. 2013. *Embarquement immédiat pour les nanotechnologies responsables. Comment poser et re-poser la question de la réflexivité?* PhD dissertation in Social and Political Sciences, University of Liège.

Van den Hoven, Jeroen. 2013. Value sensitive design and responsible innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. R. Owen, J. Bessant, and M. Heintz. Chichester: Wiley.

Vara, Ana María. 2007. Sí a las papeleras, no a la vida. En torno a una controversia ambiental inédita en América Latina. Redes. *Revista de Estudios Sociales de la Ciencia* 25: 15–49.

Varsavsky, Oscar. 1969. *Ciencia, política y cientificismo*. Buenos Aires: CEAL.

Vasen, Federico. 2016a. ¿Estamos ante un giro poscompetitividad en la política de ciencia, tecnología e innovación? *Sociologias (Porto Alegre)* 41: 242–268.

Vasen Federico. 2016b. What does a "National Science" Mean? Science Policy, Politics and Philosophy in Latin America. In: *Science studies during the Cold War and beyond*, ed. Aronova E. and Turchetti S. Palgrave Studies in the History of Science and Technology. New York: Palgrave Macmillan.

Von Schomberg, René. 2013. A vision of responsible research and innovation. In *Responsible innovation. Managing the responsible emergence of science and innovation in society*, ed. Richard Owen, John Bessant, and Maggie Heintz, 51–74. London: Wiley.

Wernink, Tessa, and Carina Strahl. 2015. Fairphone: Sustainability from the inside-out and outside-in. In *Sustainable value chain management. Delivering sustainability through the core business*, ed. Michael D'heur, 123–140. Cham: Springer.

Zwart, Hub, Laurens Landeweerd, and Arjan van Rooij. 2014. Adapt or perish? Assessing the recent shift in the European research funding arena from 'ELSA' to 'RRI'. *Life Sciences Society and Policy* 10.

Part II
Organising RRI: Application, Actors and Approaches

Chapter 7
Climate Engineering: Responsible Innovation or Reckless Folly?

Keynote Address to RRI Conference 2015

Steve Rayner

Abstract Responsible innovation has been a longstanding concern for the social sciences: dating back at least as far as the 1980s, since when a succession of technologies have been introduced with grandiose claims of life-changing benefits, only to founder in the face of under-performance and public scepticism. This paper asks whether the emerging and already controversial field of climate geoengineering will prove to be yet another chapter in this litany or whether it represents an opportunity to develop a framework for responsible innovation according to a model of guiding societal principles and technology-specific protocols. It concludes by noting that geoengineering is currently at a research impasse as technologists await a green light from social scientists before proceeding with research, while social scientists are limited to commenting on highly speculative ideas about how geoengineering might turn out in practice. Under these conditions, the values underlying debates about novel technology are unusually transparent.

7.1 Clashing Values in Innovation

My engagement with issues of responsibility in scientific and technological innovation began very early in my career. My first post-doctoral research project, in the early 1980s, examined the perception of ionising radiation among medical workers in hospitals in the United States. The 1979 nuclear accident at Three Mile Island had focused attention on potential public exposure to ionizing radiation hazards. I wanted to look at the perception of radiation hazards in a context that was perhaps less controversial, a less 'hot' context if you like, than nuclear power. So I chose to look at the use of radiation in medicine, particularly in therapeutic and diagnostic

S. Rayner (✉)
Institute for Science, Innovation and Society, Oxford University,
64 Banbury Road, Oxford OX2 6RP, UK
e-mail: Steve.rayner@insis.ox.ac.uk

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_7

contexts where it could be viewed as actually serving a positive purpose, rather than simply being the incidental hazard of another technology. It was a fascinating project (Rayner 1986). Among other things, I discovered very important cultural differences between people working in research labs and those working in clinical labs. The culture of the clinical lab required technicians to reproduce the same test over and over as exactly as possible. So creating routines and sticking to them very precisely was an important cultural value of the clinical laboratory. On the other hand, in the research lab, the values were very different. The priority here was to make progress on research projects and, unlike the workers in the clinical labs who were very careful to keep their radionuclides within the designated areas on the lab bench and to use only the special hot sinks for the disposal of radioactive materials, the technicians in the research labs said things like 'I treat it like candy'. In fact one of them even went by the amusing sobriquet of 'Three Mile Irwin'.

Following this work, I moved to the Oak Ridge National Laboratory in Tennessee: one of the US Department of Energy's research facilities. Initially they said, 'We are really interested in your work on the perception of ionizing radiation hazards in medicine because it might help us to understand why people are perfectly happy to live downstream of one of the Tennessee Valley Authority's dams but they are often uncomfortable about living close to one of TVA's nuclear power stations. If we can understand this perhaps we can better explain to people why their concerns about nuclear power are misguided'. My response was, 'Firstly, I'm not sure that I can do what you suggest and, secondly, I'm not sure that it would be ethical. I don't think it's the role of the social sciences to engage in social engineering. But I think that we can help engineer technology to the requirements of society'. I went on to propose developing ways of characterising technologies in their early stages to understand what some of their social, environmental, economic, and political implications might be and how those might mesh with the existing values that people have.

Subsequently, as a researcher in the US and Britain, I have engaged with a series of controversial technologies including nuclear power, the Internet, GM foods, nanomaterials, and most recently, climate geoengineering. In the process, I have repeatedly observed what I've come to call the 'novelty trap' (Rayner 2004). Each time one of these new technological fields emerged, their advocates promised marvellous things; nuclear power – 'energy too cheap to meter', the Internet – 'the information super highway', GM foods – 'we're going to feed the hungry', and with nanomaterials we were going to beat swords into ploughshares and produce wonderful kinds of technological utopian futures. What did we get? We got a nuclear waste legacy, the digital divide, the Flavor-saver tomato, and deodorizing socks. Not only was the outcome of these technologies much more mundane that what was promised; the promise of radical new capacities, new human powers through technologies, raised questions about new risks. Time and again, the response on the part of innovators to concern about novel risk was very consistent. They would say, 'Actually, this isn't new after all; we've been doing it a long time'. 'Nuclear energy is just another way of boiling water to drive a turbine to make electricity', or 'digital technology is really an extension of the telephone and the television and existing telecommunications', and 'GM foods are just animal husbandry and crop breeding'.

As for nanomaterials, we were told, 'You can go to the Royal Institution in London and see a big purple jar of colloidal silver created by Michael Faraday in the Nineteenth Century. Nanotechnology is really just colloidal chemistry and electronic miniaturisation'. This cycle of hype, reaction and normalisation is inherently ambiguous. On the one hand, it is a costly and socially disruptive process involving dubious argumentative reversals. But on the other hand, I also suggested that it is a form of informal technology assessment. It is reasonable to ask to what extent the proliferation of public engagement mechanisms for responsible innovation that are proposed by social scientists are driven by discomfort with the untidiness of the political world: a drive to tidy-up politics.

Responsible innovation has been a longstanding concern for the social sciences: dating back at least as far as the British sociologist David Collingridge in the early 1980s. In his classic work, The *Social Control of Technology*, Collingridge (1980) argued that, ideally, society would want to put mechanisms in place for responsible governance early in the development of a new technology, rather than wait until it has become locked in to a larger socio-technological system. When we discover, late in the day, that there are problems with a technology and try to back fill regulation, it is often too late to do so effectively. Since Collingridge, we have seen proposals for *constructive technology assessment* from Ari Rip et al. (1995), the idea of *value sensitive design* (Friedman 1996), *real-time technology assessment* coming from Arizona State University in the United States (Guston and Sarewitz 2002), the idea from the UK of *up-stream public engagement* (Wilsdon and Willis 2004), *anticipatory governance* (Barben et al. 2008) and, more recently, the idea of *responsible innovation* (Owen et al. 2013). So the challenges of governing emerging technology remain serious and ever present concerns that we try to get to grip with in different ways, although we never seem to get quite a satisfactory handle on them.

7.2 Responsible Innovation with Climate Geoengineering

This, then, is the context for the rest of this paper in which I will focus in on my recent experience with climate geoengineering. My engagement with this topic began in 2008 with an invitation from Britain's Royal Society to join a working group preparing a report on climate geoengineering, defined as 'the deliberate large-scale manipulation of the planetary environment to counteract anthropogenic climate change' (Shepherd et al. 2009). Two things are worth emphasising from the outset. One is that there were no fully fledged geoengineering technologies to assess. This remains the case today. Geoengineering proposals are very much in the concept phase. There are existing pieces of equipment that could be incorporated into geoengineering technology systems; there are computer simulations of their effects; but we are not remotely near anything like a socio-technical system capable of delivering geoengineering outcomes. The second thing I want to emphasise is that geoengineering is a broad term covering an extraordinarily wide range of heterogeneous technical practices. In that sense it is very much like the idea of

nanotechnology which, if you remember, brought together diverse fields like microbiology, electronic miniaturisation, molecular engineering and so on and so forth. Geoengineering is a portmanteau term encompassing a very wide range of practices.

7.2.1 Available Geoengineering Technologies

Convention has it that there are two approaches to geoengineering. One is to take carbon out of the atmosphere and store it somewhere. The second is to reflect some of the sun's energy away from the earth by creating some kind of umbrella or parasol effect. But there are also two ways that you can do both of these things. One is by enhancing earth system processes or, if you don't like it, you can say 'tinkering' with earth systems processes. The other is by hard engineering machines. To take carbon dioxide out of the atmosphere by enhancing earth systems processes, one might add iron to the ocean to encourage plankton blooms that would draw carbon out of the atmosphere through photosynthesis, or by engineering artificial trees that extract that carbon from the atmosphere and then inject under pressure into spent oil and gas wells.

Similarly there are two ways to reflect the sun's energy. One is by putting sulphate aerosols into the stratosphere in imitation of volcanic eruptions, which have been associated with atmospheric cooling effects. For example, the 1991 Mount Pinatubo eruption resulted in a drop in the global average temperature of about half a degree for a couple of years. A hard-engineering approach to reflect sunlight might involve putting mirrors into space, possibly at the Lagrange Point between the earth and the sun. This is probably the least talked about geoengineering option, largely because it is considered that the costs of lifting the kinds of payloads involved would be prohibitively expensive.

As James Fleming (2010) has pointed out, engineering weather and climate is not entirely novel. James Pollard Espy was known as 'the Storm King'. He was the scientist who, in the 1840s, advised the US government that it should buy up wood lots across the United States so that it could set fire to them to create convection currents that, he believed, would bring moisture-laden air into an area to create rain. This idea, by the way, was rejected largely because of opposition from the southern states who feared that it might one day be used as a weapon against the South.

The notion that 'the rain follows the plough' was the Nineteenth Century idea that ploughing arid areas, such as the US Great Plains or South Australia, would release moisture from the soil, which would then evaporate to create rain clouds, thus geoengineering the agricultural capacity of the plains by interfering with the weather system. In the Twentieth Century, there were various cloud seeding programmes designed to stimulate rain or divert hurricanes, particularly in the US from the 1940s onwards. In 1965 the first ever report to the US president on climate change (US PSAC 1965) says nothing about reducing greenhouse gas emissions:

nothing about what today we call mitigation. It is also silent about what we would describe as adaptation. It is entirely concerned with things that today we would call geoengineering.

7.2.2 Drivers Behind Geoengineering

So then the question is, why are people reviving geoengineering ideas some 50 years later? Partly it is the result of frustration with the slow pace of climate negotiations about emissions reductions. There is the idea that geoengineering techniques might be used to shave the peaks of a warming trend, to extend the period in which to introduce mitigation options. There is another concern: that we have been making global warming signals stronger by our actions to reduce other forms of air pollution. Recently introduced measures to clean up sulphur emissions from global shipping mean that we are actually putting less sulphate aerosols into the atmosphere than before. Thus, aerosols that have been masking some of the existing warming trend are being removed. Perhaps this could be corrected for by deliberately introducing new aerosols in the stratosphere. Another reason might be to tackle the effects of non-point source emissions. About 60% of the world's emissions are non-point source emissions, this makes them much harder to deal with, while the jet aircraft is probably the only technology for which we cannot currently find a low emissions alternative that will deliver the same level of service that we presently enjoy. Having something that would deal with the emissions from jet aircraft in another way, such as sucking the carbon out of the atmosphere, might be desirable, and that is one of the reasons that Richard Branson has offered a prize for somebody who can come up with a good way to do that. So, some people see commercial opportunities to develop new industries and services in this space.

But we've also had some other kinds of motivations. I have heard people say that they are not really interested in getting people to do geoengineering; they want to use the idea of geoengineering to scare people into trying much harder to reduce their carbon dioxide emissions. On the other hand, there are those who think that the prospect of geoengineering will have the opposite effect of making people feel more relaxed about cleaning up their carbon emissions.

There is also increasing recognition of the impossibility of meeting current temperature stabilisation goals through mitigation alone (e.g., UNEP 2012). The most recent assessment of the Intergovernmental Panel on Climate Change offers only one pathway that keeps a future temperature rise below 2 degrees Celsius. This pathway, RPC2.6, includes very large amounts of Biomass Energy with Carbon Capture and Storage (BECCS). That is a geoengineering carbon removal technology that would have huge implications for land use and may have implications for biodiversity. But the IPCC is essentially saying that we cannot limit warming to 2 degrees target without some kind of geoengineering.

So, there are many motivations for the recent interest in climate geoengineering. But, I do emphasise that it is generally recognised by anyone who seriously engages with geoengineering research that, at best, it is a compliment to, it's not a substitute for, conventional adaptation and mitigation technologies.

7.2.3 Concerns About Geoengineering

There are also many concerns about geoengineering. Is it technically feasible beyond the drawing board and the lab bench? Can you scale it in a timely fashion? Could you get enough BECCS put in place, for example, in time to actually keep emissions down in the way that the IPCC suggests? In the case of carbon removal, we are looking at building an industry on the scale of the fossil fuel industry in order to reverse engineer the waste disposal practices of the industrial economy over more than two centuries.

In addition to the issues of scale and timing, there are significant unknowns costs and financing. Some have claimed that putting sulphates into the stratosphere is cheap (e.g., Barrett 2008). The programme costs of putting sulphate aerosols into the stratosphere are probably fairly cheap, and probably a lot cheaper than conventional mitigation for the same amount of temperature reduction. However, as with all benefit-cost analysis, this conclusion depends on how the analyst bounds the system. If the system is bounded more broadly to include the wider potential environmental and social costs, such as the possible biodiversity implications of Biomass Energy with Carbon Capture and Storage, or the costs of possibly disrupting the Asian monsoon, upon which Asian agriculture is highly dependent, the costs are going to look very different. Both the Royal Society (2009) and the Climate Geoengineering Governance project (MacKerron 2014) reviewed estimates of geoengineering costs and found that that they are entirely sensitive to the input assumptions. If you want to make any geoengineering technology look expensive, then you chose the input assumptions that make it look expensive. If you want to make it look cheap you chose the input assumptions and you bound the system in a way that makes it look cheap. Claims about 'the incredible economics of geoengineering' are just that. They are not credible.

There is also the prospect of unintended consequences, I have already mentioned potential biodiversity implications with BECCS. Disruption of the Asian monsoon has been raised as a possible consequence of the deployment of sulphate aerosols although different computer models reveal conflicting results. There is the question of moral hazard that I've already alluded to. This is the notion that if we appear to have some kind of get out of jail free card, then people will try less hard at reducing their emissions. Concerns have also been raised about a so-called the slippery slope, which is the idea that once you go a little bit down the path of geoengineering you're going to lock yourself in. I will say more about this in a moment. There are potential issues of technical and economic lock-in. An example of technical lock-in might be that if we put sulphate aerosols into the stratosphere and continue to emit greenhouse

gasses, relying on the umbrella that we have created to cool the atmosphere, only to find that there is a significant unintended consequence, such as disruption of the Asian monsoon, that requires us to terminate the programme, then there would likely be a very sudden spike in temperature because of the accumulated greenhouse gases. This could be much more disruptive than having allowed the atmosphere to reach that temperature in a much more gradual way. Economic lock-in is a potential issue with carbon removal. Having constructed a massive industrial infrastructure for carbon removal and storage, there would be entrenched interests in keeping that industry running, even if it has a significant downside.

There are also questions of public acceptability. Environmentalists and policy makers have spent the last few decades telling people not to put sulphur into the atmosphere and not to deposit material in the oceans. Proposals to put sulphate aerosols into the stratosphere to reflect sunlight or iron into the oceans to encourage plankton to draw down carbon dioxide sound very much like arguments that a little deliberate pollution might be a good idea. This might prove to be a hard sell for public acceptability.

There are concerns about the ethics of terraforming. There are those who cleave to a notion of natural ethics who simply say that it is immoral for us to be consciously changing the earth. For those of this conviction, it is bad enough that we have been doing it unconsciously since the industrial revolution.

There are issues around governance and regulation. Who would decide to implement these technologies? How would they be governed? It is clear that the unencapsulated approaches, sometimes called 'feral' technologies, of putting things into the ambient environment, are the technologies that cause the most public concern.

We have something called the geoengineering paradox (Rayner 2010; Zürn and Schäfer 2013) to take into account as well. Which is that the technically easiest and most fast acting technology, which would be sulphate aerosols, would probably be one of the few things that would require a global treaty to implement. I am well known as a longstanding sceptic of the idea that a global treaty is really a precondition for effective climate policies (Rayner and Caine 2015). However, I really cannot envisage putting sulphate aerosols into the stratosphere without one. Imagine for a moment that India had done some stratospheric aerosol spraying just before the 2010 floods in Pakistan. It would have been a hard task to convince a Pakistani politician or person in the street that the Indians weren't responsible for those floods. Regardless of whether or not there is compelling scientific evidence of damage to any country from another's deployment of sulphate aerosols, the potential security implications of people's concerns about disruptions from the technology may make it very difficult to implement such a technology without global agreement, which, as we know from past experience of climate negotiations, may take many years, or even decades, to achieve. On the other hand, atmospheric carbon removal does not seem to require such an agreement. We could probably build carbon removal machines under existing environmental and planning laws of countries, provided we were sequestering the carbon within national boundaries (Armeni and Redgwell 2015a, b). The delay in effective deployment here arises from the scale of the industry

that would have to be in place before it would make much of a dent in the atmospheric concentrations of carbon. There would then be a further delay between taking that carbon out of the atmosphere and it having any material influence on the climate. So the technically fastest option with the highest leverage would likely be delayed for political and social reasons while the institutionally easiest solution would also take a very long time to take effect. In both cases it seems unlikely that geoengineering will be making much, if any impact on global temperatures by mid-century.

7.3 Governance Between Uncertainty and Control

There has been outspoken opposition to any idea of geoengineering. David Suzuki has described it as 'insane'. In a BBC television interview, David Attenborough described it as 'fascist'. The ETC group (2010), an NGO with a history of implacable opposition to a number of new technologies, has described it as 'geopiracy'. There is certainly reason to be concerned about the potential for technocracy associated with geoengineering. Paul Crutzen who re-launched the idea of geoengineering in a 2006 article in Climatic Change was also one of the advocates of the idea of the Anthropocene, writing in a 2002 article, 'and I do note that the daunting task that lies ahead is for scientists and engineers to guide society towards environmentally sustainable management during the era of the Anthropocene'. However, there is an equally worrying technocratic tendency in mainstream climate discourse in which some commentators (e.g. Beeson 2010) have argued for consumption controls enforced by a "good authoritarian" or that: "We need an authoritarian form of government in order to implement the scientific consensus on greenhouse gas emissions" (Shearman and Smith 2007:4). The technocratic tendency, not to say threat to democracy, is present in the wider climate discourse and is not in any way peculiar to geoengineering.

The issue of governance brings me back to the conclusion of the Royal Society report that the acceptability of geoengineering would be determined as much by social, legal and political factors as by scientific and technical factors (Shepherd et al. 2009). That was a very radical finding for the Royal Society. This report was almost not issued because all reports of the Royal Society have to be approved by its council; there were many members of the council that felt that this was going beyond the Royal Society's remit instead of sticking to just doing the narrow, technical, scientific assessment. The report recommended the development and implementation of governance frameworks to guide both the research and development (R&D) and the possible deployment of the technologies and it is to those governance frameworks for R&D that I'm now going to turn.

7.3.1 Slippery Slope Concern

The controversy over geoengineering extends even to the idea of doing research to see if these sociotechnical imaginaries can be developed into practical technologies. The question is particularly posed in relation to research into solar radiation management methods, particularly of sulphate aerosols. For example, the geographer Mike Hulme, author of *Why we disagree about climate change*, argues in his new book on geoengineering that we simply cannot know in principle enough about the way these technologies will behave in practice to responsibly implement them; since we know that now, we should not even begin research. Indeed, our ignorance about the implications of these technologies will save us from the folly of trying to develop them. He says that "the simulation models upon which aerosol injection technology would rely are like calculative cartoons" (Hulme 2014). Harvard University engineer David Keith disagrees. He has laid out a programme whereby, he suggests, that our ignorance can be reduced by carefully conducted research and that we actually have quite a lot of information based on empirical experience with volcanic eruptions and computer modelling (Keith 2013). He argues that we should not remain ignorant of our ignorance.

There is concern in some quarters that to do research is to embark on the slippery slope that I mentioned earlier. The concern here is that even starting to do research on these topics normalises them. In the recently completed three-year project on Climate Geoengineering Governance funded by the UK Economic and Social Research Council (Healey and Rayner 1995), we found ourselves wondering whether, by researching the potential governance issues, we were making thinkable something that some people think ought not to be thinkable? It's a reasonable question. However it is a very strong claim to say, as some people have done, that once the first step has been taken then development cannot be stopped. We know that patent offices are actually the graveyards of dreams. Most patents are death certificates. We spent 30 years and close to a 100 billion dollars on research into fast breeder reactors, but there are no fast breeder reactors operating anywhere today. So there is some empirical evidence that suggests that the slippery slope argument is, perhaps, overstated.

7.3.2 Dilemmas of Control for Research

There are a number of important dilemmas of control for research. I have already said we are talking about socio-technical imaginaries. There are no geoengineering technologies existing today. What then is the object of governance, regulation, control or management for geoengineering? There has been a fascinating definitional politics around deciding which technologies fall into or outside the category of geoengineering. For example, at the 2011 Asilomar Conference in California, people involved in forestry were very keen not to be classified as doing geoengineering

because they were concerned that it might subject them to additional regulatory burdens. On the other hand, biochar advocates were desperate to get that technology included in the category of geoengineering because they thought that this would be an opportunity to open up research and development funding for their particular technology. Such definitional arguments continue with the latest reports from the US National Academy of Sciences on geoengineering, which split the topic into two separate reports, one on solar radiation management and one on carbon dioxide removal, arguing that they have nothing in common, and that carbon dioxide removal is really just mitigation US NRC (2015a, b). I am dubious about this move, not least because it isolates technical interventions based on their mechanism rather than viewing technologies as complex systems that might combine multiple processes. For example, if we were to contemplate solar radiation management using sulphate aerosols, given the termination problem that I mentioned earlier (the danger of a sudden temperature spike) it would be prudent to have an exit strategy from that technology. One exit strategy that suggests itself would, of course, be atmospheric carbon removal. So if we think about a socio-technical *system* of geoengineering it might well involve doing some solar radiation management while simultaneously developing a capacity to take carbon dioxide out of the atmosphere as a way that to bring that phase of climate intervention to a close. So the notion that carbon removal and solar radiation management are completely separable is, I think, a questionable one. Either way, it is a very important to be aware that these are deeply political and not just technical considerations.

There is also the problem of how to distinguish geoengineering from basic science. If someone sprays some sulphate aerosols into the atmosphere claiming that they are doing this just to study how clouds form, then they may be doing exactly the same experiment as somebody who says that they are doing it in pursuit of developing a geoengineering technique. Does it make a difference how that experiment is defined, or is it the activity of experiments in the atmosphere that we are really wishing to govern?

Who should govern and regulate research? Should it be self-regulation? At a 2014 geoengineering conference in Berlin, it was very clear that some scientists and engineers held strong views that it was their prerogative to manage their own technological activity. They were offended at the idea that anybody should think they would not exercise appropriate responsibility in developing their research projects. On the other hand, there were other participants who clearly expressed strong opposition to anything that looked like self-regulation and did not want any geoengineering to happen until there was a global treaty in place, which would put it off a very long time.

Finally, there is the desirability of avoiding locking-in, either to an inappropriate or unacceptable technology, or to a governance architecture that is inadequate to deal with that technology as it develops or possibly prevents a technology from developing that otherwise would be socially useful. We don't have a completely blank slate. National governments already regulate activity on the ground within their borders. Experimental work for carbon dioxide removal would, in any case, have to conform to existing regulations governing environmental impact and

planning (Armeni and Redgwell 2015b). There are customary international norms imposing duties of care on states not to damage neighbours. There is, as yet, no international law of the atmosphere. The London Convention and Protocols govern experiments in international waters but there is still a vacuum for comparable geoengineering activities in the atmosphere. We have no law of the air although the Convention on Biodiversity (CBD) has expressed some interest and appetite to assert wide jurisdiction. The problems with relying on the CBD include the facts that (a) the US is not a signatory and (b) it can only exert that jurisdiction if there are impacts on biodiversity, not for other kinds of impacts.

7.4 Oxford Principles

So, do we need a fully articulated governance architecture before research can go forward? Is such an architecture even plausible? What features would it have to have? It would have to recognise this tension between top-down anticipatory and bottom-up emergence approaches to governance. Would it have to be implemented bottom-up through existing national structures for R&D funding and regulation, or top-down as part of a global comprehensive architecture. In 2010 a group of social scientists with past experience of responsible-innovation dilemmas set out to propose a few guiding principles that would be applicable to a highly heterogeneous range of technologies (Rayner et al. 2013). In articulating these principles, we did not start with the imagined attributes of the range of geoengineering imaginaries. Rather, the principles were based on our long experience with past socio-technical and environmental controversies, including nuclear power, GM foods, biodiversity, ocean dumping and nanotechnology. We focused on what we saw as the kinds of concerns that cropped up time and time again with novel technologies. Which of these should inform non-negotiable principles for responsible geoengineering R&D?

7.4.1 Five Principles

We came up with five principles, which have become known as the Oxford Principles, which we submitted to the UK House of Commons Science and Technology Committees enquiry into climate geoengineering governance.

Principle 1: Geoengineering to be regulated as a public good acknowledges that all individuals have a common interest in the good of a stable climate and, therefore, the means by which this is achieved. At the global level, the principle asserts that the state of the climate is a common concern of humankind. Hence, geoengineering research should be regulated as a public good and in the public interest at both national and international levels.

Principle 2: Public participation in geoengineering decision-making requires that those conducting geoengineering research must notify, consult and, ideally, obtain the consent of, those affected by the research activities. The extent and mode of public engagement could be different depending on the level of activity being proposed. It would be perfectly reasonable, for example, in conducting an experiment like the SPICE balloon experiment, which I will come to in a moment, to engage the people living adjacent to the Cambridgeshire airfield from which the balloon would be launched and tethered. I don't think that it is necessary to ask people in the People's Republic of China for their permission to do that. On the other hand, as the scale of potential impacts increases it would be necessary to find ways of engaging broader publics. The mode of engagement will also vary, depending on political culture. Public engagement is likely to be driven by very different logics in urban America and rural China (Wong 2013).

Principle 3: Disclosure of geoengineering research and open publication of results requires the prompt and complete disclosure of research plans and open publication of results to allow the public to assure itself as to the integrity of the process. Mindful of cases when drug companies withheld negative results of product trials in seeking licences, it is an essential feature of this principle that the results of all research, including negative results, be made publicly available. The requirement for complete disclosure and open publication of results is an appeal to the procedural value of transparency, in which the basis of decisions as well as rationales of the decision-makers must be made available.

Principle 4: Independent assessment of impacts should be conducted by a body that is intellectually and financially separate from those undertaking the research. Where techniques are likely to have transboundary impacts, assessment should be carried out through the appropriate regional and/or international bodies. Assessments should address both the environmental and socio-economic impacts of research, including mitigating the risks of lock-in to particular technologies or vested interests.

Principle 5: Governance before deployment requires that decisions with respect to deployment should only be taken with robust governance structures already in place, using existing rules and institutions wherever possible. Governance structures must have the credible capacity to enforce rules and terminate activity in the event of physically or socially deleterious effects.

Subsequently we have elaborated on Principle 5, to argue that the governance arrangements for each stage of research and development also needs to be clear before you go on to the next stage of research and development. This elaboration incorporates the idea of research protocols and stage-gate reviews.

7.4.2 Research Protocols

The key to implementation of the Oxford Principles will be in the development of research protocols for each stage of the development of the technology, from the initial idea, through computer simulation, laboratory experiments, outdoor experiments, field trials, to eventual implementation. We propose that before any activity is commenced, at each stage of development, that researchers be required to prepare a research protocol that explicitly articulates how the issues embodied in each of the Oxford Principles is to be addressed at that stage of research (Healey and Rayner 2015). Such stage gates would enable research protocols to be interrogated by a competent third party appropriate to the stage of research. While university ethics committees might be able to provide sufficient review for computer modelling, outdoor experiments might require a higher level of review which could be provided by public funding bodies, where they are sponsoring such research, or by independent review panels appointed for the purpose by national academies of science. Where an experiment has the potential for transborder impact, clearly, the review should include representatives from all potentially affected countries. The review body at each stage gate must be invested with the authority to withhold approval until it is assured that the experimental design for that stage satisfies the Oxford Principles and that it will be competently and conscientiously implemented. Where there is a risk to third parties, the review body could use the stage-gate process to require specific risk-mitigation measures. There is also the possibility that it could establish satisfactory liability arrangements in anticipation of potential damage. Most importantly, each stage gate would enable researchers and regulators to address specific issues of reversibility.

Such an approach would allow the governance structure to co-evolve with the technology step-by-step, rather than trying to anticipate at the outset what the technology will look like when it has undergone extensive development.

The research protocols and stage-gate approach also addresses a criticism that was made of the Oxford Principles in the magazine *Nature* (2012), which argued that the principles were so general that they could apply to any emerging technology. While the research protocols provide the technological specificity that *Nature* was looking for, I would welcome the widespread adoption of the Oxford Principles, or something like them, as Principles of Responsible Innovation to be applied in other areas of novel and controversial technological development.

The cancellation of the SPICE balloon experiment provides some evidence that the principles + protocols + stage gates approach has practical potential. Stratospheric Particle for Climate Engineering (SPICE) was a UK research programme to assess sulphate aerosol technology. As part of that programme, it was proposed to launch a balloon on a 1 km tether over a Cambridgeshire airfield to do some engineering tests to investigate how such an arrangement might behave if deployed on a much larger scale: a 20 km high balloon the size of Wembley Stadium that could be used as a way of continuously introducing sulphate aerosols into the stratosphere, rather than taking them up in aircraft. This proved controversial. The UK Research

Councils that were funding the research, therefore, required the balloon project to go through a series of stage-gate approvals prior to releasing the funds for the experiment. In the first stage-gate meeting the proposers were told that they had not consulted adequately with the public about the technology. The project team would have to go back and do more public consultation. However, before that second stage of consultation was completed, it came to light that two of the scientists involved in proposing this experiment had already taken out patents; they had already obtained intellectual property on the technology. At this point, Professor Matt Watson, the head of the SPICE project, decided to cancel the balloon project. In doing so he explicitly invoked Oxford Principle 1, because he said that this private acquisition of property rights was not in the public interest and public good. He also recognised the inadequacy of the public engagement that had been performed, and the absence of explicit governance arrangements for outdoor experiments, issues addressed by Oxford Principles 2 and 5 respectively (Watson 2012). Many people refer to these events as 'the SPICE balloon debacle' or 'the SPICE balloon disaster'. I refer to it as 'the SPICE balloon success' because I think it showed how this model of governance can be implemented successfully in practice. It doesn't guarantee that, of course, but it shows that it can.

So, I would argue that the Oxford Principles have proven to be robust and useful and they have been used, the stage gate model was effective, but norm building is essential. Colleagues at the Institute for Advanced Sustainability Studies in Postdam, Hubert and Reinweich (2015), have taken up the challenge of norm building by proposing a draft code of conduct for responsible geoengineering research, drawing on the Food and Agriculture Organisation's Code of Conduct for Responsible Fisheries to develop the idea of the Oxford Principles and the technology protocols even further.

7.5 Concluding Remarks

In closing, I just want to raise a few questions. Firstly, is geoengineering the right object for our attention? I have already alluded to this. I have talked about the active definitional politics of opting in and opting out. Do we really want to say that carbon dioxide removal is just mitigation, particularly in the light of the potential biodiversity impacts of a technology like BECCS? Does intention matter? Is the real concern the potential detrimental effects of the activity or is it with the motivation behind the activity? Is geoengineering, per se, the proper focus for governance or should the focus really be on any kinds of experiments in the atmosphere on some kind of scale or another? I don't know the answers to all of these questions, but I do believe that they are ones that society needs to address if it is going to consider any sort of geoengineering in a spirit of responsible innovation.

Furthermore, we should not limit our thinking just to the implications of geoengineering for society in the narrow sense of the technical or climate implications of geoengineering. We should also ask, 'What can we learn from thinking about geoengineering for responsible innovation?'

Geoengineering challenges standard narratives of enthusiastic technological innovators being stopped in their tracks by alarmed NGOs. In the geoengineering case the potential technology developers themselves took up the social scientists' call for upstream deliberation about the ethics and social acceptability of their imaginaries. Whatever the propaganda from ETC or those fringe groups who view geoengineering as part of a global conspiracy to poison us all with so-called chemtrails (Cairns 2014), this is not Dr. Strangelove. It is rather the scientific community saying, 'Ok you social scientists, you've been telling us for years we should be doing this upfront assessment, come in and do it'. It is notable that there have been more social science papers on geoengineering published in the last few years than there have been geoengineering technical papers.

But the ratio of social science to technical papers also highlights a major weakness inherent in the ideas of up-stream public engagement and up-stream technology assessment. Social scientists are in a situation where they have a choice of imaginaries, ranging from geoengineering as something that could be very useful, perhaps even inevitable, to one that sees geoengineering as extremely dangerous, even immoral. Exactly what are social scientists to engage the public about? We have tried to get publics involved but, for the most part, they are not interested in engaging because they see geoengineering to be little more than science fiction or too vague and poorly defined to get to grips with.

In the absence of a clear social license to operate emerging from public engagement and social science research, the technology developers and research funders seem reluctant to move geoengineering imaginaries away from science fiction towards concrete technical configurations that the public and social scientists can get to grips with. The prudent concern over governance that is holding technologists back from doing anything outside of the lab, also limits social science research because, as a consequence, we lack specificity about the technology that we are trying to assess. Is it Mike Hulme's imaginary or David Keith's, some hybrid of the two, or a third vision entirely?

Thus geoengineering research seems to be at a research impasse. The scope for the social sciences is highly limited until the technology develops into something less speculative than it is today. On the other hand, the technology seems unlikely to develop until it gets clearer guidance from the social scientists regarding the likely operating context and governance arrangements for the technology.

So, what can this geoengineering discourse do for society? I think it's fascinating that the values in these debates at this moment are unusually explicit. There is very little science to hide behind and that's unusual. Usually controversy only emerges once there is enough science to enable people to hide their social values behind technical claims. This is a very interesting moment. The geoengineering discourse presents us with an opportunity to explore representations of nature, debates about the good society, the role of technology in our lives, and so on. Regardless of whether it actually ever comes into fruition in material terms, geoengineering and the debate about geoengineering governance can teach us much about the governance of other global emerging technologies.

References

Armeni, Chiara, and Catherine Redgwell. 2015a. *International legal and regulatory issues of climate geoengineering governance: Rethinking the approach*. Oxford: CGG Working Paper no. 21.

———. 2015b. *Geoengineering under national law: A case study of the United Kingdom*. Oxford: CGG Working Paper no. 23.

Barben, Daniel, Erik Fisher, Cynthia Selin, and David H. Guston. 2008. 38 Anticipatory Governance of Nanotechnoiogy: Foresight, Engagement, and Integration. In *The handbook of science and technology studies*, ed. Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman. Cambridge, MA: The MIT Press.

Barrett, Scott. 2008. The incredible economics of geoengineering. *Environmental and Resource Economics* 39 (1): 45–54.

Beeson, Mark. 2010. The coming of environmental authoritarianism. *Environmental Politics* 19 (2): 276–294.

Cairns, Rose. 2014. Climates of suspicion:'chemtrail'conspiracy narratives and the international politics of geoengineering. *The Geographical Journal*: 1–15.

Collingridge, David. 1980. *The social control of technology*. London: Pinter.

Crutzen, Paul J. 2002. Geology of mankind. *Nature* 415 (6867): 23–23.

———. 2006. Albedo enhancement by stratospheric sulfur injections: A contribution to resolve a policy dilemma? *Climatic Change* 77 (3): 211–220.

Fleming, James Roger. 2010. *Fixing the sky: The checkered history of weather and climate control*. New York: Columbia University Press.

Friedman, Batya. 1996. Value-sensitive design. *Interactions* 3 (6): 16–23.

Guston, David H., and Daniel Sarewitz. 2002. Real-time technology assessment. *Technology in Society* 24 (1): 93–109.

Healey, Peter, and Steve Rayner. 1995. *Key findings from the climate geoengineering governance project*. Oxford: CGG Working Paper no. 25.

Hubert, Anna-Maria, and David Reinweich. 2015. *An exploration of a code of conduct for responsible scientific research involving geoengineering*. Potsdam: IASS and Oxford: InSIS.

Hulme, Mike. 2014. *Can science fix climate change?: A case against climate engineering*. Chichester: Wiley.

Keith, David. 2013. *A case for climate engineering*. Cambridge, MA: MIT Press.

MacKerron, Gordon. 2014. *Costs and economics of geoengineering*. Oxford: CGG Working Paper no. 13.

Nature. 2012. Editorial: A charter for geoengineering. *Nature* 485: 415.

Owen, Richard, John Bessant, and Maggy Heintz, eds. 2013. *Responsible Innovation: managing the responsible emergence of science and innovation in society*. Chichester: Wiley.

Rayner, Steve. 1986. Management of radiation hazards in hospitals: Plural rationalities in a single institution. *Social Studies of Science* 16: 573–591.

———. 2004. The novelty trap: Why does institutional learning about new technologies seem so difficult? *Industry and Higher Education* 18: 349–355.

———. 2010. 'The Geoengineering Paradox'. *The Geoengineering Quarterly*. Online newsletter, 20 March. http://www.greenpeace.to/publications/The_Geoengineering_Quarterly-First_Edition-20_March_2010.pdf. Accessed 7 Feb 2017.

Rayner, Steve, and Mark Caine, eds. 2015. *The hartwell approach to climate policy*. Abingdon: Routledge.

Rayner, Steve, Clare Heyward, Tim Kruger, Nick Pidgeon, Catherine Redgwell, and Julian Savulescu. 2013. The Oxford principles. *Climatic Change* 121 (3): 499–512.

Rip, Arie, Johan Schot, and Thomas J. Misa. 1995. Constructive technology assessment: A new paradigm for managing technology in society. In *Managing technology in society. The approach of constructive technology assessment*, 1–12.

Shearman, David, and Joseph Wayne Smith. 2007. *The climate change challenge and the failure of democracy*. Westport: Greenwood Publishing Group.

Shepherd, John, Ken Caldeira, Joanna Haigh, David Keith, Brian Launder, Georgina Mace, Gordon MacKerron, John Pyle, Steve Rayner, Catherine Redgwell, and Andew Watson. 2009. *Geoengineering the climate: Science. governance and uncertainty*. The Royal Society: London.

UNEP. 2012. *The Emissions Gap Report 2012*. United Nations Environment Programme (UNEP), Nairobi: 3.

US NRC. 2015a. *Climate intervention: Reflecting sunlight to cool the earth. Report of the national research council*. Washington, DC: National Academies Press.

———. 2015b. *Climate intervention: Carbon dioxide removal and reliable sequestration. Report of the national research council*. Washington, DC: National Academies Press.

US PSAC. 1965. *Restoring the health of our environment: Report of the environmental pollution panel. President's science advisory committee*. Washington, DC: The White House.

Watson, M. 2012. *Expert reaction to decision not to launch the 1 km balloon as part of the SPICE geoengineering research project*. Science Media Centre. http://www.sciencemediacentre.org/expert-reaction-to-decision-not-to-launch-the-1km-balloon-as-part-of-the-spice-geoengineering-research-project-2/. Accessed 7 Feb 2017.

Wilsdon, James, and Rebecca Willis. 2004. *See-through science: Why public engagement needs to move upstream*. London: Demos.

Wong, Pak-Hang. 2013. The public and geoengineering decision-making. *Techné: Research in Philosophy and Technology* 17 (3): 350–367.

Zürn, Michael, and Stefan Schäfer. 2013. The paradox of climate engineering. *Global Policy* 4 (3): 266–277.

Chapter 8
Formal and Informal Assessment of Energy Technologies

Udo Pesch, Aad Correljé, Eefje Cuppen, Behnam Taebi, and Elisabeth van de Grift

Abstract Societal controversies on the implementation of new energy technologies relate to public values that are affected by these new technologies. The process of specifying and articulating these values and assessing technologies based on those values follows both a formal and an informal trajectory. This chapter studies the interplay between such formal and informal assessment in the case of two plans for the implementation of energy technologies in the Netherlands, the project to develop carbon capture and storage in the municipality of Barendrecht and the project to have explorative drilling for shale gas in the municipality of Boxtel. What these studies reveal is that values that are specified by actors emerge from different discourses. These discourses do not emerge independently, but develop in mutual interaction, at times contributing to a progressive entrenchment of positions. To avoid such a situation, the symmetry between different claims for the public interest needs to be attended by policy makers.

8.1 Introduction

The implementation of new energy technologies frequently leads to societal controversy. Think for instance of wind parks (Breukers and Wolsink 2007; Devine-Wright 2005), nuclear power (Pidgeon et al. 2008; Taebi and Roeser 2015), carbon capture and storage (Cuppen et al. 2015), hydrogen fuels (Huijts et al. 2014; Di Ruggero 2014) as cases that have raised intense public debates. These controversies are fuelled, firstly, by societal changes. Society has become 'fluid'(cf. Bauman 2000), in the sense that people have become more individualised, more mobile and more articulate than before, which makes it hard to pre-empt on societal preferences and the legitimacy of collective decisions. Societal opinion has become more intangible than ever before because of higher general levels of education, emancipation and

U. Pesch (✉) • A. Correljé • E. Cuppen • B. Taebi • E. van de Grift
Department of Values and Innovations, Delft University of Technology, Delft, The Netherlands
e-mail: U.Pesch@tudelft.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_8

social mobility. In addition, new media technology allows effective forms of communication and collective action. Secondly, the energy system has become entangled in a so-called transition process. Whereas not that long ago the energy system could be typified as a one-directional grid that was centrally controlled, nowadays, we see a multidirectional grid with no fixed roles for producers, consumers and regulators (Pesch 2014). Moreover, questions about environmental degradation and resource depletion have made the future of the energy system highly unpredictable (Ligtvoet et al. 2015; Ligtvoet et al. 2016). Although all actors in the system are aware that the energy system cannot be maintained as it is, no one knows which new system it will evolve into. Moreover, ideas and preferences on how, how fast and by what means the system should transform diverge considerably.

Debates on energy technologies involve the articulation of public values related to the new energy technology that is to be implemented. The process of specifying and articulating values and assessing technologies based on those values follows two trajectories. There is a *formal* path of assessment in which a repertoire of procedures, standards, tools, and policy arrangements is used for establishing a collective appraisal of the new technology (Taebi et al. 2016). At the same time, a debate includes an *informal* path of value articulation. Controversy can be regarded as informal assessment (Rip 1986), as it articulates the conflicting values at stake and reveals unanticipated societal and ethical risks. This informal trajectory is characterised by the advocacy for public values that some groups consider to be underrepresented in formal forms of assessment.

Responsible innovation can be defined as the need to the include the full range of relevant values in the design of both a technology and the institutional context in which the technology is embedded (Taebi et al. 2014; Owen et al. 2013). Given the volatile character of society and the energy system, the pursuit of responsible innovation becomes a challenge in the case of energy technologies for three reasons. First, the volatile character makes it harder to *identify* the values that are connected to a technology. Second, it is difficult to *include* values that emerge in informal assessment in decision-making processes, as these 'emerge' outside of the scope of conventional circuits of policy making. And third, new institutional and technological characterisations of the energy system may make decision-making structures *outdated*. These complications urge us to think about the questions as to how the interplay of formal and informal assessment of new technologies influence technology implementation and how decision making structures can include the range of values, not only those assessed in formal procedures but also those that emerge in informal assessment. Here we will study the interplay between formal and informal assessment in the case of two plans for the implementation of energy technologies in the Netherlands, namely the project to develop carbon capture and storage in the municipality of Barendrecht and the project to have explorative drilling for shale gas in the municipality of Boxtel.[1]

[1] These cases have been studied before in research projects in which (some of) the authors have been involved. The Carbon Capture and Storage case has been studied as part of the 'The Next 50 Years' project, financed by a grant of the Energy Delta Gas Research (EDGaR) program.

These cases have been chosen for their significant impact on the Dutch debate on energy policy, especially because their contested nature led to cancellation of these projects. As such they are emblematic for the role that societal controversy can play in current discussions on new energy technologies. For discussing these cases we draw upon earlier work, in which we did in-depth studies of the formal and informal interactions based on interview and document analysis in the case of the Barendrecht project, and an investigation of values and discourses based on media analysis and interviews in the case of the shale gas project. References to the sources used are included in the respective sections. For each of the cases, we will first give a case description followed by an analysis of the actors, discourses and values that have come to play a role in them. After the cases have been discussed, we will make an analysis of the dynamic interplay between formal and informal trajectories of assessment. This analysis will give rise to reflections on the conditions for linking formal and informal assessment, all of which could enable a responsible innovation of energy projects.

8.2 Carbon Capture and Storage in Barendrecht

The emission of carbon dioxide is one of the greatest societal challenges we are facing. Finding ways to get rid of it seems a shared responsibility (Pesch 2015). However, when the Dutch government together with Shell developed plans to capture CO_2 from the Rotterdam harbour area and to store it in an empty gas field under the town of Barendrecht, protests were fierce, eventually leading to the cancellation of all carbon capture and storage (CCS) projects in the Netherlands. So, what exactly did happen in Barendrecht? How could a project that seemed beneficial to so many powerful stakeholders go so wrong?[2]

8.2.1 Case Description

The Dutch government announced the tender procedure for onshore CCS demonstration projects in 2007. At that time, CCS was considered one of the central CO_2 emission reduction options in Dutch energy and climate policy, combined with energy efficiency and renewable energy sources. In 2008, the 'Energy Report' was adopted by parliament, containing plans for a joint CCS project between the

EDGaR is co-financed by the Northern Netherlands Provinces, the European Fund for Regional Development, the Ministry of Economic Affairs and the Province of Groningen. The Shale gas case has been studied as part of the Responsible Innovation program, funded by NWO (Netherlands Organization for Scientific Research) under Grant number 313 99,007.

[2] This case description is based on: Feenstra et al. (2012) and Cuppen et al. (2015). These two studies have been primarily based on interviews and policy documents.

Ministry of Housing, Spatial Planning and Environment and the Ministry of Economic Affairs. It was expected that CCS would be commercially viable from 2020 onwards. After a selection procedure, Shell was granted 30 million Euros for a CCS project in November 2008. The project was bound to take place in the town of Barendrecht, just south of Rotterdam.

The formally required Environmental Impact Assessment[3] (EIA) to a large extent determined the interactions that took place in the beginning of the decision-making process. As a basis for this assessment a study was initiated by the NAM (the Dutch Oil Company, the biggest oil and gas producer in the Netherlands, and responsible for gas production in the fields of Barendrecht). This study, called the AMESCO-report ('Generic Environmental Impact Study on CO2 storage') brought together a number of public and private parties to analyse the possibilities for CCS in the Netherlands. Another element of the EIA was the organisation by Shell of two public hearings to inform the local community about the project plans. The first hearing (in February 2008) included presentations from Shell and the Dutch knowledge institute TNO on CCS technology, risks, the geology of the project, the EIA procedure and the AMESCO study. In these presentations, it was emphasised that the project was 'completely safe'. Furthermore, a presentation from the national government explained the necessity of CCS in general. The second public hearing (in April 2008) attracted more people (180 as compared to 60). The number of questions and concerns raised was considerably higher, and focused mainly on the risks and safety of the technology.

From the second public hearing onwards, the interest of local stakeholders increased immensely. An information meeting was organised by the local government on 18 February 2009. This meeting was attended by about 1000–1100 people, of which the (active) majority opposed to the project. Through a sequence of meetings, the local government and citizens began to develop their position and understanding about the project and the CCS technology. The local authorities found it problematic that they had not been involved in the decision on the location of the project, and they felt that the project was strongly guided by Shell's economic interests.

As part of the procedure, citizens were given 6 weeks to submit their 'viewpoints' after publication of the EIA approval in February 2009. No less than 1570 viewpoints were handed in. The committee published its approval of the Environmental Impact Assessment on 23 April 2009. The EIA procedure led to discussion about procedural issues amongst the involved parties. Firstly, the municipal government felt that not all of its viewpoints were answered satisfactorily in the expert meetings and the EIA reports. They subsequently expressed their dissatisfaction by submitting a 'viewpoint' to the EIA. In response, they received a letter from the Ministers of Economic Affairs and Spatial Planning and Environment in June 2009 in which they announced that from now on they would discuss their decisions with the local government. The second procedural issue raised by the local

[3] In Dutch: Milieueffectrapportage.

government was that no EIA procedure had been followed for the choice of the location for the CCS project (a 'plan- EIA' before a 'decision-EIA').

Although the EIA was approved, the ministers decided to postpone the decision to continue with the project because, as they said, emotions needed to cool down and additional research needed to be commissioned to address questions raised by the local community. Three studies were commissioned: a first on geologically suitable locations; a second on impacts on human health (especially psychosomatic effects such as fear); and a third on project safety. Although these studies intended to address local questions and concerns, they took quite a technocrat position in which the health and safety risks were concluded to be acceptable by technical standards. Such a conclusion, based on risk and safety regulations, ignores the community's moral considerations such as those about the distribution of risks and benefits, both spatially and temporarily.

A policy consultation group (BCO_2) was formed mid-2008 by two Ministers and a Province deputy. The consultation group was intended to improve communications between national and local government and the wider public. Until then the debate took place mostly via press releases and media, contributing to a growing discrepancy between the viewpoints of the local and national government. Several stakeholders participated in the consultation group: a member of the Provincial Executive, the regional environmental protection agency, the alderman of Barendrecht, and two representatives of national government. The group organised public information meetings and established an information centre. The information meetings did not contribute to a shared understanding however, as it merely increased the distrust of Barendrecht residents in Shell's ambitions. Also the information centre was not fully successful, the municipality of Barendrecht decided not to join the centre as they found it hard to represent a different opinion in an information centre that is paid by the state. Moreover, the number of visitors was relatively low, suggesting that citizens did not want to rely on the information given at the centre.

On the first of March 2009, the project was included in the National Coordination Regulation. This regulation offers national government agencies the opportunity to coordinate decision-making procedures for projects in the national interest. The goal of this regulation was to both shorten decision-making procedures and accelerate project completion. Before the project fell under the National Coordination Regulation, the local Barendrecht government held sole responsibility for zoning plans and licensing decisions. The new regulation gave the national government power to overrule decisions made on the local level. This regulation was perceived by the local Barendrecht government as a means to force the project on stakeholders.

On 18 November 2009, the Ministers of Economic Affairs and Housing, Spatial Planning and Environment announced their official approval of the project. On the first of December 2009, they visited the Barendrecht community in a public hearing in the local theatre to explain their decision. About 600 people attended, and even more watched the hearing via live broadcasting in the town hall. This meeting took place in a rather hostile and emotional atmosphere. The ministers explained that

they approved the project as it was ‘absolutely safe’ and CCS a necessary technology in the transition to a sustainable energy system.

Not much later, the national government fell. This meant that structural decisions about the project had to wait for the new government to be installed (based on the elections of June of that year). In November 2010, the new Minister of Economic Affairs announced that the project in Barendrecht would not be continued because of ‘lack of support’.

8.2.2 Analysis

The case presents a number of salient moments in which formal procedures and institutional arrangements have set the stage for actors to organise themselves around certain discourses and certain sets of values.

To start, the tender procedure brought together the Ministry for Economic Affairs, the Ministry for Housing, Spatial Planning and Environment and Shell. They found each other by formulating a discourse or frame that is characterised by Cuppen et al. (2015) as ‘goal-rational’, in the sense that it set out clearly circumscribed ‘ends’ that could be achieved by the right ‘means’. The ‘end’ at stake was motivated by the national policy aspiration to attain local and national emission targets. These actors became the ‘project-owners’, and having a shared discourse functioned as a coordinative force that allowed them to keep working together. This common discourse portrayed values as boundary conditions that are subservient to this overall goal of emission reduction. The values that were considered to be relevant became manifest in the second formal moment the Amesco report. This report analysed potential safety risks of a technical nature. The direct impact on the lives of local stakeholders was, however, largely ignored. In the public hearing following the report, it was stressed that the project was ‘completely safe’. As such, the discourse of the project-owners became reinforced in the formal trajectory.

In the public hearings these findings about risks, technology and geology were presented. To an increasing extent, the municipality and local residents became displeased with the emphasis on risks. Their concern did not have much to do with the safety of the project and the associated risks, but rather they were concerned that they had no say in the way this project was established and implemented. As Noordegraaf-Eelens et al. (2012) argue, the issue was not that citizens did not want to avoid risks, but they had moral questions about the acceptability of imposing risks upon specific groups – in other words, their concern was not of a technological, but of a moral nature, especially related to issues of procedural and distributive justice. They felt that they were defenceless recipients of a new technology of which the safety could neither be proved nor refuted. Moreover, for the local residents risk implied much more than mere technical aspects, but included also for instance the financial risk regarding of the price of houses. If one would live above a ‘CO_2 bomb’, as some media interpreted the project, what would it mean for the value of your property?

The moral concerns of the residents could not easily be made part of the discourse of the project-owners from industry and government who deployed a shared discourse in order to adjust their activities. Instead of setting up a genuine dialogue between stakeholders, the project-owners persevered in their own way of thinking, repeating the mantra that 'CCS is safe', supported by the facts and figure provided by the experts. This created more and more opposition and distrust. As Terwel et al. (2012) show in a survey held in 2010 among 811 local residents, people from Barendrecht felt that the whole decision-making process was unfair (a claim supported by almost 90% of the respondents), leading to a feeling that the project-owners could not be trusted.

The organisation of public information meetings and the establishment of a public information centre could not overcome the disparity between the views of local stakeholders and the project owners. They simply could not understand each other. Moreover, the introduction of the National Coordination Regulation and the visit of the ministers to Barendrecht reinforced the goal-rational character of their discourse in full form.

8.3 Shale Gas in Boxtel

Shale gas is natural gas that is trapped in shale layers which have to be 'fracked' in order to retrieve the gas. This means that drilling has to take place by injecting fluids and chemicals underground under high pressure so that the rock is fractured and the gas is released. The policy process on shale gas exploration in the Netherlands started off as tranquil as you would expect when a relatively small firm wants to do some relatively small deals with a relatively small town, but it ended as a heated national debate that as of yet knows no winner. Basically, all of the themes that were identified in the Barendrecht case reappear in the shale gas debate with more intensity and more complexity. Again we can observe a regulatory mismatch, in which local authorities can only speak out on risks and not on the public interest. We can also observe a sharp contrast between looking for societal legitimacy, either by following formal procedures or by emergent informal mobilization of public support. Moreover, we can see how there are many interactions between the formal and informal assessment: newly formed actor groups try to get an entry into formalised institutional settings, while project owners tried to follow formal forms of assessment to reduce the heat of the public debate.[4]

[4] This case description is based on Remmerswaal (2013), Dignum et al. (2015), and Metze (2014, 2013) These studies involved interviews and media analyses.

8.3.1 Case Description

In 2009, the British oil company Cuadrilla applied for an exploration permit at the Ministry of Economic Affairs. In order to receive this exploration permit, Cuadrilla needed to submit a report on the technical aspects of the drilling process and on the economic viability of their business. The Ministry asked the Dutch knowledge institute TNO for advice on drilling for shale gas in the Netherlands. In October 2009 the ministry granted an exploration permit for test drilling on shale gas to Cuadrilla for two areas: the province of North-Brabant and Flevoland (Staatscourant 2010; 2009). This concession meant that Cuadrilla was allowed to start exploring the presence of shale gas in those provinces.

Cuadrilla approached the municipality of the town of Boxtel in the province of North Brabant in the spring of 2010. On 25 August of that year the local town board agreed with exploratory drilling in their municipality. For exploration projects, which are expected to have only a temporary impact, a full-fledged Environmental Impact Assessment is not required Early October 2010 the town board informed the town council by a letter that this application was going to be granted. The procedure of the public hearing was started: neighbours and representatives of local businesses were informed, and two information meetings were organised for neighbours and council members. These meetings were not well attended.

A local permitting procedure includes the possibility of submitting 'viewpoints'; this entails that the draft decision is up for comments by the public (citizens, action groups, and companies). Fourteen of these 'viewpoints' were submitted to the local administrators, including one of the Rabobank, a major Dutch bank that has a data centre in the vicinity of the exploration site. Most of these 'viewpoints' addressed possible nuisances caused by drilling activities, such as traffic nuisance, odour nuisance, spoiling the scenery, concerns about seismographic activities (minor earthquakes) and water contamination. Moreover, some parties of the council asked questions about the environmental aspects. In spite of these concerns, the town board saw no further problems and released a permit on 11 January 2011.

Four groups of people that had submitted 'viewpoints' were not satisfied by the way the town board had included their objections in the permit, and they started a lawsuit against the decision. The main objection was that the permit had been issued for a temporary drilling activity, while there were, as it was claimed, not enough guarantees that this would indeed be temporary. Moreover, it was argued that there was not sufficient argumentation and evidence that ensured the spatial quality – more specifically the safety, air quality, seismic activities, the quality of the soil and ground water and noise nuisance. The argument was that the living environment would be 'unreasonably' damaged.

In the period between the filing of a lawsuit and the ruling of the court in October 2011, societal resistance grew; citizens of the nearby municipality Haaren were the first to object to drilling for shale gas in their neighbourhood, and formed a group called 'Shale Gas No'. The turmoil spread to Boxtel, where a similar action group was founded, existing both of local residents worried about falling house prices, and

of residents who had moved to the Brabant countryside for its natural surroundings. The action groups were especially worried about the risks and environmental damage that the drilling could bring (Milieudefensie 2013). The activists received support from the provincial environmental federations, and other activists. Soon they received support from left-wing national political parties. The provincial parliament of Noord-Brabant got concerned and sent a letter together with the municipality of Boxtel to the Minister of Economic Affairs to ask for a moratorium and an independent study on the pros and cons of shale gas exploration and production. The Minister, in response, believed the legislation and experience within the Netherlands to be sufficient for exploratory drilling and turned down the request. However, disappointed responses eventually put pressure on him and led him to request Cuadrilla to perform more studies.

Initially the social upheaval did not cause a change in direction by the ministry and the municipality and the permit was granted to Caudrilla. Still, the shale gas debate continued as more requests for shale gas exploration permits came at other places in the Netherlands were made. This led the environmental NGOs to write to several municipalities to successfully ask them to declare themselves 'shale gas free'. Moreover, the areas of interest started overlapping with the operating area of water company Vitens, which started to warn for water pollution as well, just like the water company Brabant Water. The debate gained even more attention when the movie *Gasland* was broadcasted on Dutch national television in September 2011.

Meanwhile, the Dutch administrative court decided in October 2011 that the municipality of Boxtel had not followed the right procedures. If exploration would show that natural resources were to be economically producible, a production permit could be granted for the production of the gas. Therefore, shale gas exploration was not considered a temporary activity, which was the premise of the exploration permit and Cuadrilla's permit was withdrawn. Two days later, at a Parliamentary debate, the Minister responded to the worries and requests of municipalities and announced an independent study into the risk of shale gas and coal-bed methane exploration and production. As long as the research was not finished, no exploration wells would be drilled, and no new permit applications would be taken into procedure.

After the announcement of an independent study, media attention decreased. Protest groups asked the Minister to broaden the research on safety and environment with the utility and necessity of shale gas, but the Minister responded that these are political matters and not something for this research. The Minister did promise to consult multiple stakeholders for setting up the research questions. This led to a broad range of questions, which could not all be answered in the time available for the study.

In spite of the relative peace, the shale gas debate became more national. In April 2012, a national anti-shale gas NGO ('Schaliegasvrij Nederland') was founded by local communities and environmental organisations. The protest groups continued with several protest actions and actively reported on their interpretation of shale gas developments.

At the beginning of the year 2013, earthquakes in Groningen of 3.7 on the Richter scale due to gas production by the NAM exceeded the expected 3 on the Richter scale. Although Schaliegasvrij Nederland already linked (small) tremors in Groningen to gas production in August 2012, in February 2013 more actors started raising questions about the risk of earthquakes caused by fracking. The earthquakes in Blackpool were brought up again and in the political domain more questions about soil subsidence were asked.

Also the claim that shale gas was relatively sustainable compared to other fossil fuels received criticism. In February 2013, a group of professors sharply criticised shale gas developments. Their main reproach was that policy makers had not strategically considered the impact of shale gas production on the transition to a sustainable energy supply.

In August, the research which had been requested by the minister was published. The study concluded that the risks of shale gas exploration and production were low and could be mitigated. However, the minister did not want to take any decisions without having an assessment of the study by the EIA committee. This assessment was not at all supportive for making fast decisions, as it stated that the scope of the research had been too narrow by focusing mainly on underground effects. The committee also stated that the current regulation was not sufficient to cover the risks of shale gas exploration and production and advised the Minister to use a governmental spatial planning procedure to cover both underground and above ground impacts of shale gas developments (Commissie Milieueffectrapportage 2013). The Minister decided to postpone his decision and announced a new study that would focus on all potential interesting locations so that local interests could be involved as well. The so called 'Rijksstructuurvisie' would be used together with an Environmental Impact Assessment to find out which locations within the Netherlands are most promising and to find out at which locations risks are easiest to mitigate. The Minister also announced that he wanted to involve local policy makers and the local communities of the potential locations. Moreover, the research would be used to investigate the technical options for risk mitigation together with water companies and the mining industry. The moratorium was prolonged, so for the time being – at least as long as the current cabinet is in charge – no important decisions will be taken (Ministerie van E.L.I. 2013).

8.3.2 Analysis

In comparison to the Barendrecht case, the debate on shale gas reveals a much more complex pattern of value articulation. Two features that contribute to this complexity are the scaling up of the debate from a local to a national concern and the multitude of values that have been a topic of the debate (Dignum et al. Forthcoming). To make sense of the latter, we follow Remmerswaal (2013) who reconstructed three clusters of values by analysing an extensive set of newspaper articles on shale gas that had been published in Dutch newspapers during the periods described above.

These clusters can be classified as ‘safety and environment’, ‘utility and necessity’ and ‘procedural justice’. Each of these clusters figures as a frame within which proponents and opponents express their assessment of shale gas. For instance, the ‘safety and environment’ cluster came to include issues like the possible contamination of water and later also the impact of tremors. In the ‘Utility and Necessity’ cluster, the discussion was on the question how shale gas is useful in a transition towards sustainable energy and how it can decrease dependence on other countries. In the ‘Procedural Justice’ cluster, proponents have pointed out that the Dutch laws on regulation and inspection assure safe exploration and production of shale gas, while opponents dispute the fairness of the decision making process.

The task for decision-makers had been to accommodate emerging values and issues brought forwards in the societal debate, while at the same time responding to the three key objectives that have been formalised in Dutch energy policy, which are sustainability, security of supply, and affordability. This tension can be further sketched by the distinction between *substantive* values and *procedural* values introduced by Dignum et al. (2015), who collected and analysed the values that have been forwarded in both online and traditional media. The values belonging to the *substantive* type of values relate straightforwardly to the three key objectives of Dutch energy policy presented above. In addition to those values that have become embodied in formalised arrangements such as environmental impact assessment and permitting procedures, there are *procedural* values which relate to the nature of the rules and regulations and the procedures that constitute the decision-making on the exploration and exploitation of shale gas. Three main values were identified in this group: *distributive justice*, *procedural justice*, and *accountability*. Each of these values related to different aspects of the procedure surrounding shale gas exploration and exploitation. These latter issues concern aspects such as the limitations of current legislative frameworks, access of stakeholders to the decision-making process, justice and transparency.

For each of the substantive values, we can recognise a similar dynamic to that of the Barendrecht case. There is a formal conception of such a value that is dominant, giving rise to the emergence of different publics that come to contest this conception. How shale gas exploitation may contribute to economic welfare, the transition to a sustainable energy system, and what the risks are for public health and the environment are questions to which widely divergent answers are given in the debate. Furthermore, the given structure of operationalising the substantive values in official policy invokes the resentment among publics about the legitimacy of the policy process in general – which is articulated in the emergence of the set of procedural values. In other words, the formal assessment provokes the emergence of informal assessment, which by its very nature is heterogeneous, inarticulate, and messy, and with that challenging the efficacy of decision-making processes.

This challenge necessitated policy makers to continuously develop strategies to balance the appeals of both formal and informal assessment. Given the interplay between formal and informal assessment, strategies will have the propensity to backfire – which is what can be observed in the case of shale gas. In reaction to shale gas becoming a national concern, the Minister announced a moratorium and

issued an independent study. This was a strategy to buy more time and to encapsulate informally expressed worries about shale gas by a formal form of assessment in the manifestation of an expert-based study. Indeed, the intensity of the societal debate seems to have diminished after that. However, the second leg of the strategy did not seem to work out, as the organisations that received the assignment to study the risks of shale gas were criticised for having economic interests in shale gas exploration. The critique on the announced study allowed environmental groups to spread their viewpoints on shale gas and also to create an institutional toehold. For instance, a new nationwide anti-shale gas NGO was founded by local communities and diverse environmental organisations. This new NGO could become a member of the feedback group that had to report on the study on the risks of shale gas. By having direct insight, this NGO had first-hand knowledge about the results, and it could immediately react by publishing its assessment in press releases and on a range of websites.

The additional study was not at all successful in its goal to figure as a shared basis for the debate – on the contrary. The announcement of the Minister to postpone the decision for shale gas exploration even further might be seen as another attempts to cool down the debate. Whether this attempt will work out or not is still unclear, at this moment it seems that the enthusiasm of industry for shale gas has suffered a severe blow. The profits of US shale gas endeavours have proven to be dramatically low, with sincere environmental drawbacks. The controversial nature of shale gas in Europe makes it a thorny affair; companies are quite reluctant to invest in such an unpopular resource. But most of all, the vast decline of the prices of fossil resources – to which the US shale boom has contributed substantially – makes investment in expensive forms of exploitation a financial move that is highly precarious.

8.4 The Reproduction of Positions: Contrastive Discourses

Responsible innovation revolves around the identification and endorsement of public values in the design of the technology and institutional context. What both our studies reveal is that it is not the case that different actors hold on to different sets of values. Welfare, safety, sustainability, accountability, etc. are not contested. Instead, different actors have different conceptions of one and the same value; they adhere to different discourses that recruit specific interpretations of the values at stake.

In turn these discourses do not emerge independently, but are deeply interrelated. Especially the discourse that can be associated with informal assessment is developed in reaction to the decisions and assessments that are made as part of the formal trajectory. For instance, project owners see their efforts to prove the safety of their project and their observance to official procedures as proof that they held the public legitimacy in high regard. Local actors however see these efforts as tokens of the technocratic inclination of the project owners to impose decisions in a top-down fashion. The discourses that emerge out of such interactions lead to mutual distrust,

and to a progressive entrenchment of positions: project owners will even try harder to reduce technical risk, to emphasise the improbability of risk and to stick to the official decision-making procedures, giving further nourishment to the sentiment of local actors that these project owners are not reliable.

The two cases reveal that formal forms of assessment are the hinges on which both the policy process and the societal debate rest. In the policy domain, the role of formal assessments contributes to the alignment of actors, discourses and values. They figure as 'boundary objects', which are items that inhabit intersecting social worlds that allow the negotiation process between different stakeholders to take place (Star and Griesemer 1989; Jasanoff 1990). In other words, formal assessment is instrumental for actors that represent institutional stakes to 'close down' the variety of decision-making options (Stirling 2008).

For the societal debate, formal assessment, like other official policy actions, may trigger the involvement of new actors, the formation of new discourses and the introduction of previously unarticulated values. To which extent concerns have already lingered beneath the surface is hard, if not impossible, to establish in hindsight. These concerns are too obscurely articulated for that. It is the introduction of a formal assessment that allows these concerns to become crystallised and, as such, visible. With that, formal assessment allows the 'opening up' of the process of appraisal in the sense that a wider set of actors, discourses and values is recruited that have to be given consideration – while paradoxically, these assessments at the same time allow the 'closing down' of the range of policy options for institutional stakeholders. The findings that are made public through formal procedures trigger the awareness of a wider set of societal actors that a new policy issue has been set upon the agenda, about which, obviously so, a normative appraisal still has to be formulated. For these actors that are found outside of the conventional institutional policy, the formal assessment is a starting point for deliberation, instead of an end point.

Another finding is that the democratic legitimacy of existing rules and procedures is not taken at face value. One of the sources of mutual misunderstanding between institutional stakeholders and emergent publics seems to be a contrastive understanding of the societal legitimacy of a decision. For most institutional actors, the fairness of a procedure is guaranteed by following the rules and by consulting the well-known interest groups. First, these rules have been established in democratic processes and as such are considered to be the best expression of the public interest. Second, societal legitimacy is pursued by involving the representatives of an array of public interest organisations (cf. Visser and Hemerijck 1997). The representativeness and the scope of organisations are considered to be a sufficient source of legitimacy. New actors, new discourses and new (conceptions of) values may emerge that undermine this standard way to warrant the legitimacy of a project or decision. The new groups that are established by societal stakeholders are not (yet) involved in consultation rounds on new policy decisions; moreover, a lot of citizens do not feel automatically represented by the classic NGOs (cf. Pesch 2014). Societal legitimacy is consequently sought outside of the circle of 'usual suspects' that are consulted by policy makers.

Moreover, there is a regulatory problem regarding the satisfactory allocation of responsibilities: energy technologies are often introduced as a national affair, while risks are seen as a local concern. As local authorities are only allowed to say something regarding their own jurisdiction, they have to focus on the risks of new energy technologies while refraining from any claim about the utility and necessity of these new technologies. Likewise, the people that are subjected to the risks of a new technology are not the same people that yield its benefits. So local stakeholders are compelled to focus only on risk, while neglecting other kinds of worries and arguments; at the same time, these stakeholders are not allowed to say something about the overall desirability of a certain technological development. This rigid separation of jurisdictions significantly contributes to discontent that finds its way through informal channels of assessment, while this discontent is often simply labelled as nimbyism[5] and risk-aversion.

At the same time, actors try to *strategically* use formal and informal assessment to pursue their own goals. For instance, aligning with the discourse connected to informal assessment creates 'street credibility' – so to speak – for environmental NGOs and municipalities. It gives them the opportunity to make ideological alliances with local residents and wider societal opinion. At the same time, these actor groups clearly try to influence formal assessment by administrative jurisdiction, filing lawsuits, membership of consultation platforms, and so on. One may also observe that project owners on their side try to influence informal forms of assessment by strategically deploying formal forms of assessment. Most clearly this is the case in the shale gas debate in which the Minister attempted to reduce the intensity of the controversy by announcing science-based studies. This allowed the minister to buy time, and create legitimacy by emphasising his objectivity. The Barendrecht case showed similar initiatives for instance with the establishment of a public information centre.

8.5 Finding Symmetry Between Formal and Informal Assessment

In the end, the question is about the public legitimacy of energy projects, and more concretely so, who is or who may represent the public? The answer is that there is no tangible phenomenon that can be labelled as *the* public; in its very essence society cannot be unambiguously operationalised (Schubert 1960; Pesch 2008; 2005; Huitema et al. 2007). Different parties may have aspirations to monopolise the claim for the public interest, either in a technocratic direction in which the public interest is guarded by national authorities, leading to the exclusion of arguments and sentiments of local populations, or in a populist direction, in which it is presupposed

[5] Nimby is the acronym for 'Not in my backyard', suggesting an egoistic disposition of local protestors.

that only voices from outside the institutional system are credible expressions of what the 'public' really wants.

We have stated in the introductory section that responsible innovation revolves around the endorsement of public values in the design of a technology and the institutional context in which this technology is embedded. Our analysis here shows that in case of energy technologies, the values that are to be endorsed may be hard to identify. Moreover, there is no method – at least not yet – to decide beforehand which value conceptions can be singled out as genuine, legitimate, and sound. How to identify the right value conceptions and how to decide upon their validity is an issue that will have to be explored in a future study, for now we can infer from our analyses that, in any case, the *symmetry* between different claims for the public interest needs to be attended by policy makers. Instead of taking existing institutional arrangements for granted, we have to accept the dynamics that pertain to both society and the energy system. This implies that, rather than an ex ante assessment of ethical and societal aspects, responsible innovation involves a continuous process of assessment (Taebi et al. 2014; Stilgoe et al. 2013; Taebi et al. Forthcoming). The question becomes how can we link formal and informal assessment of an energy technology in a symmetrical way?

To do so, we have to emphasise that there is no party which has the exclusive appeal to represent 'the' public. Ideally, a decision is based on a dialogical form of deliberation: no party should prescribe the rules of the game in terms of which interests and discursive frames are valid, and which modes of expression are acceptable. A level playing field should be created among all actors involved which should allow a dialogue about knowledge claims about the impacts of the technology, local concerns, and the conditions under which a new technology would become acceptable (Cuppen 2009; Roeser and Pesch 2015; Taebi et al. 2014; Taebi forthcoming). To have symmetry of assessment, decision making trajectories should allow flexibility and diversity (Cuppen et al. forthcoming). It should be possible to renegotiate how the public interest is formulated in real-life decisions (Pesch 2014), and decision making processes should be capable of accommodating divergent normative perspectives (Stirling 2010).

At the same time, we have to acknowledge that in contemporary society there will always be a fundamental asymmetry in the way political leverage is distributed. Collective decisions are the result of agenda-setting processes in which a small segment of actors have much more ability and opportunity to define which issues will receive public notice (Kingdon 1984). Project owners simply have a head start; they have the chance to develop a shared discourse in relative isolation that includes a clear idea of goals, values, and preferences. Other actors can only react in response to this discourse, their values and preferences can never be known beforehand. Only by acknowledging the potential of emerging values and by being prepared to adjust decisions, technological designs and institutional provisions, innovation processes in energy projects can be managed responsibly.

References

Bauman, Zygmunt. 2000. *Liquid modernity*. Cambridge: Polity.

Breukers, Sylvia, and Maarten Wolsink. 2007. Wind power implementation in changing institutional landscapes: An international comparison. *Energy Policy* 35 (5): 2737–2750. doi:10.1016/j.enpol.2006.12.004.

Commissie Milieueffectrapportage. 2013. Brede afweging schaliegas mist nog. http://api.commissiemer.nl/docs/mer/p00/p0023/persbericht_effectstudie_schaliegas.pdf. Accessed 18 Aug 2016.

Cuppen, Eefje. 2009. *Putting perspectives into participation: Constructive conflict methodology for problem structuring in stakeholder dialogues*. Oisterwijk: Boxpress.

Cuppen, Eefje, Suzanne Brunsting, Udo Pesch, and Ynke Feenstra. 2015. How stakeholder interactions can reduce space for moral considerations in decision making: A contested CCS project in the Netherlands. *Environment and Planning A* 47 (9): 1963–1978.

Cuppen, Eefje, Udo Pesch, Sanne Remmerswaal and Mattijs Taanman. Forthcoming. Normative diversity, conflict and transition: shale gas in the Netherlands. *Technological Forecasting and Social Change*. doi:https://doi.org/10.1016/j.techfore.2016.11.004

Devine-Wright, Patrick. 2005. Beyond NIMBYism: Towards an integrated framework for understanding public perceptions of wind energy. *Wind Energy* 8 (2): 125–139.

Di Ruggero, Olga. 2014. *Anticipating public acceptance: The hydrogen case*. Delft: Delft University of Technology.

Dignum, Marloes, Aad Correljé, Eefje Cuppen, Udo Pesch, and Behnam Taebi. 2015. Contested technologies and design for values: The case of shale gas. *Science and Engineering Ethics* 22 (4): 1171–1191. doi:10.1007/s11948-015-9685-6.

Dignum, Marloes, Udo Pesch, and Aad Correljé. Forthcoming. Frames of reference and the interpretation of values in the Dutch shale gas debate. In *New perspectives on responsible innovation*, ed. Roland Orrt, Martijn Blaauw and Jeroen van den Hoven. Dordrecht: Springer.

Feenstra, C.F.J., T. Mikunda, and S. Brunsting. 2012. What happened in Barendrecht?! Case study on the planned onshore carbon dioxide storage in Barendrecht, the Netherlands. http://www.globalccsinstitute.com/sites/www.globalccsinstitute.com/files/publications/8172/barendrecht-ccs-project-case-study.pdf. Accessed 18 Aug 2016.

Huijts, Nicole M.A., Eric J.E. Molin, and Bert van Wee. 2014. Hydrogen fuel station acceptance: A structural equation model based on the technology acceptance framework. *Journal of Environmental Psychology* 38: 153–166.

Huitema, Dave, Marleen Van de Kerkhof, and Udo Pesch. 2007. The nature of the beast: Are citizens' juries deliberative or pluralist? *Policy Sciences* 40 (4): 287–311. doi:10.1007/s11077-007-9046-7.

Jasanoff, Sheila. 1990. *The fifth branch: Science advisers as policymakers*. Boston: Harvard University Press.

Kingdon, John W. 1984. *Agendas, alternatives and public policies*. Boston: Little Brown.

Ligtvoet, Andreas, Eefje Cuppen, Kas Hemmes, Donna Mehos, Udo Pesch, Jaco N. Quist, and Olga Di Ruggero. 2015. De komende 50 jaar gas in Nederland–perspectieven en robuuste strategieën.

Ligtvoet, Andreas, Eefje Cuppen, Olga Di Ruggero, Kas Hemmes, Udo Pesch, Jaco Quist, and Donna Mehos. 2016. New future perspectives through constructive conflict: Exploring the future of gas in the Netherlands. *Futures* 78: 19–33.

Metze, Tamara. 2013. What the frack? Development of the controversy about hydraulic fracking for shale gas in the Netherlands. (June 27, 2013). doi: 10.2139/ssrn.2285995

———. 2014. Fracking the debate: Frame shifts and boundary work in Dutch decision making on shale gas. *Journal of Environmental Policy & Planning* ahead-of-print: 1–18.

Milieudefensie. 2013. *Factsheet schaliegasvrije gemeenten*. https://milieudefensie.nl/publicaties/factsheets/factsheet-schaliegasvrije-gemeenten. Accessed 11 Aug 2016.

Ministerie van Economische Zaken, Landbouw en Innovatie 2013. *Brief aan de Tweede Kamer – Vervolgstap schaliegas*. https://www.rijksoverheid.nl/documenten/kamerstukken/2013/08/26/brief-aan-de-tweede-kamer-schaliegas-resultaten-onderzoek-en-verdere-voortgang. Accessed 11 Aug 2016.

Noordegraaf-Eelens, Liesbeth H.J., Michel van Eeten, Marjolein Februari, and Jony Ferket. 2012. Waarom Burgers risico's accepteren en waarom bestuurders dat niet zien. repub.eur.nl/pub/38337/metis_183032.pdf. Accessed 11 Aug 2016.

Owen, Richard, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and David H. Guston. 2013. A framework for responsible innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, 27–50. Somerset: Wiley.

Pesch, Udo. 2005. *The predicaments of publicness: An inquiry into the conceptual ambiguity of public administration*. Delft: Eburon.

———. 2008. The publicness of public administration. *Administration and Society* 40 (2): 170–193. doi:10.1177/0095399707312828.

———. 2014. Sustainable development and institutional boundaries. *Journal of Integrative Environmental Sciences* 11 (1): 39–54.

———. 2015. Publicness, privateness, and the management of pollution. *Ethics, Policy & Environment* 18 (1): 79–95.

Pidgeon, Nick F., Irene Lorenzoni, and Wouter Poortinga. 2008. Climate change or nuclear power—No thanks! A quantitative study of public perceptions and risk framing in Britain. *Global Environmental Change* 18 (1): 69–85. doi:10.1016/j.gloenvcha.2007.09.005.

Remmerswaal, Sanne. 2013. *The dynamics in the societal debate on shale gas in The Netherlands*. Master thesis Delft University of Technology.

Rip, A. 1986. Controversies as Informal Technology A ssessment. *Science Communication* 8: 349–371

Roeser, Sabine, and Udo Pesch. 2015. An emotional deliberation approach to risk. *Science, Technology & Human Values*, On-line first.

Schubert, Glendon. 1960. *The public interest: A critique of the theory of a political concept*. Glencoe: The Free Press.

Staatscourant. 2009. *Besluit opsporingsvergunning Noord-Brabant*. https://zoek.officielebekendmakingen.nl/stcrt-2009-16000.html. Accessed 11 Aug 2016.

———. 2010. *Opsporingsvergunning koolwaterstoffen Noordoostpolder*. https://zoek.officielebekendmakingen.nl/stcrt-2010-9431.html. Accessed 11 Aug 2016.

Star, Susan Leigh, and James R. Griesemer. 1989. Institutional ecology, translations' and boundary objects: Amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907–39. *Social Studies of Science* 19 (3): 387–420.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580.

Stirling, Andy. 2008. "Opening up" and "Closing down" power, participation, and pluralism in the social appraisal of technology. *Science, Technology & Human Values* 33 (2): 262–294.

———. 2010. Keep it complex. *Nature* 468 (7327): 1029–1031.

Taebi, Behnam, Aad Correljé, Eefje Cuppen, Marloes Dignum, and Udo Pesch. 2014. Responsible innovation as an endorsement of public values: The need for interdisciplinary research. *Journal of Responsible Innovation* 1 (1): 118–124.

Taebi, Behnam and Sabine Roeser, eds. 2015. *The Ethics of Nuclear Energy: Risk, Justice and Democracy in the post-Fukushima Era*. Cambridge: Cambridge University Press.

Taebi, Behnam, Aad Correljé, Eefje Cuppen, Elisabeth Van de Grift, and Udo Pesch. 2016. *Ethics and the impact assessment of large energy projects*. 2016 IEEE International symposium on ethics in engineering, science and technology conference, Vancouver, 13–14 May 2016.

Taebi, Behnam, Sabine Roeser, and Ibo Van de Poel. Forthcoming. Responsible innovation of nuclear energy technologies: Social experiments, intergenerational justice, and emotions. In *New perspectives on responsible innovation*, ed. Roland Ortt, Martijn Blaauw and Jeroen van den Hoven. Dordrecht: Springer.

Taebi, Behnam. Forthcoming. Bridging the gap between social acceptance and ethical acceptability. *Risk Analysis*. https://doi.org/10.1111/risa.12734.

Terwel, Bart W., Emma ter Mors, and Dancker D.L. Daamen. 2012. It's not only about safety: Beliefs and attitudes of 811 local residents regarding a CCS project in Barendrecht. *International Journal of Greenhouse Gas Control* 9: 41–51.

Visser, Jelle, and Anton Hemerijck. 1997. *A Dutch miracle: Job growth, welfare reform and corporatism in the Netherlands*. Amsterdam: Amsterdam University Press.

Chapter 9
Social Learning and Identity: Some Implications for RRI

Lotte Asveld

Abstract The core of the concept of responsible research and innovation (RRI) is the principle of mutual responsiveness; that is, actors should engage in a transparent and interactive exchange of values, concerns and hopes regarding a new technology to arrive at a shared perspective. As such, RRI can be considered a form of social learning. However, whether such mutual responsiveness is feasible depends on the identities of the actors involved. Identities consist of moral framework and a collection of social roles. Identities influence our willingness and capacity to engage in social learning exercises such as RRI. In this paper, I argue that taking the issue of identity into account can help structure RRI exercises to enhance their effectiveness. It can also make us more precise about which societal goals we can achieve through RRI and for which goals we need additional measures.

9.1 Introduction

The responsible research and innovation (RRI) approach can be understood as an attempt to align new technologies with societal concerns and needs. RRI is intended to help designers and manufacturers of new technologies identify and accommodate public concerns when developing a new technology by engaging with a wide range of relevant actors (Van den Hoven and Jacob 2013). As such, RRI can be considered a tool to answer questions about the direction in which we would want to use available scientific and technical knowledge (cf. Owen et al. 2012).

However, RRI processes do not take place in a social vacuum. Actors engaged with a specific innovation have social identities that influence their willingness to engage and/or their capacity for engagement with specific other actors. Such identities can be recognised in public debates on new technologies such as synthetic biology, nanotechnology or energy technologies, where they become manifest as the social roles and values of participants.

L. Asveld (✉)
Department of Biotechnology & Society, Delft University of Technology, Delft, The Netherlands
e-mail: l.asveld@tudelft.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_9

For instance, the spokesperson for an environmental NGO may not want to be perceived as supporting the products of a big company if his usual peer group considers that company to be environmentally harmful. Hence, he may refrain from constructing a shared vision on an innovation trajectory with representatives from this company. Likewise, the CEO of a big company might be unwilling to adjust procedural operations within her company to allow for social reflections on its methods of production. She might perceive a conflict between the stability of her company and opening up her production methods to external influences (Blok and Lemmens 2015).

In addition, technology influences the way we understand ourselves and the moral values we cherish (Swierstra, Chap. 2, this volume). This implies that any understanding about the desirability of applications is necessarily related to the understanding we have of ourselves and of who we desire to be. We assess technologies on the basis of our values, the impact they will have on our particular lives and the kind of future we see for ourselves. Hence, the concept of identity is highly relevant to the process of RRI and should be taken into account. This paper is an attempt to explicate the role of identity for RRI processes.

The question addressed in this paper is: "In what way can the identity of involved actors be expected to influence RRI and what lessons might be drawn from this for designing RRI?" I argue that identity should be understood as a combination of social role and moral framework. These two aspects are related and both have specific ramifications for RRI. The importance of moral frameworks and the need to reflect on these when evaluating the social desirability of a technology has been extensively described in the literature (Kupper et al. 2007; Cuppen 2012; van de Poel and Zwart 2010; Kahan 2012; De Witt et al. 2015; Asveld and Stemerding 2017). The aspect of social role has been used to identify the possible societal ramifications of new technologies, for instance, in the value sensitive design approach (Davis and Nathan 2015). Roles in this case are used as a spotlight to identify societal effects of a technology. Which roles will be affected by a particular technology and in what way?

How social roles in turn affect the propensity to construct a shared vision, or how they influence perceptions of acceptable designs, has however received little attention. We claim that taking identity into account helps us improve the design of RRI exercises to achieve socially robust innovations. In addition, including the notion of identity enables us to arrive at outcomes of RRI processes that can be transported across contexts, and hence allows for the upscaling of those outcomes.

9.2 RRI as Social Learning

To achieve socially robust innovations, the aim of the RRI approach is to include a broad spectrum of actors in the process of innovation. RRI has been described as:

> A transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to 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). (Von Schomberg 2011, p.9)

As such, RRI can be considered a social learning exercise. Social learning implies that individuals not only learn from each other, but also achieve learning results as a group (Röling 2002; van Mierlo et al. 2010; Armitage et al. 2008). More specifically: "Social learning may be defined as a change in understanding that goes beyond the individual to become situated within wider social units or communities of practice through social interactions between actors within social networks" (Reed et al. 2010). RRI can hence be considered a form of social learning, since the aim is to produce a new understanding within a network of social actors related to a new technology. It is not only the designers of new technologies who learn to accommodate social concerns while engaged with RRI; other social actors can also learn to formulate a new perspective on a newly emerging technology.

Social learning that takes place within the setting of RRI is usually geared towards such moral questions as: is this really a desirable technology? Who benefits from this technology and who bears the burdens? What should be the targets for innovation? (Owen et al. 2012). Such questions are usually characterised by a wide range of possible answers, stemming from the wide range of diverse moral frameworks present in society (Kupper et al. 2007). As such, the issues that are addressed within RRI can be understood as morally ambiguous, implying that it is unclear what the 'right' answer is, because there are many different viable positions stemming for the various moral frameworks of the actors involved (Asveld and Stemerding 2018).

A social learning exercise in which participants hold varying perspectives can result either in a shared perspective that also includes shared supporting values (because these were there already), or in a perspective that is compatible with each of the participants' values, although these values are not necessarily shared (van de Poel and Zwart 2010; Röling 2002). For instance, when a scientist and an environmental activist both deem it undesirable to apply synthetic biology to agriculture on a large scale, the scientist may hold this position because she thinks more research is needed to identify the possibly risks, whereas the activist may think that synthetic biology is a technology that reinforces the root cause of environmental problems, namely the concentration of power in the hands of a few. They arrive at a shared evaluation of the technology but on the basis of different values, namely precaution and economic equality. In contrast, if the scientist and the activist both think that the synthetic biology application should be temporarily or definitely stopped out of precaution, then their agreement also extends to their values.

A third option for some form of agreement between various actors is a compromise. In this case the actors do not arrive at a shared understanding; rather, they construct a way of dealing with new technology that is partly compatible with their individual values. Each of the actors gains something and loses something. If a scientist believes, for instance, that synthetic biology is safe to apply to agriculture, but the environmental activist believes it can have detrimental effects on global economic equality, they might agree that synthetic biology should be allowed only when it is applied on a small scale and is publicly financed. This solution does not comply completely with each of their respective positions, but it does go some way to accommodate their concerns.

To arrive at a shared perspective, or even just a compromise, as an outcome of social learning it is important to acknowledge the various moral frameworks from which actors operate, since these determine the options for arriving at a shared perspective. In addition, actors are embedded in social structures that determine their economic and other dependencies and that entail institutional and social constraints on their engagement in RRI processes. These two elements are captured by the notion of identity and should be explicitly taken into account to increase the chance of a socially robust innovation trajectory.

9.3 Identity

Identity is a conception of who one is and who one wants to be. As Taylor (1989, p. 27) puts it:

> To know who I am is a species of knowing where I stand. My identity is defined by the commitments and identifications which provide the frame or horizon within which I can try to determine from case to case what is good, or valuable, or what ought to be done, or what I endorse or oppose.

As such, identity consists of both social roles and a specific worldview, or framework as Taylor calls it, comprising values, norms, assumptions and convictions (cf. Hedlund-De Witt 2013). These two elements are intertwined. Socially defined identities are a source of values (Appiah 2010). A Catholic cherishes the teachings of the Pope, whereas a feminist strives for gender equality. Identities usually consist of a variety of roles that together make up the overarching value orientation of the individual.

9.3.1 Moral Frameworks

Identity provides a moral framework with which we and other people are judged. Our moral framework helps us to evaluate the world around us and to find our way through it (Taylor 1989; Asveld 2008). This moral identity framework contains the resources from which we form our moral judgements. It can provide us with a sense of belonging when we share a set of values with other individuals, or with a sense of uniqueness when we find out that others hold different values (Appiah 2010). The moral character of this framework may not always be easily discernible. Norms and values may be intertwined with other elements of an identity, for instance cultural interpretations of specific roles.

Individual moral frameworks can be linked to socially prevalent frameworks, which can be termed worldviews. Worldviews are culturally dominant frameworks of meanings (Hedlund-de Witt 2013; Kahan 2012). Worldviews consist of coherent structures of values, beliefs and attitudes. Individuals' frameworks will never completely coincide with these worldviews, but individuals tend to hold a certain

worldview more than another. Individual frameworks are more complex and nuanced than worldviews, but the individual frameworks often take shape with reference to the overarching sociocultural worldviews (Asveld 2008). Worldviews help us to grasp some of the recurring tensions in debates on new technologies even if we have not explicated the depth of each individual framework involved in such tensions (Kupper et al. 2007; Asveld and Stemerding 2017).

Worldviews are the subject of a wide body of literature, in which they are usually divided into four distinct categories. Defining characteristics include such issues as an orientation towards the local or the global (Douglas and Wildavsky 1982; De Vries and Petersen 2009), perceptions of the role of the government (De Vries and Petersen 2009), the vulnerability of the natural environment (Douglas and Wildavsky 1982; Kahan 2012), our responsibilities towards that environment (Dryzek 2013) and expectations about the possible beneficial effects of technology (Hedlund-De Witt 2013).

Worldviews relate to social roles, as many moral outlooks coincide with specific social roles. The CEOs of multinational companies can be expected to have a more global oriented worldview in which the role of government is minimal and mostly confined to assuring safety and security. They can be expected to generally subscribe to the idea that environmental problems can be sufficiently addressed through innovation and regulation. In contrast, environmental activists will generally subscribe to the idea that innovation and regulation are insufficient to achieve a sustainable world, and that we need a change in attitude, that is, we need to minimise our consumerist attitude.

9.3.2 Social Roles

As described above, moral frameworks are often connected to our professional identities and other social roles, although our personal moral framework and our professional values need not necessarily overlap. For instance, a physician who smokes may be acting against his professional values, but not necessarily against his personal values (except when the smoking arises out of a weakness of will). Especially for activists, their moral framework is usually highly constitutive to their social role. Providing a moral evaluation of current issues is their core business. The moral framework can be said to make up the core of their identity as activists. This is why it may be difficult to get them to open up to an alternative view.

Social learning can be said to require a willingness to reflect on one's own perspective and assumptions, and to consider them with a certain openness to allow for the creation of a shared perspective with others. We should not expect individuals to easily change their moral framework, precisely because identity is constitutive to our self-understanding and hence to our stability as persons. But even if an individual does not change her moral framework, she should still be willing to openly reflect on her values and the ensuing assessment of technologies to allow for social learning.

This willingness to reflect might, however, be constrained by some aspects of the individual's identity. Because our moral frameworks are connected to our social

identity, the extent to which our moral convictions are open to social learning varies (cf. Reed et al. 2010). Moral frameworks are to some extent externally dictated, because they are part of the social role the individual takes on. For instance, an academic is supposed to value truth-finding over activism, and a parent is supposed to value the long-term wellbeing of his children over his personal immediate pleasure. We cannot expect people to abandon the moral judgements associated with moral values that are fundamental to a cherished social role. For instance, we cannot ask a scientist to put aside his scientific standards for a moment so that he might be more open to construct a shared vision with someone who rejects the scientific method.

An individual can of course always chose to abandon a social role should the moral framework change and she no longer feels comfortable with it. An activist criticising synthetic biology might change her mind about this and give up her position with an NGO and become a consultant helping synthetic biology companies to produce sustainable technologies. But as long as she takes on a specific social role, her options for social learning, namely to reflect on her moral framework in interaction with others and to arrive at a shared perspective, will be constrained by the social expectations related to her social role.

To understand how this impacts on RRI exercises and how we can accommodate these constraints in RRI, we should make a distinction between those aspects of identity that are externally dictated and those that are more open to individual interpretation. These latter aspects of identity leave room for a specific individual perception, without undermining the communicative value or the credibility of identity. Some basic features of social roles are indispensable for them to convey social meaning. To be a mother, one has to have (or to have had) a child. But, to be a mother, one need not have given birth to a child as there are other means available to have a child. To be considered a scientist, one needs to conform to scientific method, but one does not necessarily need to be a member of a university faculty. This difference between indispensable and flexible aspects of identity is explained below with reference to form and content features (Laden 2001).

9.3.3 Form and Content Features

Laden (2001, p. 108) puts forward the concepts of form features and content features of identity, which may be helpful distinctions in considering the tension between individual frameworks, societal worldviews and social roles within which individual frameworks take shape. This tension may be understood as that between individual identity and its determination by social categories and social roles.

In my discussion of these form and content features, I stay close to Laden's conceptualisation of them. Form features of identity are always externally dictated, whereas content features can, but need not be, externally dictated. Externally dictated in this context implies: socially constructed in such a way that it cannot be easily altered by an individual. In contrast, content features of identity depend specifically on the interpretation of the individual.

> (…) the content features of an identity might vary widely without the identity itself having changed. They can, in this sense, be contingent features of the identity they characterize. Form features, however, derive from the place a given identity occupies in a more complicated, social system. The form characteristics of an identity are determined by the part it plays, and by what is required of someone playing that part, within a larger social context. (Laden 2001, p. 108)

I think the concept of marriage nicely illustrates the difference between form and content features, namely what it implies in social interactions to be married. Important aspects of form features are, for instance, the legal obligations that exist between husband and wife, such as the obligation to take care of each other. Content features pertain to how a marriage and its accompanying values are understood by a specific couple. Form features are mainly public in character, whereas content features are primarily private in character. However, private aspects of identities are to a substantial degree always necessarily determined by form features, whereas form features are influenced by private aspects of one's identity.

The effects of each of these features on the identity of the individual will vary per situation and per individual. The legal obligation of loyalty between married partners can, for instance, imply a wide range of actions. For some couples it might primarily refer to sexual exclusivity, whereas for others it might imply that they always give each other priority over other relationships such as friends or family.

Both the form and the content features of an identity are eventually subject to change, but less so in the case of the former than the latter, as Laden states (2001, p. 110). Again, I should like to illuminate these features, as I understand them, by referring to the institute of marriage. In many societies, marriage forms a stabilising institution. Subjecting it to constant revision would undermine social coherence. Many aspects of the concept of marriage therefore belong to the form features of identity within a particular society.

However, due to changes in the content features of certain identity constituting concepts, form features may eventually change. In the case of marriage, for instance, the fact that gay marriages are now legal in the Netherlands is to some extent due to changing attitudes towards what marriage means. Whereas people used to conceive of marriage as something with a practical value, a coalition between two parties that would produce offspring, the understanding of marriage has shifted more and more to the expression of a romantic, unique connection between two persons: a symbol of love rather than social custom. Since gay couples cannot produce children, but can form loving attachments, the existence of gay marriage fits the modern understanding of marriage more than the older understanding of marriage. In this case, the form features have changed due to alterations in the content features.

Form features of social roles indicate shared values between individuals. This understanding of form features is compatible with Laden's approach, although he does not explicitly mention the moral aspect. An example of such a shared value is the value of reasonable pluralism, as indicated by Rawls (2005). He states that most people are likely to share a belief about reasonable pluralism that allows them to

engage in a political debate with each other. According to Laden (2001), such a belief should be considered a form feature of citizenship. Comparable beliefs exist for other social roles. For instance, physicians vow to do no harm and a law enforcer is supposed to emanate neutrality with regard to religious convictions.

We attach certain rights and obligations to specific identities, for instance to government officials or legal spouses. We expect government officials to act for the sake of the common good, for instance, and we expect legal spouses to take care of each other. If individuals do not comply with these moral rights and obligations, their identity may not immediately alter in a legal or social sense, but it may be questioned. Government officials who accept extraordinary gifts from third parties or wives who mistreat their husbands have acted contrary to the rights and obligations we usually attach to these identities, which may lead to their being expelled from their social roles.

According to Laden (2001, p. 110), the more public an identity is, the more its articulation is constrained by form features. The identity of citizen is for instance highly constrained by form features. To act out her role as citizen, that is to act in political contexts, an individual needs to take account of the rights and duties that are typically associated with such a role, because she interacts and offers reasons for specific claims on other citizens, that is, these are public reasons that ought to appeal to anyone who is also a citizen. In this context, Laden suggests the form feature of overridingness as associated with the identify of citizen: in public deliberations it should override other, non-political identities.

This implies that in public debates we are not supposed to make claims that are based on a specific, non-public identity, such as being a mother or having a specific ethnic background. Such identities can lend support to claims, but they are not overriding. For instance, an individual states that as a mother she cannot accept any unsustainable technologies, because they would endanger the future of her child. Such support might add to the strength of her argumentation, since by referring to motherhood, she invokes recognisable moral concerns legitimated by the social role of mother and the rights and obligations socially attached to it. However, referring to motherhood is not a compelling argument, because motherhood is not a public but a private identity. A claim supported by a public identity, namely that of citizen, could be: “As a member of this society, I cannot accept unsustainable technologies because they would endanger of the future of our society as a whole.”

When an individual makes claims on others as a mother within the private sphere, these concern primarily her family. For instance, when an individual claims that the time spent watching television should be restricted, or that the family should give some of its household budget to charity, the need to justify the claim she makes is limited solely to that family. The individual can support her claims with reference to her identity as a mother. “As your mother I feel you need to watch less television” or “As the mother in this family, I feel we should donate money to charity”. This appeal does carry a moral weight in the context of the private family.

9.4 RRI, Social Learning and Identity

We have established that identity consists of a moral framework and a collection of social role(s) and that identity is shaped by both form and content features. Including identity in the design of RRI processes has several consequences. First, we can use identity to select participants for social learning exercises, based on either their moral framework or their social role. Second, we can design social learning exercises such that the form features of identity are aligned with the purpose of the exercise, in either a public or a sheltered setting. Third, form features of identity allow us to transfer lessons learned across different contexts, thereby creating an inclusive agenda for a public debate on synthetic biology. Fourth, form features of identity can help to identify which issues are suitable for exercises involving individual actors and which issues require a different approach. I elaborate on these consequences in turn.

9.4.1 Selection of Participants

Considering the social roles as well as the moral frameworks of actors is a way to ensure that a wide range of perspectives are represented in an RRI exercise. It allows for the selection of participants based on the expectation that they will have a *specific perspective* on the matter at hand that can be expected to contribute to the quality of the outcome of a social learning exercise, because such exercises usually benefit from a diversity of perspectives (Cuppen 2012).

In approaches such as technology assessment, which can be considered a predecessor of RRI (Rip 2014), random citizens are sometimes invited, such as in the case of citizens' juries and focus groups. In these cases, the main aim is to broaden the scope of arguments in order to increase the quality of the evaluation of a technology, and to increase the democratic justification of technological developments (Joss and Bellucci 2002). However valid these motivations may be, if the aim of RRI is to ensure that innovations are developed with a view to addressing actual societal concerns and challenges, the associated learning exercises should at least involve those actors who might fuel actual controversies (Bogner 2012) and those who will be using the technology (Schot 2001). Such actors can be found by selecting specific moral frameworks (Cuppen 2012), as well as by taking the aspect of social role into account (Davis and Nathan 2015).

Hence, effective RRI should involve not only private persons whose main concern is their personal valuation of a technology, but also individuals whose evaluation of a technology is mediated or also mediated by a specific social role relevant to the innovation at stake. We may, for instance, invite a physician to give her opinion about developments in industrial biotechnology, but this might not necessarily produce anything useful in the context of RRI. We should rather ask her to share her opinion on medical biotechnology, which relates to her actual practice. In such a

case, she can reflect on an innovation that will actually arrive in the practice that she represents.

This is not a call to only involve experts, because also laypeople can have social roles that make them indispensable for a wide ranging RRI exercise. For instance, a patient suffering from the disease targeted by the biotechnological device can be invited too. It is also not a call to exclude people who do not occupy a social role that is related to the technology at stake, but still have an interest because they have strong opinions about it. They can also be valuable contributors to RRI.

It is a call, however, not to include random individuals, but to consider the identity of participants for effective RRI. Eligibility may arise from a relevant social role or from a strongly voiced moral framework. Such expression of a moral framework may have arisen in relation to another technology. Since RRI is supposed to lead to an assessment of technology before a controversy arises, it may be hard find individuals expressing opinions related to the targeted technology. In such a case, opinions on comparable technologies serve as an indication of relevant opinions.

This gives rise to the question whether participants who are selected because of their social role should stick to that specific role, or whether it is legitimate for them to invoke other elements of their identity. A scientist can, for instance, be invited to present her assessment of the risks attached to a specific synthetic biology application; that is, to partake in the exercise as a scientific expert. The scientist may, however, feel the need to actively advocate the use of that specific application because of her moral assessment of that application, namely that it should be put to use despite the risks. She is then acting more as an activist than a neutral scientist.

It may be considered a form feature of scientists participating in public debates or social learning exercises to limit their input to scientific facts.[1] However, scientists are not solely scientific beings. They may be moved by values other than purely scientific ones, such as wanting to offer help to patients. A scientist may also be acting out her role as a mother, because she has a child who is suffering from a disease that may be cured using this specific technology. Awareness of different aspects of identity can help us distinguish between the different kinds of input an individual may contribute to a social learning exercise debate. In the above example, the scientist can be allowed to voice all her concerns, but she may be asked to explicitly acknowledge which statements follow from her identity as a scientist and which from her identity as a private person, such as a mother.

Distinguishing between different aspects of identity can help individuals reflect on their perspective on a technology: which part of it arises from their social role and which from their particular individual framework. This might in turn lead to an evaluation of that social role when the individual decides she is uncomfortable with its dictates; for instance, when a physician feels as an individual that she should sometimes withhold treatment because a patient feels ready to die. To make such a course of action available to her without her losing her identity as a physician,

[1] Whether such a clear distinction between facts and moral, particular values can indeed be made is a matter of debate. The point is that a scientist should strive to represent those elements of his knowledge that he believes are factual.

requires a public debate resulting in a change in the form features of that specific identity. As such, awareness of the distinction between the public and the private aspects of our identity can result in an agenda for public debate. I return to this point further down.

9.4.2 Design of the Exercise

Our identity affects our willingness to engage in social learning. If we take the notion of identity into consideration in this process, it helps us understand the extent to which we can expect participants in RRI exercises to be "mutually responsive to each other", as propagated by Von Schomberg (2011). Some social roles allow for such reflection and the creation of a shared perspective only under specific circumstances.

We should take into account that the form features of identity are dictated by the social environment. When they do indeed constitute a barrier to effective social learning, we might consider changing the social environment. An exchange of arguments necessary for social learning such as RRI, can take place in public settings. In such settings, identities and related perceptions are usually robust, implying that they are not open to change. Depending on our purposes, we might therefore change the setting to a more private one. The spokesperson for an environmental NGO may, for instance, be more willing to reflect on her moral framework within a closed setting, where she is sheltered from the public eye. This is an almost self-evident way to diminish constraints caused by form features, as they are publicly dictated.

A politician whose ideology compels him to reject subsidies on technologies for renewable energy, will not easily change his mind during a public debate on such technologies. He engages in public debates to put forth his arguments, not necessarily to learn. However, willingness to engage in social learning might be impacted by alternative social settings. A politician may not be willing to change his mind in public, but he may be willing to do so behind closed doors when discussing matters with his fellow party members, as he will have time to reconstruct his perception away from the public eye.

However, it is also possible to make use of the form features of identity in a constructive way by appealing to them. Such an appeal might best be undertaken in a public setting. For instance, a CEO might not be compelled to reflect on his framework solely because he is interacting with someone who thinks differently, such as a representative of an NGO. But since he is expected to safeguard the interests of his company and these interests may be hurt by an NGO's allegations that his products are unsustainable, the CEO is publicly committed to reconsider his assumptions, because the allegations of the NGO might affect his company's commercial prospects. In this case, the form features may have a positive effect on the learning process.

In addition, form features can provide an instrument to scale up learning processes. They can provide a pathway to carry learning effects over to other contexts, and thus not be limited to the specific learning context. It is possible that the learning

effects achieved in a public setting in which the form features of identity are strongly present, can be transported to a context in which other individuals occupy similar social roles. Such an effect can be expected for lessons learned for each specific role. When a policymaker learns about certain strong sentiments among the public about the need for the regulation of risks associated with synthetic biology, such insights will be useful for other policymakers. When a scientist learns that other actors expect him to communicate possible applications of a new technology at a very early stage, this will be relevant to all scientists working on synthetic biology. As such, lessons learned by individuals can be transported to other individuals occupying a similar role. As explicated further down, insights from social learning exercises can also have ramifications for the structures that determine the social roles.

It may also happen that social learning effects within a sheltered setting can be transported to other contexts. Representation has been a notable issue in many social learning exercises. Many learning exercises are necessarily small scale because they require intense interaction between the participants. The question then is how the results of these exercises can be translated to other settings. In a sheltered setting, individuals can be expected to loosen their social identities somewhat, but they will not completely abandon them. Any insights pertaining to their social identities can be shared with others who also carry that identity.

9.4.3 A Socially Encompassing Agenda

When social learning is hindered by an individual's social role, and specifically by the form features of this role, this indicates that reaching a shared perspective requires more than an exchange between individuals, even if these individuals represent larger groups or organisations. Or rather, precisely because these individuals are connected to larger groups and are determined by social structures, an effective RRI process requires more than a meeting of individuals. Even if the individuals in such a meeting arrive at a shared perspective, the social roles they have outside this meeting might prevent them from taking this new shared perspective to another context.

A policymaker participating in an RRI exercise may, for instance, be convinced of the need for stricter regulation of synthetic biology, but may find out later that his colleagues in another department are working on a trade agreement aimed at reducing regulations. The policymaker may, of course, try and argue her case with her colleagues, but it is possible that she will not win the argument because there are other, competing public goods such as economic gain, clashing with her focus on the value of safety. Her social role as a public servant will in such a case trump her personal moral judgements.

Another example is when the CEO of a company states that he is not able to modify the design of his product, because breathing down his neck he has investors who wish to see a swift return on their investment. Even if the CEO is convinced of the need to change his product because of the social concerns of other actors, his obligations as a CEO may prevent him from acting upon this newly gained insight.

Such cases make clear that the involvement of individuals in social learning exercises, such as RRI, may not always be sufficient to achieve a truly socially robust new perspective on an emerging technology. The social structures in which these actors operate also determine whether a new perspective gains wide acceptance. When social learning exercises solely involve individuals and do not target the social structures that determine these social actors, they risk becoming obsolete.

However, social learning exercises involving individuals can help us identify which issues still need to be resolved in the public debate, such as macro-economic issues or issues relating to shared societal values, such as those captured in form features of identity. RRI can provide valuable input into wider questions of a desirable society. What economic incentives do we need to stimulate truly sustainable innovation? How do we want to measure the impact of science? How much intervention from the government do we consider desirable in innovation trajectories? Such questions might not always be answered during individual RRI exercises, but such exercises can and should provide useful insights and input for the public and political debates about the societal embedding of innovation.

9.5 Conclusions

Identity can be conceived of as consisting of a moral framework and a collection of social roles. These elements are intertwined. Identity has an influence on the propensity of individuals for social learning. Taking this influence into account has ramifications for the design of effective RRI exercises.

Taking the issue of identity into account can help structure RRI exercises to enhance their effectiveness. First, it can be a helpful concept in designing effective RRI for synthetic biology. It can provide guidelines for assuring a wide ranging representation by selecting participants who represent various prevalent worldviews in relation to biotechnologies and by selecting participants who occupy relevant social roles. Explicating the role of identity can also help to establish which issues are best addressed in a sheltered exercise and which in a public debate.

Second, explicit consideration of the identity of participants in social learning exercises can help in linking particular learning exercises to wider societal networks and societal structures of meaning, whereby RRI does not remain confined to individual learning, but becomes truly social learning. This is mainly due to the concept of form and content features of identity. Possible conflicts between personal concerns (content features) and form features of one's identity can serve as an indication that a specific issue should be put on the agenda for public debate.

Third, insights relating to the form features of identity, as in social expectations relating to a specific social role, may be transported to other contexts, allowing for a scaling up of learning effects. Being aware of various social roles as part of one's identity can also help participants distinguish between private and public arguments.

Lastly, acknowledging the importance of identity for social learning exercises can also induce us to become more modest about our expectations with regard to RRI and its outcomes. In the field of synthetic biology, value orientations and perceptions of new technologies are often fiercely opposed. Controversy may simply be unavoidable given the dictates of social roles such as some environmental activists and risk-seeking biotechnological entrepreneurs. The ideal that the exchange of viewpoints and mutual responsiveness will lead to some form of societal agreement might sometimes simply be farfetched.

If that is indeed the case, we might instead start looking for workable alternatives to societal consensus. I think that for such alternatives, identity can also prove to be a valuable platform. The issue of identity helps us to distinguish between public and private claims and to co-construct the strength of our appeal with others. What are the claims I should feel compelled by in the public debate? But also, and importantly, which claims can I ignore, which claims do not hold an appeal to me given my social role and my moral framework? What does it imply for my identity when I accept certain claims and not others? Or more constructively: based on my private and my public identities, what reasons do I have to support or instigate a certain innovation trajectory?

The answers to such questions might not necessary lead to societal consensus with regard to innovations, but they might help us create clearer, more coherent and more comprehensive narratives to explain and shape prospective innovation trajectories. Such innovation supporting narratives can also be considered a form of social robustness and hence a desirable outcome of RRI.

References

Asveld, Lotte. 2008. *Respect for autonomy and technological Risk*. PhD Thesis, Delft University Press.

Asveld, Lotte, and Dirk Stemerding. 2017. Social learning in the bioeconomy: The case of Ecover. In *Experimentation beyond the laboratory. New perspectives on technology in society*, ed. Ibo van de Poel, Donna Mehos, and Lotte Asveld. London: Ashgate Publishers.

Appiah, Kwame Anthony. 2010. *The ethics of identity*. Princeton: Princeton University.

Armitage, Derek, Melissa Marschke, and Ryan Plummer. 2008. Adaptive co-management and the paradox of learning. *Global Environmental Change* 18 (1): 86–98. doi:10.1016/j.gloenvcha.2007.07.002.

Blok, Vincent, and Pieter Lemmens. 2015. The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation. In *Responsible innovation 2: Concepts, approaches, and applications*, ed. Bert-Jaap Koops, Ilse Oosterlaken, Henny Romijn, Tsjalling Swierstra, and Jeroen van den Hoven, 19–35. Cham: Springer.

Bogner, Alexander. 2012. The paradox of participation experiments. *Science, Technology & Human Values* 37 (5): 506–527.

Cuppen, Eefje. 2012. Diversity and constructive conflict in stakeholder dialogue: considerations for design and methods. *Policy Sciences* 45 (1): 23–46. doi:10.1007/s11077-011-9141-7.

Davis, Janet, and Lisa P. Nathan. 2015. Value sensitive design: Applications, adaptations, and critiques. In *Handbook of ethics, values, and technological design: Sources, theory, values and*

application domains, ed. Jeroen van den Hoven, E. Pieter Vermaas, and Ibo van de Poel, 11–40. Dordrecht: Springer.
De Vries, Bert J.M., and Arthur C. Petersen. 2009. Conceptualizing sustainable development: An assessment methodology connecting values, knowledge, worldviews and scenarios. *Ecological Economics* 68 (4): 1006–1019. doi:10.1016/j.ecolecon.2008.11.015.
De Witt, Annick, Patricia Osseweijer, and Robin Pierce. 2015. Understanding public perceptions of biotechnology through the "Integrative Worldview Framework". *Public Understanding of Science*: 1–19. doi:10.1177/0963662515592364.
Douglas, Mary, and Aaron Wildavsky. 1982. *Risk and culture: An essay on the selection of technical and environmental dangers*. Berkeley: University of California Press.
Dryzek, John S. 2013. *The politics of the earth: Environmental discourses*. Oxford: Oxford University Press.
Hedlund-De Witt, Annick. 2013. *Worldviews and the transformation to sustainable societies*. PhD thesis, Vrije University.
Hoven, J.V.D., and K. Jacob. 2013. *Options for strengthening responsible research and innovation*. Brussels: European Commission.
Joss, Simon, and Sergio Bellucci. 2002. Participatory technology assessment. In *European perspectives*. London: Center for the Study of Democracy.
Kahan, Dan M. 2012. Cultural cognition as a conception of the cultural theory of risk. In *Handbook of risk theory*, ed. Sabine Roeser, Rafaela Hillerbrand, Per Sandin, and Martin Peterson, 725–759. Dordrecht: Springer.
Kupper, Frank, Linda Krijgsman, Henriette Bout, and Tjard de Cock Buning. 2007. The value lab: Exploring moral frameworks in the deliberation of values in the animal biotechnology debate. *Science and Public Policy* 34 (9): 657–670. doi:10.3152/030234207x264944.
Laden, Anthony Simon. 2001. *Reasonably radical: Deliberative liberalism and the politics of identity*. Cambridge, MA: Cambridge University Press.
Owen, Richard, Phil Macnaghten, and Jack Stilgoe. 2012. Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39: 751–760.
Rawls, John. 2005. *Political liberalism*. New York: Columbia University.
Reed, Mark, Anna Clair Evely, Georgina Cundill, Ioan Raymond Albert Fazey, Jane Glass, Adele Laing, Jens Newig, Brad Parrish, Christina Prell, and Chris Raymond. 2010. What is social learning? *Ecology and Society* 15 (4): r1.
Rip, Arie. 2014. The past and future of RRI. *Life sciences, society and policy* 10 (1): 17.
Röling, Niels. 2002. Beyond the aggregation of individual preferences. In *Wheelbarrows full of frogs: Social learning in rural resource management*, 25–48.
Schot, Johan. 2001. Towards new forms of participatory technology development. *Technology Analysis & Strategic Management* 13 (1): 39–52. doi:10.1080/09537320120040437.
Stilgoe, J., Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580.
Taylor, Charles. 1989. *Sources of the self: The making of the modern identity*. London: Harvard University.
van de Poel, Ibo, and Sjoerd D. Zwart. 2010. Reflective equilibrium in R & D networks. *Science, Technology & Human Values* 35 (2): 174–199. doi:10.1177/0162243909340272.
van de Poel, Ibo. 2016. An ethical framework for evaluating experimental technology. *Science and Engineering Ethics* 22 (3): 667–686.
van Mierlo, Barbara, Cees Leeuwis, Ruud Smits, and Rosalinde Klein Woolthuis. 2010. Learning towards system innovation: Evaluating a systemic instrument. *Technological Forecasting and Social Change* 77 (2): 318–334. doi:10.1016/j.techfore.2009.08.004.
Von Schomberg, Rene. 2011. *Towards responsible research and innovation in the information and communication technologies and security technologies fields*. Luxembourg: Publication Office of the European Union.

Chapter 10
Decision-Making in Water Governance: From Conflicting Interests to Shared Values

Klara Pigmans, Neelke Doorn, Huib Aldewereld, and Virginia Dignum

Abstract The development of water infrastructure is a long and complex process that involves multiple stakeholders, multiple scales, various sub-systems and relations of dependence among stakeholders. Stakeholder participation is increasingly seen as an indispensable element of water policymaking. The failure to address stakeholders' underlying values, however, may create or exacerbate conflicts. In this chapter, we address the difficulty of approaching stakeholder participation in terms of conflicting interests. We illustrate this with an urban flood prevention case, followed by a categorisation of the difficulties presented by such processes. Instead of pursuing an interest-oriented approach, we suggest taking a step back in order to discern the influence of differing conceptions of shared values on multi-stakeholder decision-making processes. The goal of this chapter is to achieve a better understanding of the difficulties entailed in interest-driven decision-making processes in water governance, and how it could be beneficial to pursue a value-sensitive approach in such situations.

10.1 Introduction

In the face of climate change, the growing global population, environmental pollution and resource scarcity, the governance of our complex world has become a major challenge. Water infrastructure is a good example of a system that is becoming more complex, because it is facing climate change and resource allocation whilst simultaneously having to deal with increasing levels of stakeholder participation. To cope with these challenges, such systems need effective and sustainable governance.

Water governance can be described as the governance of a socio-technical system, characterised by a complex technological infrastructure and an interrelated network of independent stakeholders. Given the dynamics and dependencies within this system, both policymakers and scientists advocate stakeholder participation

K. Pigmans (✉) • N. Doorn • H. Aldewereld • V. Dignum
Delft University of Technology, Delft, The Netherlands
e-mail: k.a.m.pigmans@tudelft.nl; n.doorn@tudelft.nl; huib.aldewereld@hu.nl; m.v.dignum@tudelft.nl

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_10

(WMO 2009; Huitema et al. 2009). When policies for developing water infrastructure need to be discussed in order to reach a decision, however, conflicts between stakeholders tend to be the rule rather than the exception.

Scholars have discussed the difficulties presented by participative decision-making in terms of uncertainty and disagreement (Hommes et al. 2009), disillusionment (Reed 2008), and a lack of shared perceptions of the nature of a problem (Pahl-Wostl 2002). Each stakeholder has its own conception of the problems underlying the decision-making process, because each has different interests, whether these relate to profits, safety, or a quiet living environment with nice views. Interests can be defined as "matters people feel they should strive for, for themselves" (Van de Poel and Royakkers 2011) because they are individually involved in these matters. Entrepreneurs have an interest in making a profit, because without profits, their companies will go bankrupt; someone who bought their house in a particular area because it overlooks an expanse of water has an interest in preserving this view; and so forth. These interests can differ so considerably that they may conflict with each other, sometimes resulting in stakeholders refusing to accept a proposed policy. This defensive behaviour can severely delay the decision-making process and thereby threaten the development of infrastructure.

To date, how stakeholder decision-making processes in water governance might be improved has been investigated thoroughly with reference to knowledge-sharing (Pahl-Wostl 2002; Hommes et al. 2009; Kolkman et al. 2005). Ample research has been conducted on how stakeholders disagree on conceptions of the knowledge needed to influence the process (or solve the problem). Yet, differing conceptions of underlying values have received little attention so far.

Values have been defined as "lasting convictions or matters that people feel should be strived for in general, and not just for themselves, to be able to lead a good life or to realize a just society" (Van de Poel and Royakkers 2011). Values have also been described as enduring beliefs (Rokeach 1973) that are universally acknowledged (Schwartz 1994) and as guiding principles (Cheng and Fleischmann 2010). Rokeach (1973) focused on the connection between values and behaviour, making a distinction between terminal values and instrumental values. Schwartz (1994), by contrast, described the relation between values, including power, hedonism and security (Pigmans et al. forthcoming). We use the definition of values as stable, enduring guiding principles of what people generally think is important in life (Cheng and Fleischmann 2010). One example of such a value is "water safety", a concept that is vague and abstract enough to be generally acknowledged (Jacobs 1999). How this value is interpreted in practice, of course, is a different matter altogether.

There is a tendency in the literature to address conceptions of values in rather theoretical terms (Jacobs 1999; Joss and Brownlea 1999). In this research, we want to understand how conceptions of values influence actual decision-making processes, in order to gain more insight into such processes. The value "water safety", for example, could be perceived as a spatial planning challenge or as flood mitigation; these are two conceptions of the same value. By shifting the emphasis to conceptions of shared values, we focus on diverse perceptions of current commonalities,

instead of the current differences that need to be bridged in order to reach commonalities in future.

The aim of this paper is to provide a better understanding of the role and conception of values in the decision-making process and in interactions between stakeholders. In order to give a clear indication of the context, in Sect. 10.2 we discuss the background to participatory decision-making in water governance. In Sect. 10.3, in order to illustrate the complexities and conflicts that can characterise these processes, we describe an example of a multi-stakeholder decision-making process in water governance. Next, in Sect. 10.4, we discuss how categorising and formalising the various social aspects could enable us to understand the influence of conceptions of values in the decision-making process. In order to put our ideas into perspective, we then discuss related work in Sect. 10.5. Finally, we present our conclusions and recommendations for future research in Sect. 10.6.

10.2 Background

Stakeholder participation is increasingly seen as an indispensable element of water governance, both as a means of democratisation (Dryzek 1997; Perhac 1998; Maasen and Weingart 2005) and as a way to improve decision-making (Brunner et al. 2005; Pahl-Wostl 2007; Raadgever et al. 2012). The importance of stakeholder participation is also recognised in several international treaties and agreements, such as the agreements made at the Rio Earth Summit and the Aarhus Convention. In 1992, the Rio Earth Summit included the involvement of the public at all levels of decision-making in the definition of sustainable development (WMO 2009), which has since become an essential part of the "Integrated Water Resources Management" paradigm.

Perhaps the most important step in the juridical recognition of stakeholder participation was the adoption of the Convention on Access to Information, Public Participation in Decision-making and Access to Justice in Environmental Matters by the United Nations Economic Commission for Europe. This Convention, usually referred to as the Aarhus Convention, was signed by 46 states and the European Union (EU). Since 2001, the EU has been applying Aarhus-type principles in its legislation. The Aarhus Convention establishes the right of citizens and organisations to participate in environmental decision-making, including the possibility to comment on proposals for policies and interventions affecting or relating to the environment. This provision is implemented in both the European Water Framework Directive (2000/60/EC; preamble and article 14) and the European Floods Directive (2007/60/EC; article 9.3).

Although it is recognised in all of the above-mentioned treaties, however, it has proved difficult to put participation into practice (Reed 2008). As of yet, participatory approaches have not been systematically included in water governance in most countries, although considerable effort has been made to involve stakeholders in drafting water policy on an experimental basis.

For example, a relatively large number of projects have been initiated by water authorities to ensure the involvement of key stakeholders in the implementation of the European Water Framework Directive and the European Floods Directive (Doorn 2016). Due to their ad hoc nature, the degree of success of these stakeholder-involvement projects remains largely unknown, and will depend on the specific participatory goals in question. On the one hand, stakeholder participation can improve the quality of decision-making by opening up the decision-making process and making better use of information (Huitema et al. 2009). On the other hand, multi-stakeholder decision-making processes seem to be characterised by complexity, uncertainty and disagreement (Kolkman et al. 2005). The outcomes of these stakeholder projects can be disappointing, especially if stakeholders are thought to be representing different interests. When stakeholders negotiate about fixed interests, disagreements can seem insurmountable; these interests can differ so greatly that they can conflict with each other. This sometimes results in stakeholders refusing to accept a proposed policy, which can severely delay the decision-making process.

Previous research suggests that stakeholder involvement processes are more effective if they focus on important values rather than the interests that stakeholders represent (Glenna 2010; Doorn 2016). In some cases, failing to address the values that underlie environmental conflicts can even create and exacerbate conflict, resulting in a stalemate of conflicting vested interests (Rikoon and Goedeke 2000; Wilshusen et al. 2003). Discussing values with stakeholders may facilitate the implementation of water policy as part of an approach to water governance that can account for more than the interests of the best negotiators alone. In turn, this may also lead to more support for the chosen policy solutions at a later stage.

10.3 The Case of Water Governance in the Netherlands

In order to understand the complexities of decision-making processes involving multiple stakeholders, we studied one such process in an urban flood prevention case in the Netherlands. We examined vision and planning documents, process development documents, direct and reported communications between the stakeholders, observations of the interactions between the stakeholders performed by external process managers, memos, interview transcripts in which the concerns of stakeholders were discussed, and evaluation reports.

The Netherlands has 23 water boards. Rather than overlapping with the boundaries of the country's provinces, these are defined by natural watersheds. The water boards are functionally decentralised water authorities with exclusive water management responsibilities. They are responsible for flood management, regional water governance and the purification of waste water in the region. One such water board is Waterschap De Dommel in the South-eastern Netherlands. Centred around the river Dommel, this water board focuses on five themes: dry feet (flood management), clean water, enough water, sustainability and beautiful water. In its

decision-making processes, the water board has found that conflicting interests can slow down the realisation of water governance by 10 or sometimes even 20 years.

In the Netherlands, stakeholder involvement is organised mainly around interests: if one has an interest in the decision-making process, one is allowed to object to certain measures or submit one's point of view on specific plans. The development of the water infrastructure typically involves property or land that is owned by citizens and agricultural companies. Property-owners will therefore be involved in the case of concrete measures that influence them directly. In addition, they generally have the right to compensation if they need to make investments in order to comply with new policies, or even full compensation in case of expropriation. In the case of decentralised regulations or more high-level policies that also influence the property-owners, however, the latter are not taken into account in the decision-making process.

In the case of concrete measures, stakeholders are involved after a specific technical solution has been chosen, but before its practical development has been fine-tuned. This has created a situation in which stakeholders know their rights and might have already defended them in earlier encounters with governmental authorities. The case that we studied can be described as having unfolded in three steps, namely: determining the technical solution based on the expressed need; involving stakeholders; and developing the solution. This section will focus on the first two steps.

10.3.1 From Need to Technical Solution

In this case, the decision-making process started with the need to develop the water infrastructure in order to prevent floods in urban areas. This need was initially described by the province and the water board in long-term regional plans and vision documents, and was determined more specifically in the project planning documents produced by the water board and the municipality. The design of the water infrastructure had to incorporate both technically-advanced water systems and the social dependencies at stake. The water board explored the possibilities for technical solutions, potentially together with external technical partners, in an early phase of the process.

By planning the technical development of the water infrastructure step by step, the relevant stakeholders could be identified. Since watermills were located in the area that needed to be developed, the watermill-owners were stakeholders in the process. The private property-owners who lived in the area were also stakeholders, as were the property developers who were active in the area, large-scale cattle farmers and organic farmers. Finally, the water board, the municipality and the province were also stakeholders.

There was a possibility that the water board or the municipality might have to buy private or commercial properties in order to be able to develop the land as needed. This could necessitate citizens having to sell part or even all of their property,

including their homes. Local, large-scale farmers needed to make adjustments to their farms in order to comply with the new water policy. The amount of compensation that needed to be budgeted depended on the amount of land that had to be bought and on the investments that local companies or farmers needed to make. The budget for compensation potentially gave an indication of the degree of social complexity of the process.

The water board and the municipality had to agree on one of the proposed technical solutions, which needed to be developed in alignment with the networks of independent stakeholders.

10.3.2 From Technical Solution to Stakeholder Involvement

Involving the stakeholders was a step characterised by social complexities on many levels, such as agreeing on procedures, bridging differing organisational cultures, conflicting interests, alignment of knowledge-sharing, and means and expectations of communication. In this step, the roles of the stakeholders became known. Each of the stakeholders had its own goals, whether these were professional or personal: from developing wetlands to preventing floods, increasing profits, or living quietly by the water. As soon as the stakeholders became part of the decision-making process, these goals became interests they wanted to secure and, if necessary, defend. This defensive attitude could become particularly apparent when private or commercial interests seemed to conflict with the interests of the governmental authorities.

Certain stakeholders seemed to share certain interests: farmers, for instance, shared an interest in the continuity of their business activities. However, their strategies could differ to such an extent that their interests entailed conflicting objectives, such as the interests of small-scale ecological farming versus large-scale cattle farming. At first glance, the governmental authorities also appeared to share the same interests. If we take a closer look at the societal roles of the governmental organisations, however, we see that their interests were not that similar: the water board had an interest in there being sufficient clean water, whereas the province focused on spatial planning, the economy and culture.

When stakeholders became defensive about their own interests, the governmental focus on formal communication only, in the form of written exchanges or informative meetings for all stakeholders, could be interpreted as an attempt to withhold information from the stakeholders. Informal discussions at the homes of property-owners were highly valued by the latter, as the time and effort taken to listen to them created a feeling of trust. Yet, due to the thoroughness of this form of communication, these discussions took a lot of time to prepare and hold, and were therefore not always used as a means of communication. Moreover, entrepreneurs who did not wish to speak to the authorities because of earlier disputes, but who represented a local political party, used their influence in local politics to lobby indirectly. The use of both indirect and exclusively formal approaches could increase the social distance

between the stakeholders. Different but unspoken expectations, governmental authorities pursuing new ways of working, fragmented responsibilities and a lack of experience in collaborative decision-making could also undermine the effectiveness of the process. A different perspective could lead to new insights about how processes marked by such complexity could still result in socially acceptable outcomes.

10.4 Representing the Influence of Conceptions of Values

The case described in Sect. 10.3 shows that focusing on interests can result in a negotiation process. In this research, we suggest that negotiation alone is not the ideal way to reach a decision, since this implies compromise and ultimately reaching a solution that may not take these interests satisfactorily into account. In order to take a different approach, we first want to gain a better understanding of what initially happens in such processes.

The conflicts in the urban flood protection case were investigated and evaluated by the stakeholders involved, resulting in a list of recommendations on this specific case. Despite this, the evaluation did not produce tenable insights that have the potential to prevent similar difficulties arising in future. If we were to view the process more abstractly, however, we might understand how social aspects influence the process. In order to do so, we need to model the social interactions in this case. This will provide an opportunity to understand how social aspects influence one another and the process as a whole.

In addition to a thorough understanding of these aspects, we want to understand what happens when the focus on individual interests shifts to a focus on shared values. Each of the stakeholders can and will have a different conception of these values. If we consider the value "water safety", for example, many citizens would consider the risk of a flood once every 500 years a very abstract possibility and therefore acceptable, whereas a water board would insist on taking measures for the case that such a threat were to materialise. Nevertheless, we assume that if the stakeholders were to agree on the importance of the overarching value, then the starting point of the process would be agreement on this value, instead of a search for how best to defend one's own interests. One way of exploring this assumption is by modelling it.

In order to represent the complex social phenomena that occur in decision-making involving multiple stakeholders, we need to select a modelling framework that has sufficient modelling capability to represent all of the important aspects of the problem. For this, we can refer to the agent organisation approach taken by Dignum and Padget (2013) to model the interactions between stakeholders with and within the organisational structures of which they form a part. We also refer to the framework used by Ghorbani et al. (2013) to model individuals and institutions as the key components for capturing, analysing and understanding the domain and its complexities. We aim to build upon their research by making values and conceptions of values the major component relating to the social structure, and thereby

searching for common ground rather than differences in interests. In this subsection, we therefore explore the characteristics that are required to model stakeholder participation, the influence of values and conceptions of these values.

In themselves, values do not provide a sufficient picture to understand why particular social phenomena occur in multi-stakeholder decision-making. For a complete picture, we need to consider the influence, direct or indirect, of conceptions of values on other aspects of the decision-making process. The value-related elements that shape multi-stakeholder decision-making can be divided into four (interrelated) categories, based on the urban flood prevention case: *actors and relations*, *processes*, *communication*, and *norms* (or institutions). The concept of an *actor* is required in order to model the objectives and interests of the different stakeholders that are taking part; the *relations* among the actors influence the way in which the process is unfolding; the *process* describes the interactions between the actors as they come to an agreement; *communication* signifies the way in which the interactions take place; and the *norms* influence the way actors behave/interact. We will discuss each concept in detail below.

The *actors and relations* signify the stakeholder map; who the stakeholders are; what their objectives are, both individually and with respect to the decision-making process; and their relations with other stakeholders, such as relations of dependency or power relations. By its very nature, multi-stakeholder decision-making is a multi-actor problem: different parties are involved, each with its own views, objectives, and so forth. There might be differences in aggregation level for each of the participants: some might be individuals, such as a watermill-owner; some might act individually, but belong to a group, such as farmers; some form part of, or represent, a formal organisation, such as municipalities. There is a need to represent each of these different aggregation levels. More importantly, we need a way to draw links between a particular actor's different aggregation levels. For instance, to understand the position of "the group of farmers", one needs to understand the values and objectives that they all share. In addition, one should also be able to see how parts of that group – for example, small-scale ecological farmers vs. large-scale cattle farmers – or even individuals, such as those located close to the affected area vs. those located further away, affect that position.

The *processes* describe the way in which the interactions between the actors occur. These processes can be flexible or highly formalised – that is, strict and transparent – procedures. This also includes the frequency of interactions and the means of interaction used. The frequency and means used can contribute to the transparency and openness of the process, which, in turn, can affect the confidence of the stakeholders involved in the process and the authorities. There may be multiple procedures available to achieve similar goals, and, depending on the actors, it may be possible to choose a particular procedure. If the procedures allow significant leeway, the actors participating in the procedure determine – together, though not always in a transparent manner – how the procedure is executed. The case described above involved two governmental organisations with rather different processes: one organisation had a culture of acting pragmatically and rapidly, whereas the other

organisation had a more bureaucratic culture, where acquiring written permission within the hierarchy continued to play an important role in the process.

Communication refers to the ways in which interactions can take place; the kind of languages spoken (formal/informal), the vocabulary used and its performative function, in the form of speech acts such as locutions, illocutions and perlocutions (see Austin 1975; Searle 1969). To understand what is happening in the process and how that corresponds to the objectives and values of the actors, one may need to understand the speech acts used, such as assertives, directives and declaratives. Without an understanding of their performative function, utterances in the process cannot be linked correctly to their impact on the decision-making process. Another formal aspect of communication that could be of interest is the way the arguments put forward by the stakeholders are constructed and influence each other. To understand the dialogue between the stakeholders, we need to examine the rebuttal and undercutting status of the arguments used (see, e.g., Dung 1995). This formal representation of the language used also forms part of the communication aspect of the representation. If we apply this to the case, we need to take the communication and arguments that were used by the stakeholders into account in order to understand their stances. These interactions can be described with reference to their letters, emails, documented discussions and arguments.

Lastly, the processes are influenced by *norms*. Norms regulate the behaviour of stakeholders by describing the actions they must (or must not) perform in specific situations (Da Silva Figueiredo and Torres da Silva 2013). This includes, but is not limited to, rules of behaviour pertaining to interaction (politeness and clarity, for example; often classified as *social norms*), but also formalised rules such as organisational policies and laws (also classified as *legal norms*). These norms determine, or rather identify, why particular parties are behaving in a particular way. For instance, the fact that a municipality blocks a particular decision might have nothing to do with its interests/preferences, but rather with the enforcement of legal norms. As we shall explain further in Sect. 10.5, there is a clear link between norms and values. In the case above, norms can be identified by explicitly discussing values, norms and objectives with the stakeholders.

The requirements for modelling stakeholder participation can be summarised as follows:

- Identify and specify underlying values.
- Identify and specify stakeholders, including:
  - Stakeholder conception(s) of values;
  - Individual stakeholder objectives;
  - Relation(s) to global objectives;
  - Internal stakeholder structure (if stakeholder is a group or organisation);
  - Relation to other stakeholders.
- Identify and specify processes and procedures, including:
  - Ability to identify expectations of the process;
  - Ability to model the actual process;
  - Ability to model conflicts between expected and actual processes.

- Identify and specify communication, including:
  - Language(s) used (formal/informal);
  - Performative function of utterances;
  - Relation(s) between utterances (e.g., rebuttal and undercutting).
- Identify and specify norms, including:
  - The relation to the values;
  - The influence of the norms on the process;
  - Detection of conflicts between norms of various sources;
  - Detection of conflicts between norms and the fulfilment of (individual or global) objectives.

Given the complexity of the decision-making processes, this social emphasis is needed in order to unravel the dependencies and interactions between the actors, so as to understand how they influence the process and the outcomes. This necessitates the extension of existing frameworks, as none of the current frameworks that focus on the social aspects of stakeholder interaction (such as OperA (Dignum and Padget 2013), MAIA (Ghorbani et al. 2013), OMNI (Dignum et al. 2004)) covers values, value conceptions, the translation from values to norms, and the relations between them.

10.5 Related Work

To date, how participatory decision-making processes in water governance might be improved has been investigated thoroughly with a focus on knowledge-sharing. The problem of differing conceptions of the required knowledge is broadly acknowledged, and ample research has been conducted on how stakeholders disagree on the conceptions of the knowledge needed to influence the process or to solve the problem. Yet, differing conceptions of underlying values have thus far received little attention. Even when conceptions of values are addressed, this tends to be in rather theoretical terms, focused on a specific value.

Pahl-Wostl (2002) describes an integrated assessment methodology to support complex decision-making processes. She argues that there are different perceptions of the kind of knowledge required, depending on the stakeholders' perspectives. The methodology creates a knowledge base for collecting together all of the scientific disciplines that are perceived as relevant, in order to integrate social learning into the decision-making process.

Kolkman et al. (2005) claim that these decision-making processes typically contain values or assumptions of which stakeholders are unaware, or that remain hidden because they were not openly communicated during the process. They advocate a technique for eliciting these values or assumptions, whereby they map the mental models of the stakeholders. These mental models are then used to stimulate the learning process during decision-making.

Hommes et al. (2009) elaborate on the different perceptions of the central problem by describing knowledge gaps between the stakeholders as the cause of these differences. They bridge stakeholders' perceptions by initiating a knowledge-creation process.

Our research builds strongly on the research that has identified differing conceptions of the central problem that is addressed during the process. If social interactions are disrupted too much, however, as happened in the case described in the previous section, stakeholders no longer care about knowledge or knowledge gaps anymore (see Sect. 10.3). In such cases, it might be better to start from an approach that steps away from current social relations and searches for common ground on the more abstract level of values (Glenna 2010; Doorn 2016). For this reason, we do not focus on conceptions of knowledge or on the active creation of a knowledge-base as part of the process; instead, we take a step back from the process and switch the focus from conflicting interests to shared values.

Apart from the focus on knowledge-sharing as a means of improving water governance, previous studies have discussed actors' perceptions and their role in the decision-making process from a procedural justice perspective in more general terms. Rawls (1995) argues that stakeholders should strive for *consensus*, based on the existing overlap between the different conceptions, to arrive at a generally-supported conception of the value at stake. Habermas (1995), by contrast, argues that stakeholders should strive for *mutual understanding* of moral points of view by discussing different perceptions openly and freely. Joss and Brownlea (1999) also take this procedural justice perspective, but with an emphasis on the perceived fairness of the process. In general, the differing perceptions of fairness are explained in terms of the degree to which stakeholders have control over the process. Joss and Brownlea suggest that this focus on control should shift to a focus on the type of relations between the stakeholders. More specifically, they describe a focus on the perceived neutrality of the decision-making party, trust in the decision-making party's motives, and the decision-making party's respect for stakeholders' rights. In order to make this shift in focus, they advocate having a true understanding of different conceptions of the process and of their influence on the process. This understanding will then be captured in the resulting socially-accepted decisions. Yet, aside from an admission that it might be difficult, the authors do not specify *how* this understanding should be realised.

Finally, our approach is inspired by the *value hierarchy* described by Van de Poel (2013). He argues that the relations between values, norms and design requirements can be seen as a value hierarchy, with values at the top, norms in the middle and design requirements at the bottom of the pyramid. The formulation of the norms and design requirements depends on the specific context of the values at stake. He describes environmental sustainability as a value, for example, with low ammoniac emissions as a norm and a reduction in litter as a design requirement. This hierarchical relation is applicable in the context of poultry husbandry systems, which he uses as an example. In the case of water governance, other norms and therefore other design requirements would be derived from the value "environmental sustainability". Furthermore, moving through the hierarchy from top to bottom, the relation is

a specification; that is, norms are described as specifications of a value. Moving through the hierarchy from the bottom to the top, the relation is 'for the sake of': a design requirement is in place for the sake of a norm, and a norm is in place for the sake of a value.

In our research, we are not necessarily focusing on design or design requirements. However, the translation of values into less abstract concepts and vice versa provides a new and relevant perspective on values.

10.6 Conclusions and Recommendations for Future Research

Stakeholder participation is increasingly seen as an indispensable element of water policymaking and has been enshrined in law since the Aarhus Convention. Actually putting stakeholder participation into practice, however, has turned out to be challenging. Taking values into account in multi-stakeholder decision-making processes may facilitate the implementation of water policies that represent more than just the interests of the best negotiators. This, in turn, may lead to more support for the chosen policy solutions at a later stage.

In cases such as that presented in Sect. 10.3, where social interactions are disrupted to the extent that stakeholders no longer care about knowledge sharing, focusing on the interests or knowledge of stakeholders is a less successful approach. These kinds of cases occur more commonly than one would expect, and they show that there is a need to find common ground between the stakeholders (on a more abstract level). Focusing on underlying values may offer one means of returning to constructive interaction between the stakeholders, as opposed to defensive behaviour. The inclusion of values is not trivial, however, especially given that people may have different conceptions of the same value.

In this chapter, we presented a motivating case for the use of values in participatory decision-making and described the elements needed to model this social phenomenon, in order to understand the relation between values and the resulting decision-making process. In addition to values, elements of interest include the stakeholders, their objectives, the relations between the stakeholders, the processes used for interaction, the language used, and the norms governing the process.

So far, we have only presented a rough sketch of the elements needed to model and understand a value-sensitive decision-making process. We could draw an analogy with the modelling requirements and frameworks used by agent organisations. However, further research is required to show whether such frameworks would indeed be suitable, and what kind of extensions would be required.

In order to gain a better understanding of multi-stakeholder decision-making processes and to go beyond the Dutch case, other cases will also need to be analysed. Ideally, these should be cases in developing countries, where the conceptualisation of values may differ significantly from the Dutch situation. While the circumstances

in such countries are very different, the multi-stakeholder decision-making processes are just as complex in terms of social interaction, with multiple scales, sub-systems and dependencies between the stakeholders.

With this broader understanding of these processes, our aim is to explore how to extend the list of requirements for modelling multi-stakeholder decision processes, and to gain deeper insight into the influence of values and conceptions of values in these processes.

Acknowledgements We want to thank the reviewers for their constructive and critical feedback, which helped us to sharpen the focus of this chapter. Moreover, we would like to thank Waterschap De Dommel for the case material. This work forms part of the Values4Water project, subsidised by the *Responsible Innovation* research programme, which is partly financed by the Netherlands Organisation for Scientific Research (NWO) under Grant Number 313-99-316. The work of Neelke Doorn is supported by the NWO under Grant Number 016-144-071.

References

Austin, John. 1975. *How to do things with words*. Oxford: Oxford University Press.

Brunner, Ronald, Toddi Steelman, Lindy Coe-Juell, Christina Cromley, Christine Edwards, and Donna Tucker. 2005. *Adaptive governance: Integrating science, policy and decision-making*. New York: Columbia University Press.

Cheng, An-Shou, and Kenneth Fleischmann. 2010. Developing a meta-inventory of human values. *Proceedings of the American Society for Information Science and Technology* 47(1): 1–10.

Da Silva Figueiredo, Karen, and Viviane Torres da Silva. 2013. Identifying conflicts between norms and values. In *Coordination, organizations, institutions, and norms in agent systems IX*. Cham: Springer.

Dignum, Virginia, and Julian Padget. 2013. Multiagent organizations. In *Multiagent systems*, ed. Gerhard Weiss. Cambridge, MA: MIT Press.

Dignum, Virginia, Javier Vázquez-Salceda, and Frank Dignum. 2004. OMNI: Introducing social structure, norms and ontologies into agent organizations. In *Programming multi-agent systems*, ed. Rafael H. Bordini, Mehdi Dastani, and Amal El Fallah Seghrouchni, 181–198. Heidelberg: Springer.

Doorn, Neelke. 2016. Governance experiments in water management: From interests to building blocks. *Science and Engineering Ethics* 22 (3): 755–774. doi:10.1007/s11948-015-9627-3.

Dryzek, John. 1997. *The politics of the earth: Environmental discourses*. Oxford: Oxford University Press.

Dung, Phan Minh. 1995. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence* 77 (2): 321–357.

Ghorbani, Amineh, Pieter Bots, Virginia Dignum, and Gerard Dijkema. 2013. MAIA: A framework for developing agent-based social simulations. *Journal of Artificial Societies and Social Simulation (JASSS)* 16 (2): 1–19.

Glenna, Leland. 2010. Value-laden technocratic management and environmental conflicts: The case of the New York city watershed controversy. *Science, Technology & Human Values* 35 (1): 81–112.

Habermas, Jurgen. 1995. Reconciliation through the public use of reason: Remarks on John Rawls's political liberalism. *The Journal of Philosophy* 92 (3): 109–131.

Hommes, Saskia, Joanne Vinke-de Kruijf, Henriette Otter, and Geiske Bouma. 2009. Knowledge and perceptions in participatory policy processes: Lessons from the delta-region in the Netherlands. *Water Resources Management* 23: 1641–1663.

Huitema, Dave, Erik Mostert, Wouter Egas, Sabine Moellenkamp, Claudia Pahl-Wostl, and Resul Yalcin. 2009. Adaptive water governance: Assessing the institutional prescriptions of adaptive (co-)management from a governance perspective and defining a research agenda. *Ecology and Society* 14 (1): 26.

Jacobs, Michael. 1999. Sustainable development as a contested concept. In *Fairness and futurity*, ed. Andrew Dobson. Oxford: Oxford University Press.

Joss, Simon, and Arthur Brownlea. 1999. Procedural justice: Considering the concept of procedural justice for public policy- and decision making in science and technology. *Science and Public Policy* 26 (5): 321–330.

Kolkman, Marinus, Matthijs Kok, and Anne van der Veen. 2005. Mental model mapping as a new tool to analyse the use of information in decision-making in integrated water management. *Physics and Chemistry of the Earth* 30: 317–332.

Maasen, Sabine, and Peter Weingart. 2005. *Democratization of expertise? Exploring novel forms of scientific advice in political decision-making*. Dordrecht: Springer.

Pahl-Wostl, Claudia. 2002. Participative and stakeholder-based policy design, evaluation and modeling processes. *Integrated Assessment* 3 (1): 3–14.

Pahl-Wostl, Claudia. 2007. Transitions towards adaptive management of water facing climate and global change. *Water Resources Management* 21 (1): 49–62.

Perhac, Ralph. 1998. Comparative risk assessment: Where does the public fit in? *ScienceTechnology & Human Values* 23 (2): 221–241.

Pigmans, Klara, Huib Aldewereld, Virginia Dignum, and Neelke Doorn. Forthcoming. The role of values. In *Coordination, organizations, institutions and norms in agent systems XII*, ed. Stephen Cranefield, Samhar Mahmoud, Julian Padget, and Andre Rocha. Springer.

Raadgever, G., Erik Mostert, and Nick van de Giesen. 2012. Learning from collaborative research in water management practice. *Water Resources Management* 26 (11): 3251–3266.

Rawls, John. 1995. Political liberalism: Reply to Habermas. *The Journal of Philosophy* 92 (3): 132–180.

Reed, Mark. 2008. Stakeholder participation for environmental management: A literature review. *Biological Conservation* 141: 2417–2431.

Rikoon, J., and Theresa Goedeke. 2000. *Anti-environmentalism and citizen opposition to the Ozark man and the biosphere reserve*. Lewiston: Edwin Mellen Press.

Rokeach, Milton. 1973. *The nature of human values*. New York: Free Press.

Schwartz, Shalom. 1994. Are there universal aspects in the structure and contents of human values? *Journal of Social Issues* 50 (4): 19–45.

Searle, John. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Van de Poel, Ibo. 2013. Translating values into design requirements. In *Philosophy and engineering: Reflections on practice, principles and process*, 253–266. Dordrecht: Springer.

Van de Poel, Ibo, and Lambèr Royakkers. 2011. *Ethics, technology, and engineering*. Malden: Wiley-Blackwell.

Wilshusen, Peter, Steven Brechin, Chrystal Fortwangler, and Patrick West. 2003. Contested nature: Conservation and development at the turn of the twenty-first century. In *Contested nature: Promoting international biodiversity with social justice in the twenty-first century*, ed. Steven Brechin, Peter Wilshusen, Chrystal Fortwangler, and Patrick West, 1–12. Albany: State University of New York Press.

WMO. 2009. *Integrated flood management: Concept paper*. Geneva: World Meteorological Organization.

Part III
RI in the Business Context

Chapter 11
A Framework for Responsible Innovation in the Business Context: Lessons from Responsible-, Social- and Sustainable Innovation

Rob Lubberink, Vincent Blok, Johan van Ophem, and Onno Omta

Abstract While the concept of Responsible Innovation is increasingly common among researchers and policy makers, it is still unknown what it means in a business context. This study aims to identify which aspects of Responsible Innovation are conceptually similar and dissimilar from social- and sustainable innovation. Our conceptual analysis is based on literature reviews of responsible-, social-, and sustainable innovation. The insights obtained are used for conceptualising Responsible Innovation in a business context. The main conclusion is that Responsible Innovation differs from social- and sustainable innovation as it: (1) also considers possible detrimental implications of innovation, (2) includes a mechanism for responding to uncertainties associated with innovation and (3) achieves a democratic governance of the innovation. However, achieving the latter will not be realistic in a business context. The results of this study are relevant for researchers, managers and policy makers who are interested in responsible innovation in the business context.

R. Lubberink (✉) • O. Omta
Department of Management Studies, Wageningen University, Wageningen, The Netherlands
e-mail: rob.lubberink@wur.nl

V. Blok
Management Studies Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands

Philosophy Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands

J. van Ophem
Department of Economics of Consumers and Households, Wageningen University, Wageningen, The Netherlands

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_11

11.1 Introduction

The European Commission wants to accelerate innovation and technological development to address the 'Grand Challenges' of our time, such as global warming, ageing populations and resource scarcities. They state that "*Europe's future is connected to its power to innovate. The Innovation Union, an action-packed initiative for an innovation-friendly Europe, is the solution*" (European Commission 2013, p. 2).

Although technology and innovation have a positive connotation, one can question whether they are inherently good (Von Schomberg 2013). Innovations can have short-term advantages but also come with uncertainties, questions and dilemmas regarding the future impacts and consequences (Stilgoe et al. 2013). The combustion engine for instance is nowadays essential for transportation but also one of the main causes of CO2 emissions. Likewise the effective insecticide DDT turned out to be very harmful to the environment as well.

Responsible Innovation is an emerging concept that aims to prevent or deal with problems that arise with innovation. This is done by taking social and ethical aspects into account and by balancing economic, socio-cultural and environmental aspects (Blok and Lemmens 2015). Burget et al. (2017) state that "*Responsible Innovation is essentially an attempt to govern research and innovation in order to include all the stakeholders and the public in the early stages of research and development. The inclusion of different actors and the public is, in turn, meant to increase the possibilities to anticipate and discern how research and innovation can or may benefit society as well as prevent any negative consequences from happening*" (p. 15).

Responsible Innovation borrows processes and tools from work in Bioethics, Technology Assessment and Ethical, Legal and Social Aspects (ELSA) (Burget et al. 2017). These approaches do not study the whole spectrum of purposes, processes, products and implications of the innovation, but they primarily investigate the research stage while often overlooking the important final stages of innovation, such as commercialisation. The added value of Responsible Innovation in comparison to ELSA is that it focuses on economic valorisation, industry collaboration and socio-economic benefits (Zwart et al. 2014). Van den Hove et al. (2012) argue that Responsible Innovation goes beyond creating just economic growth, as it aims at benefitting people by meeting their needs and by providing economic, environmental and social sustainability.

The concept of Responsible Innovation in a business context faces three major challenges. First, Responsible Innovation lacks definition and clarity. It is a 'big word' that gives some direction but its contents are flexible and open (Bos et al. 2014). Correspondingly, the boundaries between the different underlying dimensions of the Responsible Innovation framework are blurred (Owen et al. 2013). Second, empirical research in the field of Responsible Innovation is lacking (Blok et al. 2015). This is because this field of research is relatively new, and was introduced in a top-down manner by policy makers (Burget et al. 2017), and is defined and understood in different ways (Bos et al. 2014; Burget et al. 2017). Third, Responsible Innovation has a narrow view on innovation as it focuses on science (Lettice et al. 2013) and technological development (Ribeiro et al. 2016) and fails to include commercialisation (Pellé and Reber 2014). This is remarkable because

commercialisation is an essential stage of an innovation process and also, most innovations take place in the private sector (Baregheh et al. 2009). Consequently, it is still unknown what the concept of Responsible Innovation entails in the business context (Blok and Lemmens 2015).

We suggest that previous work on social innovation and sustainable innovation is used to advance the concept of Responsible Innovation in the business context. One reason is that social- and sustainable innovation are already embedded in the business context. Social innovation research has been more practice-oriented and predominantly studied in the context of entrepreneurship (Choi and Majumdar 2014), while corporate sustainable innovation has already received considerable attention from researchers, managers, and policy makers (Adams et al. 2016). Second, we argue that social- and sustainable innovation are conceptually overlapping with Responsible Innovation, since each of these three innovation approaches is considered to involve innovations *for* society and *with* society.

In this chapter we analyse where the current concept of Responsible Innovation shares conceptual similarities and dissimilarities with social innovation and sustainable innovation with regard to: the *inputs* for innovation, the innovation *processes*, and the subsequent *outputs* and implications of these innovations for society. At the conclusion of this study we synthesize the results and lay the basis for the concept of Responsible Innovation in the business context. Our aim is to inspire future research on Responsible Innovation in the business context by shifting the discussion from responsible science towards Responsible Innovation. Consequently, three research questions need to be answered:

In what way is Responsible Innovation conceptually overlapping with social- and sustainable innovation in regard to purpose, process, products and implications of the innovation?

In what way is Responsible Innovation conceptually distinctive from social- and sustainable innovation in regard to purpose, process, products and implications of the innovation?

What do these conceptual similarities and dissimilarities mean for our understanding of Responsible Innovation in the business context?

Since social- and sustainable innovation are defined in different ways by different streams of researchers, we argue that our proposed concept of Responsible Innovation should not be based on just a limited set of definitions. We expect that literature reviews of responsible-, social- and sustainable innovation research provide better insights of the different perspectives on each of these concepts. Therefore, this chapter contains a conceptual analysis of literature reviews and does not involve a meta-analysis or empirical research.

The remainder of this chapter is structured as follows. In the Literature Review, the concepts of responsible-, social- and sustainable innovation are explained with information from review articles. First, the concept of Responsible Innovation is explained, which is followed by a section where the concept of social innovation is explained. Subsequently, the conceptual similarities and dissimilarities between responsible- and social innovation are presented. The same structure is followed for sustainable innovation. In the final section we will integrate these findings and develop our understanding of Responsible Innovation in the business context.

11.2 Responsible Innovation

11.2.1 Input of Responsible Innovation

Responsible Innovation is a new and upcoming concept triggered by the call for innovations that respond to the grand challenges of our time (Von Schomberg 2014) such as climate change, food security and poverty. The innovation that is necessary for finding solutions comes with uncertainties regarding their development and their future implications (Stilgoe et al. 2013). These complex challenges or 'wicked problems' can be seen as inputs for Responsible Innovation (Blok and Lemmens 2015).

The future implications of innovations cannot always be predicted during the development of the innovation. Responsible Innovation acknowledges this inherent uncertainty and it aims to achieve governance of the innovation to accommodate the uncertainty of future implications (Stilgoe et al. 2013). Other reasons to initiate Responsible Innovation can be due to public policy demands, to increase the odds of public acceptance, to better foresee possible implications, to deliver societal benefits and to develop better novel practices (Ribeiro et al. 2016).

11.2.2 Throughput of Responsible Innovation

Owen et al. (2012) and Stilgoe et al. (2013) developed a more democratic governance framework for innovation that is based on contemplating the purpose(s) of the innovation instead of focusing on avoiding detrimental implications (Ribeiro et al. 2016). More specifically, stakeholders and members of the public are involved early in the innovation process to deliberate about the innovation at stake, which helps innovators to think carefully about the purpose of the innovation. Furthermore, the deliberation should involve discussions on how the development of the innovation can be responsive to the inherent uncertainties that come with innovation. Hence, their anticipatory governance of innovation is based on a collective duty of care that requires alternative constructions of (co-)responsibility (ibid.).

Von Schomberg (2012) has a similar focus on a democratic governance of innovation and defines the process Responsible Innovation (i.e. the throughput) as:

> … a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to 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)"(Von Schomberg 2012, p. 9).

It is widely acknowledged that there are several conceptualisations and definitions of responsible (research and) innovation (e.g. Burget et al. 2017; Gianni and Goujon 2014; Wickson and Carew 2014). Accordingly, there are multiple approaches developed for Responsible Innovation, for example approaches that focus on

evaluation of the benefits, impacts, unanticipated risks and ethical implications of the innovation (e.g. Technology Assessment). However, the framework developed by Owen et al. (2012) and Stilgoe et al. (2013) is one of the most dominant approaches in Responsible Innovation (Ribeiro et al. 2016). Furthermore, the systematic literature review by Burget et al. (2017) identified four dimensions that are recurring throughout the literature on Responsible Innovation. These are the same four dimensions that comprise the framework for Responsible Innovation developed by Owen et al. (2012) and Stilgoe et al. (2013): anticipation, reflexivity, inclusion, and responsiveness. These four dimensions are further discussed as they are considered to be key for the throughput of Responsible Innovation.

Anticipation involves system thinking about any known, likely, plausible and possible implications of the innovation that is to be developed (Stilgoe et al. 2013). It plays an essential role in the beginning of the innovation, and requires that the actors involved in the innovation understand the dynamics that help to shape the innovation (Burget et al. 2017). Furthermore, the complexities and uncertainties that come with innovation are acknowledged and explicitly taken into account (Stilgoe et al. 2013). Therefore, the 'imaginations' of future implications do not serve to predict futures, but to envision desirable futures and organise resources to meet those desirable futures. The challenge here is to make certain imaginations more concrete while at the same time being receptive for other views. This needs to be done at a time when it can be constructive, but not too late to adjust the innovation (ibid.). This requires early inclusion of stakeholders and the wider public who engage in *"a dedicated attempt to anticipate potential problems and assess available alternatives"* (Wickson and Carew 2014, p. 2).

Reflexivity is about critically scrutinising one's own activities, commitments and assumptions, and being aware of the limits of knowledge and the fact that one's reality might not be universally held (Stilgoe et al. 2013). Innovators need to reflect on their value systems and theories and how these affect the development of the innovation. Furthermore, innovators need to blur the lines between their role responsibility and their wider moral responsibilities (ibid.). Wickson and Carew (2014) found that reflecting on underlying values, assumptions and beliefs, was a recurring theme in the different conceptualisations of Responsible Innovation, which can be enhanced by early inclusion of stakeholders and the public.

Inclusion is the dimension that comes back in all articles on Responsible Innovation as it is vital for proper implementation of the other three dimensions (Burget et al. 2017). Inclusion is the actual involvement of stakeholders and the wider public via dialogue or other ways to enhance the democratic governance of innovation. Aspects of Inclusion are intensity, openness, and quality of the discussion. Actors have to initiate discussions and to question the social, political and ethical implications of the innovation (Stilgoe et al. 2013). One could say that Responsible Innovation involves an *"active engagement of stakeholders for the purpose of substantively better decision-making and mutual learning"* (Wickson and Carew 2014, p. 2).

Responsiveness is having the capacity to change shape or direction in response to values of stakeholders, values of the wider public and changing circumstances.

Furthermore, it is about actually adjusting courses of action while recognising the insufficiency of knowledge and control, and responding to new knowledge, perspectives, views and norms that emerge when innovating. This in turn requires a collective institutionalised response and co-responsibility for responsible development of the innovation (Owen et al. 2013). Or as Wickson and Carew (2014, p. 2) put it: *"a willingness among all participants to act and adapt according to these ideas"*.

11.2.3 Output of Responsible Innovation

When it comes to the output of Responsible Innovation, we have to consider the actual products of the innovation process and their implications for society. It is clear from the reviews (Burget et al. 2017; Ribeiro et al. 2016) that the output of Responsible Innovation processes are primarily considering science and technological development. However, Blok and Lemmens (2015) suggest that we should widen our conception of innovation and include non-technological innovation as well, such as social innovations.

The overall goal embedded in the different conceptualisations of Responsible Innovation is to take social and ethical aspects into consideration with regard to the development of the innovations (Ribeiro et al. 2016) and its marketable products (von Schomberg 2012). When it comes to the impacts of innovations, there are two approaches to determine whether the impact of an innovation can be considered 'responsible'. According to the procedural approach (e.g. Stilgoe et al. 2013), the stakeholders develop and agree upon norms and moral judgments by engaging in deliberation (Pellé and Reber 2014, p. 41). The rightness/goodness of norms depends on the quality of stakeholder inclusion and deliberation. These norms can be translated into conditions that the innovation outcomes and their impacts should meet. The substantive approach builds primarily on prior given norms and moral judgments to determine if the outcomes and impacts of innovation processes can be deemed responsible (ibid.). For example, Von Schomberg (2013) builds on the normative anchor points presented in the European Treaty (e.g. sustainable development, social justice and protection, equality, and sustainable economic growth). Translated into broad innovation requirements, it means that Responsible Innovations should be societally desirable, sustainable, and ethically acceptable (Von Schomberg 2013).

11.3 Social Innovation

Social innovation is anything but a new phenomenon (Mumford 2002) and most of the research and definitions of social innovation are introduced by people who solved practical problems, instead of scholars who developed social innovation

theory (Caulier-Grice et al. 2012). Consequently, publications on social innovation have been mostly practice-oriented (Choi and Majumdar 2014).

However, the term social innovation is nowadays commonly, but not consistently, used by scientists (Moulaert et al. 2005) as it is conceptualised and defined in different ways (Cajaiba-Santana 2014; Choi and Majumdar 2014). For example, the term social innovation is not only used as a synonym for (unintended) social change, but also for intangible innovations that are designed with an intention to achieve specific ends (Choi and Majumdar 2014). However, social innovation often takes part in the entrepreneurial context where it encompasses innovations that are "explicitly aiming at the creation of social value and thus at positive social change. Hence, in this case, the 'social' denotes that the purpose of social innovation is to meet pressing social needs and to improve human and environmental well-being" (Choi and Majumdar 2014, p. 27). For example innovations that result in better access to healthcare, education or equal opportunities for income generation (ibid.)

The fact that social innovation is conceptualised and defined in different ways by different schools of researchers is also observed by van der Have and Rubalcaba (2016) who conducted a systematic network- and bibliometric analyses of social innovation.[1] This multiplicity of research schools that hold different perspectives on social innovation makes it hard, if not impossible, to achieve a consensus on the meaning of the concept (Choi and Majumdar 2014). Therefore, we argue that it is more appropriate to do a conceptual analysis based on literature reviews on social innovation (e.g. Choi and Majumdar 2014; Sharra and Nyssens 2010; van der Have and Rubalcaba 2016) instead of doing a conceptual analysis based on a single definition of social innovation.

11.3.1 Input of Social Innovation

The purpose of social innovation is to enhance social- and/or environmental well-being by addressing social needs or by solving social problems (Choi and Majumdar 2014) that are not being met by government or market actors (Sharra and Nyssens 2010). Also Van der Have and Rubalcaba (2016) observed that social innovations aim to meet common goals, solve social (-technical) challenges, or address matters of local development. More specifically, they identified an academic community that views social innovations as solutions to social (-technical) challenges, primarily directed to sustainability of climate, environment and health provisions (ibid.).

[1] For more information regarding the history of social innovation as a scientific concept and how different scientific communities influenced the scientific discourse on the concept, please see Choi and Majumdar (2014) and Van Der Have and Rubalcaba (2016). Since this goes beyond the aim of this chapter, it is not thoroughly discussed here.

11.3.2 Throughput of Social Innovation

Regarding the process of social innovation, there are two distinct streams of researchers that have a process-oriented understanding of social innovation (Van Der Have and Rubalcaba 2016). Researchers who investigate social innovation from a community psychology perspective understand social innovation as a process for systemically introducing change in social systems to solve (complex) social problems. Researchers investigating social innovation from a creativity research perspective aim to understand how new ideas of social relationships and social organisation are developed to generate and implement solutions to meet a common goal (ibid.). These two schools were also identified by Choi and Majumdar (2014).

There is also a stream of researchers who focus on the role of social innovation in local development (Choi and Majumdar 2014; van der Have and Rubalcaba 2016). They understand social innovation as: "satisfying human needs through (an empowering) change in the relations between local civil communities and their governing bodies" (van der Have and Rubalcaba 2016, p. 1928). This cluster pays special attention to the role of institutions and inclusive forms of collaboration in social innovation processes (ibid.). That collaboration is important in social innovation becomes clear in the review Sharra and Nyssens (2010) who found that the major characteristic of the social innovation process is the involvement of "a complex network of formal and/or informal partnerships between various stakeholders" (Sharra and Nyssens 2010, p. 7). Likewise, Dawson and Daniel (2010, p. 16) describe social innovation as a "process of collective idea generation, selection and implementation by people who participate collaboratively to meet social challenges". Social innovation is seen a collective endeavour where innovators and stakeholders (primarily target beneficiaries) reflect upon the purpose and end of the social innovation (Choi and Majumdar 2014). Especially practice-led research regarding social innovation stresses a dual objective, namely developing innovative solutions for societal problems while at the same time making sure that societal stakeholders have the capacity to act (ibid.).

11.3.3 Output of Social Innovation

The review by Sharra and Nyssens (2010) revealed that all conceptions of social innovation outputs share the element of novelty, meaning that these innovations can be new to the user, context, or application. Social innovations are distinguished from inventions by the fact that they are 'in use' and contribute to human and social life (van der Have and Rubalcaba 2016) which is similar to market adoption that makes the difference between (technological) innovations and inventions.

Social innovations can be found along a formalisation continuum. On one end, one can find highly formalised social innovations that are well-defined and have specific properties (e.g. the ethical and modular smartphone by *Fairphone*). On the

other end of the continuum one finds social innovations that are less formalised. These less formalised social innovations (e.g. minority empowerment program) are consisting of several services and smaller interventions that are continuously adjusted in response to the target group who act as co-creators (Choi and Majumdar 2014). Furthermore, van der Have and Rubalcaba came to a similar observation as Choi and Majumdar (2014), which is that different streams of researchers investigating social innovation do support the idea that:

"[Social innovation] has an important commonality in sharing two 'core conceptual elements': [social innovation] encompasses 1) a change in social relationships, -systems, or -structures, and 2) such changes serve a shared human need/goal or solve a socially relevant problem" (van der Have and Rubalcaba 2016, p. 1932).

More specifically, Choi and Majumdar (2014) state that "the dimension of change processes points not only to sustainable and long-lasting, systemic changes induced by social innovations, but also to the contexts, settings, and their specific structures in which social innovations are embedded" (p. 30). However, like any other actor engaged in innovation, also social innovators can experience resistance coming from different interests and power relations, or changing roles and mental models (ibid.).

11.4 Similarities and Dissimilarities Between Responsible Innovation and Social Innovation

11.4.1 Input

Science and technological development alone will not be able to tackle grand societal challenges (Sabadie 2014). Therefore, social innovations are increasingly understood as means to solve grand challenges in societies (Benneworth et al. 2015). Therefore, supported by the systematic literature reviews on social innovation, we argue that the grand societal challenges of our times do not only function as inputs for Responsible Innovation but also for social innovation. Responsible Innovation is also initiated to accommodate the inherent uncertainty that comes with innovation. However, in the literature reviews we did not find any indications that this also holds for social innovation.

11.4.2 Throughput

Social innovation is partly overlapping with Responsible Innovation when it comes to anticipation. Social innovators aim to better understand the needs, dislocations, dissatisfactions and blockages of target beneficiaries, which subsequently helps in "generating ideas […] and identifying potential solutions" (Mulgan 2006, p. 149).

Subsequently, social innovators find ways to bring the social change that is necessary to solve social problems that the people face (Sharra and Nyssens 2010). Social innovation seems to be less engaged in foreseeing detrimental implications that the innovation could bring.

Social innovation does reflect on the purpose for innovation and the ends that they want to achieve (Choi and Majumdar 2014). Furthermore, successful social innovators reflect on their actions and commitments as they evaluate the actual impact of their social innovations (Mulgan 2006). However, in the literature reviews we did not find any indications that social innovators engage in second-order reflexivity, meaning that they reflect how their own theories and value systems have an influence on the development of their social innovation. This is where Responsible Innovation differs from social innovation, as Responsible Innovation aims to increase awareness of different perceived realities and value systems between stakeholders and innovators.

Social- and Responsible Innovation particularly stress the importance of stakeholder inclusion, especially the people who might be affected by the innovation. However, there are differences between social- and Responsible Innovation when it comes to the reasons for stakeholder inclusion. Social innovation involves stakeholders primarily for better understanding the social problem or the societal needs that have to be addressed by the innovation. The same holds for Responsible Innovation, but in addition Responsible Innovation includes stakeholders also to facilitate more pluralistic visions of the implications innovation (Ribeiro et al. 2016). This should not only involve envisioning beneficial implications but also possible detrimental implications. Furthermore, it seems that social innovation does not aim to involve all relevant stakeholders during an innovation process, as it primarily focuses on co-creation with its target beneficiaries. Besides, social innovation does not involve stakeholders to question the desirability of social change and enhanced social- and/or environmental well-being.

When it comes to responsiveness Mulgan (2006) found that successful social innovations are developed by engaging in trial-and-error, experimenting and following hunches; followed by developing, prototyping, and piloting first versions of the solution for further improvement. Social innovation often involves a collective response by stakeholders who cooperatively generate, select and implement ideas to solve a social problem (Dawson and Daniel 2010; Sharra and Nyssens 2010). Social innovations are continuously adapting to the context in which they are developed, and to the needs of its target beneficiaries who act as co-creators (Choi and Majumdar 2014). Target beneficiaries are especially involved as co-creators for social innovations that are less formalised.

11.4.3 Output

Responsible Innovations and social innovations are both revolving around novel solutions that can take many forms. However, Responsible Innovation is primarily involved in the governance of science and technological development (Benneworth et al. 2015), whereas social innovation is about developing innovations that result in the social change necessary for solving social problems. Therefore, social innovation could be informative for opening-up the narrow view on innovation that can be found in Responsible Innovation research. Furthermore, researchers in social innovation distinguish social innovations from social inventions by stating that the latter are not in use. This cannot be said for the current notion of Responsible Innovation, which does not differentiate between responsible science and technological development. Hence, Responsible Innovation could also involve inventions by scientists that are not turned into marketable products yet.

11.5 Sustainability-Related Innovation

There is a rather diverse knowledge base coming from research on innovations that address sustainability, which includes concepts like green-, eco-, environmental- and sustainable innovation. These concepts are used interchangeably (Schiederig et al. 2012) even though there are different research communities that provide different lenses on how to innovate for sustainability (Franceschini et al. 2016).[2] Schiederig et al. (2012) identified six aspects that are recurring in the different definitions of sustainable innovation concepts.

1. Sustainable innovations can appear in different forms like products, processes, services or business models.
2. Sustainable innovations have a market orientation, meaning that they satisfies needs and are competitive on the market.
3. Sustainable innovations should reduce environmental impact, preferably have no environmental impact
4. The full life-cycle of the innovation should be considered when assessing the sustainability effect of the innovation.
5. Sustainable innovations can be driven by economic or ecological motivations.
6. Sustainable innovations can set new standards of sustainability for firms.

[2] For more information regarding the history of sustainable innovation as a scientific concept and how different scientific communities influenced the scientific discourse on the concept, please see Franceschini et al. (2016) and Schiederig et al. (2012). Since this goes beyond the aim of this chapter, it will not be thoroughly discussed here.

11.5.1 Input of Sustainability-Oriented Innovations

Sustainability-oriented innovation processes are initiated to pursue sustainable development. The International Union for Conservation of Nature (IUCN) was first to introduce the term 'sustainable development' and defined it as *"the integration of conservation and development to ensure that modifications to the planet do indeed secure the survival and well-being of all people"* (Schiederig et al. 2012, p. 181). More specifically, sustainable innovation is driven by grand challenges such as: increasing energy consumption, climate change, dependency on fossil fuels, pollution and water shortages (Charter and Clark 2007). The motivations to address the grand challenges can be driven by social or environmental motivations, but also economic motivations as companies can see potential competitive advantages by responding to the grand challenges (ibid). The latter is more present in research on 'green innovation' that relates sustainable innovation more directly to management and competition objectives (Franceschini et al. 2016).

11.5.2 Throughput of Sustainability-Oriented Innovations

Adams et al. (2016) conducted a systematic literature review to identify, analyse and synthesise sustainability-oriented innovation practices and processes at firm-level. They argue that firms can engage in sustainable innovation on three different levels. Firms at the lower level are engaging in 'operational optimisation' and have an:

"internally oriented perspective on sustainability, referring to a 'doing the same things but better' approach directed toward reducing harm through reactive, incremental improvements driven by compliance or proactively pursuing efficiencies. These are activities characteristically technical, stand-alone and insular" (Adams et al. 2016).

These companies could be of primary interests to scientists engaged in 'eco-innovation', as Franceschini et al. (2016) found that these scientists investigate issues around technology design and products that primarily lead to efficiency gains. Since Responsible Innovation aims to go beyond compliance (Stilgoe et al. 2013), we do not consider this level of sustainable innovation to be relevant for Responsible Innovation.

Firms at higher levels of sustainable innovation operate closer to the ideal of Responsible Innovation. Adams et al. (2016) state that at a higher level of sustainable innovation, firms include the social aspect into the notion of sustainability as well. The 'organisational transformers' involve companies that engage in innovation activities that are more people-oriented. Furthermore, their sustainability-oriented innovations are not treated as insular events, and the idea of sustainability is embedded throughout the firm and preferably along the value chain. A small but growing number of firms go even further and make a more radical shift in philosophy. These firms aim to think beyond the firm by reflecting with other stakeholders,

including the public, on the role of their business and its innovations for a desirable future. These so-called 'system builders' focus more on developing networks of workable relations, including unconventional stakeholders and the public, who collaboratively create sustainability value. Such novel collaborations are important for engaging in dialogue, gaining legitimacy, finding opportunities for knowledge acquisition, and finding opportunities for responsive solutions (Adams et al. 2016).

11.5.3 Outputs of Sustainability-Oriented Innovation

In the end, innovation processes result in sustainable innovations when the products, processes or business models have reduced negative externalities and preferably have no negative impact at all. In order to critically evaluate the impact of sustainable innovation, it is required that one takes the full life-cycle of the innovation into account (Schiederig et al. 2012).

The final outcomes of sustainability-oriented innovations can appear in many forms since they can be technological (like in eco-innovation), related to services (also known as servitisation), but also systems-shaping innovations that consist of interconnected sets of innovations (Mulgan and Leadbeater 2013). The implications of systems-shaping innovations are that they shift cities, sectors, economies or other systems on a more sustainable path (Draper 2013), which is necessary when addressing grand challenges.

11.6 Similarities and Dissimilarities Between Responsible Innovation and Sustainable Innovation

11.6.1 Input

Grand societal problems or 'wicked problems' are not only inputs for responsible- and social innovations but also for sustainability-oriented innovations. This holds especially for system-shaping sustainable innovations, which are necessary for responding to grand challenges that are too large for single firms to solve on their own. Again, Responsible Innovation aims to accommodate for the uncertainty that innovations could have negative implications. However, in the literature reviews we did not find any indications that this also holds for sustainable innovation.

11.6.2 Throughput

Adams et al. (2016) state that organisations that start developing systems-shaping innovations initiate, mobilise, inspire and lead the change towards workable relationships with private, public and civil society partners. These workable

relationships are not only important for constructive dialogues to collectively define the problem, but they are also beneficial for knowledge acquisition and the search for solutions (Mirata and Emtairah 2005). Furthermore, the discussions with stakeholders aim to steer innovations in the right directions by discussing the role that the firm and its innovations can play in desirable futures (Adams et al. 2016).

Organisations engaging in sustainability-oriented innovations do reflect on the outcomes of their innovations. Successful firms reflect on their actions and commitments by measuring and disclosing the impacts of the innovation. Furthermore, organisations reflect on the role that they can play in developing system solutions for complex grand challenges that they cannot solve on their own. These organisations are:

> *"leaving behind the prevailing economic paradigm to reframe the purpose of the firm in society: a part of society, not apart from it". […] "They adopt a logic of collaboration and invest in system solutions to derive new shared value propositions from the entire socio-technical and ecosystem network to make a positive impact" (*Adams et al. 2016*, p. 192).*

It is therefore fair to assume that those organisations that are engaged in finding systems-shaping solutions think beyond their role responsibilities and reflect on their wider moral responsibilities as well, which is also a core characteristic of reflexivity in Responsible Innovation (Stilgoe et al. 2013).

Sustainability-oriented innovators engage in dialogues with stakeholders beyond their supply-chain, such as civil society actors and unconventional stakeholders like community action groups or social entrepreneurs. However, also important differences could be observed. While these stakeholders are included in sustainable innovation to better define the problem and its possible solutions, the literature does not suggest that they question the social, political and ethical implications of possible solutions. Therefore, it seems that the discussion focuses on desirable implications of sustainable innovation, while possible detrimental implications receive negligible attention.

Again, innovations involving operational optimisation are predominantly developed in response to legislation and regulation (Adams et al. 2016), which is not similar to responsiveness as it is understood in Responsible Innovation literature. Organisations engaged in organisational transformation or system-building innovations for sustainability, are more inclined to develop innovations that require mutual learning and collective problem solving (Adams et al. 2016). Firms are more successful in developing sustainable innovations if they are more responsive to weak signals coming from their immediate stakeholder environment. Not only does this require absorptive capacity and connections with stakeholders, but also proper internal knowledge management processes. Without proper knowledge management processes, firms will fail to develop system-changing solutions even though they do engage in stakeholder collaborations (Ayuso et al. 2011). While Responsible Innovation does acknowledge the importance of internal knowledge management processes, it remains underexposed in Responsible Innovation literature. It is even

less discussed how to manage such processes. Seebode et al. (2012) found that organisations that want to develop system-shaping solutions need to learn how to follow novel pathways, how to work with other stakeholders, and how to find new ways of knowledge management. The advantage of sustainable innovation literature is that there is more practice-based information how organisations can engage in organisational learning, which remains underexposed in Responsible Innovation literature.

At the highest level of sustainable innovation, stakeholders are consulted during the earliest stages of innovation to find out how firms and innovations can play a role in desirable futures. However, the reviews did not provide any information how firms proceed after this initial stage. Therefore, it remains unknown whether innovators and stakeholders are mutually responsive throughout the innovation process. Research by Blok et al. (2015) confirms a tendency by firms to be transparent towards stakeholders and to deliberate with them during the initial stages of the innovation process and close to implementation of the innovation, but not during the stages in between. Therefore, there are no indications that sustainable innovation is a fully democratic and transparent innovation process like the ideal of responsible research and innovation aims to be.

11.6.3 Outputs

Sustainable innovations at a lower level focus on operational optimisation, which often result in technology-based innovations that lead to efficiency gains (Adams et al. 2016). However, recent sustainability oriented innovations increasingly involve systems-shaping solutions that consist of "interconnected set[s] of innovations, where each influences the other, with innovation both in the parts of the system and in the ways in which they interconnect" (Mulgan and Leadbeater 2013, p. 4). Adams et al. (2016) links this observation to Draper's conception of sustainability, which can be seen as "set of actions that shift a system – a city, a sector, an economy – onto a more sustainable path" (Draper 2013, p. 11). Therefore, the similarity is that both responsible- and sustainable innovation involve complex innovations that enhance sustainable development.

However, the review by Adams et al. (2016) does not provide any evidence that sustainability-oriented innovations explicitly account for the normative anchor points of responsible research and innovation like social justice, equality, and sustainable economic growth. Adams et al. (2016) state that some sustainability-oriented innovators even aim to depart from the economic paradigm. Therefore, future research could investigate what the role of these different normative anchor points are for innovation in the business context (Table 11.1).

Table 11.1 Overview of the conceptual overlap and the differences between responsible-, social- and sustainable innovation

	Responsible innovation	Social innovation	Sustainable innovation
Input for innovation	The grand challenges (Burget et al. 2017; Von Schomberg 2014; Wickson and Carew 2014)	The Social needs and problems that are not being met by the government or market actors (Choi and Majumdar 2014; Mulgan et al. 2007)	The climate-related grand challenges (Charter and Clark 2007) that are often complex (Adams et al. 2016)
	The uncertainty regarding innovations' future impacts (Stilgoe et al. 2013)		The business opportunity to increase profits by developing a sustainable innovation (Adams et al. 2016; Franceschini et al. 2016; Schiederig et al. 2012)
	The embedding of innovation in society (Ribeiro et al. 2016 and von Schomberg 2014)		
Throughput of innovation (i.e. the process)	Taking into account innovations implications and assessing alternatives (Wickson and Carew 2014)	Collectively defining the problem and searching for solutions (Mulgan 2006) and understanding the implications of social innovations (Sharra and Nyssens 2010)	Comparing different innovations' impacts based on full life-cycle assessments (Schiederig et al. 2012) and engaging in scenario thinking (Adams et al. 2016)
	Reflecting on the effect of the underlying norms, values and beliefs on the innovation at stake (Stilgoe et al. 2013)	Deliberating by consulting whether the needs of target beneficiaries are met	Deliberating with stakeholders how the firm and its innovations can help to achieve system transformation for desirable futures (Adams et al. 2016).
	Deliberating with stakeholders about the underlying norms and values that should guide the innovation in the desirable direction	Assuming that values such as social equality and sustainability are desirable and translating them into innovation requirements	Reducing the environmental harm per unit(lower level). Including the social dimension in to sustainability next to environmental- and economic dimensions (medium level). Deliberating with stakeholders what desirable futures are and what role the firm & innovation could play (high level) (Adams et al. 2016).
	Or normative anchor points of the European Treaty are used as normative goals for Responsible Innovation		
	Reflecting on activities, commitments and assumptions (Stilgoe et al. 2013)	Reflecting on the social impact and setting of new goals (Mulgan 2006)	Assessing the impact of the innovation over its full life-cycle (Schiederig et al. 2012) and disclosure of its sustainability performance (Adams et al. 2016)
	Reflecting on wider moral responsibilities next to role responsibilities (Stilgoe et al. 2013)	Taking responsibility to solve societal needs and problems that others do not address (Choi and Majumdar 2014)	Reframing the purpose of the firm from being apart from society to being part of society (high level) (Adams et al. 2016).

	Involving and deliberating with the relevant stakeholders throughout a transparent innovation process (Von Schomberg 2012) to make better decisions and learn from each other (Wickson and Carew 2014).	Including stakeholders (primarily target beneficiaries) for better understanding of the addressed social need or problem (Sharra and Nyssens 2010; Choi and Majumdar 2014).	Including stakeholders to increase the knowledge base of the firm, to improve search activities, to enhance social legitimacy, and to develop responsive solutions (Adams et al. 2016). Helps to enhance mutual learning (Boons et al. 2013) and improve decision-making (Bos-Brouwers 2010).
	Engaging with stakeholders in general and members of the public in particular (Owen et al. 2012; Stilgoe et al. 2013)	Deliberating with stakeholders in general and the target beneficiaries in particular (Mulgan 2006; Sharra and Nyssens 2010; Choi and Majumdar 2014)	Engaging with supply-chain partners [lower level]. Engaging with stakeholders that represent the innovation system during the earliest stages of the innovation process [higher level]. Members of the public are seldom involved to enhance foresight (Adams et al. 2016).
	Acting and adapting to the results from stakeholder inclusion and deliberation (Wickson and Carew 2014)	Generating, selecting and implementing innovative ideas with other actors to meet social challenges (Dawson and Daniel 2010).	Realising mutual responsiveness among supply chain actors [lower level].
		Less formalised innovations are developed and adjusted according to the innovation context and needs of target beneficiaries (Choi and Majumdar 2014).	Realising responsiveness by developing an innovation agenda that responds to the desirable futures projected by involved stakeholders [higher level]. (Adams et al. 2016)

(continued)

Table 11.1 (continued)

	Responsible innovation	Social innovation	Sustainable innovation
Output for innovation	Innovations that are societally desirable, sustainable and ethically acceptable (Von Schomberg 2014)	Innovations that enhance social- and/or environmental well-being (Choi and Majumdar 2014; Sharra and Nyssens 2010).	Innovations with reduced environmental impact on society, preferably none (Chalmers 2013; Schiederig et al. 2012) that balances social-, environmental- and economic considerations.
	Predominantly new and emerging sciences and technologies.	Innovations that induce the social change necessary for addressing the societal need or problem (Choi and Majumdar 2014) that are already implemented in practice (van der Have and Rubalcaba 2016)	Sustainable innovation goes beyond technological solutions and increasingly involves services and business-model and organisational innovation (Adams et al. 2016; Boons and Lüdeke-Freund 2013)
	Along a formalisation continuum from tangible (e.g. drones) towards intangible (e.g. financial products, GMO, etc.)	Social innovations can be found along a continuum of specificity of the innovation's properties & characteristics	Not only technology-based innovations but also other innovations. Sometimes sustainable innovation consists of a set of interrelated innovations (Adams et al. 2016) that shift a system onto a more sustainable path (Draper 2013)

11.7 Discussion and Conclusions

The aim of this chapter was to identify conceptual similarities and differences between Responsible Innovation and social- and sustainable innovation, and what this means for Responsible Innovation in the business context. Due to the multiplicity of conceptualisations and definitions that can be found in each of the three innovation concepts, we considered it legitimate to base our conceptual analysis on literature reviews of responsible-, social- and sustainable innovation. The research objectives of the literature reviews that were consulted were different. For example, literature reviews aimed at analysing and synthesising innovation activities (e.g. Adams et al. 2016) explicate the different understandings of innovation between scientific schools (such as Franceschini et al. 2016 and van der Have and Rubalcaba 2016) or aimed at outlining the characteristics of innovation in different contexts (e.g. Choi and Majumdar 2014).

The findings from our conceptual analysis indicate that social- and sustainable innovation are conceptually overlapping with Responsible Innovation on several aspects of the input, throughput and output of innovation. However, the explicit focus on determining the underlying norms and values for innovation is what discriminates Responsible Innovation from social- and sustainable innovation. These underlying norms and values for Responsible Innovation can be determined based on the results of deliberation with all relevant stakeholders (i.e. procedural approach) or they can be predetermined (i.e. substantive approach).

The conceptualisations in the literature reviews of social and sustainable innovation indicate that both innovation concepts are primarily based on the substantive approach. For example, it is predetermined that social innovation encompasses innovations that create social change to serve a shared human need or to solve a societally relevant problem, which subsequently enhances social and/or environmental well-being (Choi and Majumdar 2014; van der Have and Rubalcaba 2016). Even though there is deliberated whether the societal needs are met, the aim of the deliberation is not to discuss values such as social equality and sustainability. It is also not deliberated whether values can be conflicting, or how values are translated into innovation requirements. Similarly, ‘sustainability’ revolves around reduction of environmental impact for the lowest level of sustainable innovators, whereas at the medium level the social dimension is included as well. However, a small, but growing, number of sustainable innovators involve stakeholders for consultation. Here they reflect on the role that the firm and its innovations could play in a future desirable society (Adams et al. 2016). While this approaches the ideal of Responsible Innovation, the reviews did not reveal if and how the innovation agendas of the firms are responsive to the stakeholders. One can question whether such consultation without formal vote or say is in accordance with the deliberative democracy that Responsible Innovation aims to achieve. While one can argue if such a democratic governance of innovation is desirable in societies outside Europe and North-America (Macnaghten et al. 2014) the major challenge is how to achieve democratic governance of emerging science and innovations (Stilgoe et al. 2013).

We argue that it is highly questionable whether a democratic governance of innovation in the business context could be achieved in our current political and socio-economic system. First of all, because one cannot expect that companies become transparent during innovation as it will jeopardize the information asymmetries on which their market opportunities depend (Blok and Lemmens 2015). Second of all, inclusion of all relevant stakeholders is questionable, because Responsible Innovations respond to grand challenges that involve a wide variety of stakeholders (Weber and Khademian 2008). In reality, companies can only manage a limited number of different stakeholders in their network (van Geenhuizen and Ye 2014). Third of all, the final decision-making authority regarding the innovation strategy is restricted to the company (Blok and Lemmens 2015) as the board is responsible for the return on investment, and has to act on behalf of its shareholders and serve shareholder interests. This dominant role of shareholders is even embedded in corporate law (Heath 2011). Hence, it is questionable if all stakeholders can be treated alike, not to mention if a company can be responsive to the demands of all stakeholders. In conclusion, since we question the possibility to meet the requirement of a democratic governance of innovation in the business context, and since we did not encounter it in the literature reviews on social- and sustainable innovation, we propose not to consider democratic governance as a necessary condition for Responsible Innovation in the business context.

Another reason why Responsible Innovation is dissimilar to social- and sustainable innovation is that it requires stakeholders to reflect on the innovation trajectory and on how this trajectory could be made responsive to the inherent uncertainty that comes with innovations. Even though Stilgoe et al. (2013) proposes that Responsible Innovation should not focus on negative implications (Ribeiro et al. 2016), it seems that it is still a point of difference between Responsible Innovation and social- and sustainable innovation. Therefore, we propose that the procedural approach that can be found in the current notion of Responsible Innovation should also apply for Responsible Innovation in the business context.

However, there are important similarities between Responsible Innovation and social- and sustainable innovation. For example, responsible-, social-, and sustainable innovation provide insights how innovations can be developed that respond to the grand challenges, which can subsequently enhance social and/or environmental well-being. Social innovation is for example informative for finding out how to be responsive to the needs of target beneficiaries and how to co-create with them. Sustainable innovation is informative for developing system-changing solutions that respond to grand challenges, while taking the social-, environmental- and economic considerations into account. We see two reasons why social- and sustainable innovation can function as ‘points of departure’ for our understanding of Responsible Innovation in the business context. First, because the results of our analysis indicate that social- and sustainable innovation are conceptually overlapping with Responsible Innovation on multiple aspects regarding the input, throughput and output of innovation. Second, because research regarding social- and sustainable innovation is more practice-oriented and more embedded in the business context than Responsible Innovation.

Based on evidence presented in the reviews on social- and sustainable innovation we derive two essential preconditions for effective implementation of Responsible Innovation in the business context. These preconditions are based on the innovation practices of system-building firms that are described in the review by Adams et al. (2016), as these firms are currently innovating closest to the ideal of Responsible Innovation.

First of all, firms need to diffuse the notion of sustainability throughout the firm, and consider themselves part of society and not apart from it. This requires that the values and aspirations of the board and the owners are in line with the notion of sustainability. This notion is that sustainability is not an attribute of a single firm, instead it can only be applied at systems level, which requires collaboration with actors from private industry, public sector and involves civil society partners and investment in systems solutions. This new approach to innovation needs to be communicated throughout the firm, and integrated in the incentives and reward systems of employees (Adams et al. 2016; Armstrong et al. 2012). These actions ensure that Responsible Innovation becomes part of the company culture (Armstrong et al. 2012). Social- and sustainable innovation literature can inform how this could be achieved at strategic and operational level. This is necessary since new research (Blok et al. 2017) shows the discrepancy between the implementation of Responsible Innovation at the strategic level and at the operational level in companies.

The novel collaborations with a variety of stakeholders help to engage in dialogue, gain social legitimacy, find opportunities for acquiring new knowledge, and also help to find creative and responsive solutions. However, even though firms might engage in stakeholder collaborations, they will fail to develop system-changing solutions if there is a lack of internal knowledge management processes (Ayuso et al. 2011). The stakeholders need to learn how they can find, form and perform within the new innovation systems (Adams et al. 2016). This can be done by experimenting and learning with new approaches to sustainability, while simultaneously maintaining the existing business model. This allows firms to adjust the knowledge management processes without risking their business model, while at the same time developing an effective management approach that integrates foresight and novel collaborations with stakeholders (ibid).

Which consequences does our proposal have for the concept of Responsible Innovation in the business context? In Responsible Innovation in the business context, anticipation is similar to the understanding of anticipation in Responsible Innovation literature. Anticipation in Responsible Innovation in the business context therefore involves proactive engagement in activities enhancing foresight that take place at the start of the innovation process (Stilgoe et al. 2013). Anticipation is about better understanding the dynamics between the innovation and the wider eco-system in which it is developed and implemented. This also requires that stakeholders are involved in the discussion about what they consider to be desirable futures, and what the roles are of the firm and its innovations in those futures (Adams et al. 2016). Additionally, it is important that not only the environmental and economic implications are taken into account, but also the social, political and ethical implications of the innovation. It is important to acknowledge that stakeholder inclusion

and enhanced reflexivity does not necessarily lead to ethical outcomes and justifications (Pellé and Reber 2015) especially because it is unlikely that a democratic governance of innovation takes place in the business context. Furthermore, Responsible Innovation should still take into account that innovation can have unforeseen negative implications as well. Adopting a more procedural approach whereby the norms and values guiding the innovation are scrutinised by others than the innovators themselves, could help to become aware of the socio-political and ethical implications of innovation. Unfortunately, the literature reviews did not reveal any information on how this can be achieved effectively when innovating in the business context.

Reflexivity in the business context consists of two components. The first is measuring and disclosing the impact of the innovation, which can subsequently act as a driver for enhancing the performance of the innovation (Adams et al. 2016). This means that one assesses how the innovation performs compared to the desirable implications that were discussed at the start of the innovation process. The second is reflecting on the firm's role responsibilities but also its wider moral responsibilities. Firms need to be aware that they are part of society and not apart of it. However, the reviews did not provide insights whether companies investigate how their value systems and theories influence the subsequent development of their innovations. Furthermore, they did not reveal if companies assess whether their processes of anticipation, reflexivity, inclusion and responsiveness are in line with public values. Therefore we conclude that social- and sustainable innovation are not helpful for implementing this so-called second-order reflexivity as part of Responsible Innovation in the business context.

Firms involve stakeholders in their innovation process for three reasons. First, to achieve better foresight thinking, and to reflect on the role of the firm and their innovations in society (Adams et al. 2016). Second, to translate their underlying values for innovation into innovation requirements that result in innovation outcomes that are aligned with the needs of the target group. Third, to be able to adjust their innovation in response to new knowledge and changing stakeholder needs (Adams et al. 2016). In line with some findings in responsible- and sustainable innovation, we argue that foresight thinking and reflecting on the role of the firm (and their innovations) in society will be beneficial if such discussions take place with stakeholders that are representative for society. However, it is not likely that this is taking place throughout the innovation process, instead this more likely takes place at the start of the innovation process. Furthermore, as already mentioned before, it cannot be expected that this innovation process is transparent.

Also in the business context, firms aim to develop innovations that respond to grand societal challenges and they aim to make sure that the innovation becomes properly embedded in society. Hence it is essential to deliberate with stakeholders about the role of the firm and its innovations in a desirable future. Social innovation is primarily engaged with the target beneficiaries who can act as co-creators, whereas sustainable innovation aims to include representative stakeholders of the innovation system during the earliest stages of the innovation. What follows from the literature reviews is that firms should engage in good working relationships with

stakeholders as it allows them to quickly respond to 'weak signals' such as new knowledge or changing stakeholder needs and values (Holmes and Smart 2009). It is the responsibility of the company that aims to develop the innovation to initiate, mobilise, inspire and lead the change towards workable relationships with stakeholders in order to achieve such a mutual responsiveness. Furthermore, companies need to find new ways to develop proper internal knowledge management processes, as well as processes that help to develop innovations that respond to grand challenges and changing stakeholder needs.

Some final remarks have to be made with regard to the conclusions of this chapter. This chapter reflects on the concept of Responsible Innovation and critically examines what it could entail in the business context. This was done based on literature reviews regarding responsible-, social- and sustainable innovation for reasons explained throughout this chapter. However, it should also be noted that this approach has its drawbacks. For example, the literature reviews had different aims than this chapter, and were written from the perspective of social- or sustainable innovation, which is different from Responsible Innovation. These different aims and scientific lenses affect the analysis and synthesis of the literature, and subsequently the conclusions that are drawn in these literature reviews. Hence, it cannot be ruled out that relevant information for the concept of Responsible Innovation was omitted from the results and conclusions of these reviews. We further have to acknowledge that the business context is portrayed in this chapter as a homogeneous entity. This was done to contrast Responsible Innovation in the business context from the current notion of Responsible Innovation that focuses predominantly on science and technological development. However, we acknowledge that the business context is rather heterogeneous in practice. Nevertheless, we think that this chapter can serve as a starting point for further conceptualisation and subsequent implementation of Responsible Innovation in the business context. Therefore, it aims to inspire future work by researchers and practitioners who are interested in Responsible Innovation in general, and the business context in particular (Table 11.2).

Table 11.2 Overview of the main characteristics of the current concept of Responsible Innovation and the main characteristics of Responsible Innovation in the business context

	Responsible Innovation	Responsible Innovation in the business context
Anticipation	Proactive foresight activities to understand system dynamics between innovation and innovation eco-system	Proactive foresight activities to understand system dynamics between innovation and innovation eco-system
	Stakeholder inclusion to envision desirable futures to steer innovations in desirable direction	Stakeholder inclusion to understand the role of the firm and its innovations in desirable futures
	Being aware of possible negative (unforeseen) consequences	Being aware of possible negative (unforeseen) consequences

(continued)

Table 11.2 (continued)

	Responsible Innovation	Responsible Innovation in the business context
Reflexivity	Reflecting on norms, actions and commitments	Measuring of the innovation's performance and disclosure of the results
	Being aware of subjectivity of knowledge and that perceived realities are not universally held	Reflecting on wider moral responsibilities next to role responsibilities
	Reflecting on the effect of underlying value systems and beliefs on the development of the innovation	
Inclusion	Inclusion of all relevant stakeholders including members of the public	Inclusion of stakeholders representing the innovation system, the target beneficiaries and preferably members of the public
	Involvement of stakeholders throughout a transparent and interactive process	Openness towards involved stakeholders during the initial innovation stages and testing and launching the innovation. No transparency during the development of the business case and the innovation itself
Responsiveness	The innovators and involved stakeholders are responsive to the results ensued from anticipation, reflexivity and inclusion.	Translation of desirable futures into requirements for innovation
		Adjustment of innovation in the light of new knowledge and stakeholder needs, especially target beneficiaries
	Mutual responsiveness by being co-responsible for the development and implications of innovation	Focus on proper internal knowledge management processes
		Company remains primary decision-maker and responsible for the development of the innovation
Innovation output	Focus on science and technological advancements	Innovations that involve complex systems-shaping solutions (often consisting of interrelated sets of innovations)
	Innovation outcomes can be found along a formalisation continuum	
		Innovations can be found along a formalisation continuum

References

Adams, Richard, Sally Jeanrenaud, John Bessant, David Denyer, and Patrick Overy. 2016. Sustainability-oriented innovation: A systematic review. *International Journal of Management Reviews* 18 (2): 180–205. doi:10.1111/ijmr.12068.

Armstrong, Margaret, Guillaume Cornut, Stéphane Delacôte, Marc Lenglet, Yuval Millo, Fabian Muniesa, Alexandre Pointier, and Yamina Tadjeddine. 2012. Towards a practical approach to responsible innovation in finance: New product committees revisited. *Journal of Financial Regulation and Compliance* 20 (2): 147–168. doi:10.1108/13581981211218289.

Ayuso, Silvia, Miguel Ángel Rodríguez, Roberto García-Castro, and Miguel Ángel Ariño. 2011. Does stakeholder engagement promote sustainable innovation orientation? *Industrial Management and Data Systems* 111 (9): 1399–1417. doi:10.1108/02635571111182764.

Baregheh, Anahita, Jennifer Rowley, and Sally Sambrook. 2009. Towards a multidisciplinary definition of innovation. *Management Decision* 47 (8): 1323–1339. doi:10.1108/00251740910984578.

Benneworth, Paul, Effie Amanatidou, Mónica Edwards-Schachter, and Magnus Gulbrandsen. 2015. *Social innovation futures: Beyond policy panacea and conceptual ambiguity*. No. 20150127. TIK working papers. Oslo. http://digital.csic.es/bitstream/10261/132837/1/tikwp201501127.pdf.

Blok, Vincent, and Pieter Lemmens. 2015. The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation. In *Responsible innovation 2*, ed. Bert-Jaap Koops, Jeroen van den Hoven, Henny Romijn, Tsjalling Swierstra, and Ilse Oosterlaken, 2nd ed., 19–35. Cham: Springer. doi:10.1007/978-3-319-17308-5_2.

Blok, Vincent, Linda Hoffmans, and E.F.M. Wubben. 2015. Stakeholder engagement for responsible innovation in the private sector: Critical issues and management practices. *Journal on Chain and Network Science* 15 (2): 147–164. doi:10.3920/JCNS2015.x003.

Blok, Vincent, Tjidde Tempels, Edwin Pietersma, and Leon Jansen. 2017. Exploring ethical decision making in responsible innovation: The case of innovations for healthy food. In *Responsible innovation 3*, ed. J. van den Hoven, N. Doorn, T. Swierstra, B.-J. Koops, and H. Romijn. Dordrecht: Springer.

Boons, Frank, and Florian Lüdeke-Freund. 2013. Business models for sustainable innovation: State-of-the-art and steps towards a research agenda. *Journal of Cleaner Production* 45 (April): 9–19. doi:10.1016/j.jclepro.2012.07.007.

Boons, Frank, Carlos Montalvo, Jaco Quist, and Marcus Wagner. 2013. Sustainable innovation, business models and economic performance: An overview. *Journal of Cleaner Production* 45 (April): 1–8. doi:10.1016/j.jclepro.2012.08.013.

Bos, Colette, Bart Walhout, Alexander Peine, and Harro van Lente. 2014. Steering with big words: Articulating ideographs in research programs. *Journal of Responsible Innovation* 1 (2): 151–170. doi:10.1080/23299460.2014.922732.

Bos-Brouwers, Hilke Elke Jacke. 2010. Corporate sustainability and innovation in SMEs: Evidence of themes and activities in practice. *Business Strategy and the Environment* 19 (7): 417–435. doi:10.1002/bse.652.

Burget, Mirjam, Emanuele Bardone, and Margus Pedaste. 2017. Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and Engineering Ethics* 23 (1): 1–19. doi:10.1007/s11948-016-9782-1.

Cajaiba-Santana, Giovany. 2014. Social innovation: Moving the field forward. A conceptual framework. *Technological Forecasting and Social Change* 82: 42–51. doi:10.1016/j.techfore.2013.05.008.

Caulier-Grice, Julie, Anna Davies, Robert Patrick, and Will Norman. 2012. Social innovation overview: A deliverable of the project: The theoretical, empirical and policy foundations for building social innovation in Europe. (TEPSIE), European Commission–7th Framework programme. Brussels, Belgium.

Chalmers, D. 2013. Social innovation: An exploration of the barriers faced by innovating organizations in the social economy. *Local Economy* 28 (1): 17–34. doi:10.1177/0269094212463677.

Charter, M, and T Clark. 2007. Sustainable innovation: Key conclusions from sustainable innovation conferences 2003–2006 organised by the centre for sustainable design. http://bic-innovation.com/static/bic/knowledge_base/documents/T160433.pdf.

Choi, Nia, and Satyajit Majumdar. 2014. Social innovation: Towards a conceptualisation. In *Technology and innovation for social change*, ed. Satyajit Majumdar, Samaptri Guha, and Nadiya Marakkath, 7–34. India: Springer. doi:10.1007/978-81-322-2071-8.

Dawson, Patrick, and Lisa Daniel. 2010. Understanding social innovation: A provisional framework. *International Journal of Technology Management* 51 (1): 9. doi:10.1504/IJTM.2010.033125.

Draper, Stephanie. 2013. *Creating the big shift: System innovation for sustainability*. London. http://www.forumforthefuture.org/sites/default/files/images/Forum/Documents/SI document v4.2 web spreads_1.pdf.

European Commission. 2013. *Innovation union: A pocket guide on a Europe 2020 initiative*. Luxembourg: Research Policy and Organisation. doi:10.2777/59336.

Franceschini, Simone, Lourenço G.D. Faria, and Roman Jurowetzki. 2016. Unveiling scientific communities about sustainability and innovation. A bibliometric journey around sustainable terms. *Journal of Cleaner Production* 127: 72–83. doi:10.1016/j.jclepro.2016.03.142.

Gianni, Robert, and Phillipe Goujon. 2014. WP2 current theory and practice task 2.3 construction of an analytical grid. *Analytical grid report to EC*. Current theory and practice. http://www.great-project.eu/deliverables_files/deliverables02.

Heath, Joseph. 2011. Business ethics and the "End of History" in corporate law. *Journal of Business Ethics* 102 (Suppl): 5–20. doi:10.1007/s10551-011-1192-3.

Holmes, Sara, and Palie Smart. 2009. Exploring open innovation practice in firm-nonprofit engagements: A corporate social responsibility perspective. *R&D Management* 39 (4): 394–409. doi:10.1111/j.1467-9310.2009.00569.x.

Lettice, Fiona, Kulwant Pawar, and Helen Rogers. 2013. *Responsible innovation: What challenges does it pose for the new product development process?* The Hague, The Netherlands.

Macnaghten, P., R. Owen, J. Stilgoe, B. Wynne, A. Azevedo, A. de Campos, J. Chilvers, et al. 2014. Responsible innovation across borders: Tensions, paradoxes and possibilities. *Journal of Responsible Innovation* 1 (2): 191–199. doi:10.1080/23299460.2014.922249.

Mirata, Murat, and Tareq Emtairah. 2005. Industrial symbiosis networks and the contribution to environmental innovation. *Journal of Cleaner Production* 13 (10–11): 993–1002. doi:10.1016/j.jclepro.2004.12.010.

Moulaert, F., F. Martinelli, E. Swyngedouw, and S. González. 2005. Towards alternative model(s) of local innovation. *Urban Studies* 42 (11): 1969–1990. doi:10.1080/00420980500279893.

Mulgan, Geoff. 2006. The process of social innovation. *Innovations: Technology, Governance, Globalization* 1 (2): 145–162. doi:10.1162/itgg.2006.1.2.145.

Mulgan, Geoff, and Charlie Leadbeater. 2013. *Systems innovation*. London, United Kingdom. http://www.nesta.org.uk/sites/default/files/systems_innovation_discussion_paper.pdf.

Mulgan, Geoff, Simon Tucker, Rushanara Ali, and Ben Sanders. 2007. *Social innovation: What it is, why it matters and how it can be accelerated*. Skoll Centre for Social Entrepreneurship. http://eureka.bodleian.ox.ac.uk/761/1/Social_Innovation.pdf.

Mumford, Michael D. 2002. Social innovation: Ten cases from Benjamin Franklin. *Creativity Research Journal* 14 (2): 253–266. doi:10.1207/S15326934CRJ1402_11.

Owen, Richard, Phil Macnaghten, and Jack Stilgoe. 2012. Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39 (6): 751–760. doi:10.1093/scipol/scs093.

Owen, Richard, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and David H. Guston. 2013. A framework for responsible innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. Richard Owen, John Bessant, and Maggy Heintz, 27–50. Chichester: Wiley.

Pellé, Sophie, and Bernard Reber. 2014. Responsible innovation models report. Current theory and practice. http://www.great-project.eu/research/Responsible_Innovation_Model_Report_versionforsubmission.docx.

Pellé, S., and B. Reber. 2015. Responsible innovation in the light of moral responsibility. *Journal on Chain and Network Science* 15 (2): 107–117. doi:10.3920/JCNS2014.x017.

Ribeiro, Barbara E., Robert D.J. Smith, and Kate Millar. 2016. A mobilising concept? Unpacking academic representations of responsible research and innovation. *Science and Engineering Ethics* 23 (1): 81–103. doi:10.1007/s11948-016-9761-6.

Sabadie, Jesús Alquézar. 2014. Technological innovation, human capital and social change for sustainability. Lessons learnt from the industrial technologies theme of the EU's research framework programme. *Science of the Total Environment* 481 (May): 668–673. doi:10.1016/j.scitotenv.2013.09.082.

Schiederig, Tim, Frank Tietze, and Cornelius Herstatt. 2012. Green innovation in technology and innovation management – An exploratory literature review. *R&D Management* 42 (2): 180–192. doi:10.1111/j.1467-9310.2011.00672.x.

Seebode, Dorothea, Sally Jeanrenaud, and John Bessant. 2012. Managing innovation for sustainability. *R&D Management* 42 (3): 195–206. doi:10.1111/j.1467-9310.2012.00678.x.

Sharra, Roméo, and M. Nyssens. 2010. *Social innovation: An interdisciplinary and critical review of the concept*. Louvain, Belgium. http://innovationsociale.lu/sites/default/files/2009_SI_anInterdisciplinary%2526CriticalReviewofConcept_Sharra-Nyssens.pdf.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580. doi:10.1016/j.respol.2013.05.008.

van den Hove, Sybille, Jacqueline McGlade, Pierre Mottet, and Michael H. Depledge. 2012. The innovation union: A perfect means to confused ends? *Environmental Science & Policy* 16: 73–80. doi:10.1016/j.envsci.2011.11.006.

van der Have, Robert P., and Luis Rubalcaba. 2016. Social innovation research: An emerging area of innovation studies? *Research Policy* 45 (9): 1923–1935. doi:10.1016/j.respol.2016.06.010.

van Geenhuizen, Marina, and Qing Ye. 2014. Responsible innovators: Open networks on the way to sustainability transitions. *Technological Forecasting and Social Change* 87 (September): 28–40. doi:10.1016/j.techfore.2014.06.001.

von Schomberg, René. 2012. Prospects for technology assessment in a framework of responsible research and innovation. In *Technikfolgen abschätzen lehren*, ed. Marc Dusseldorp and Richard Beecroft, 39–61. Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-93468-6.

———. 2013. A vision of responsible research and innovation. In *Managing the responsible emergence of science*, ed. Richard Owen, John Bessant, and Maggy Heintz, 51–74. London: Wiley. http://onlinelibrary.wiley.com/doi/10.1002/9781118551424.ch3/summary.

———. 2014. The quest for the 'Right' impacts of science and technology: A framework for responsible research and innovation. In *Responsible innovation 1*, ed. Jeroen van den Hoven, Neelke Doorn, Tsjalling Swierstra, Bert-Jaap Koops, and Henny Romijn. Dordrecht: Springer. doi:10.1007/978-94-017-8956-1_3.

Weber, Edward P., and Anne M. Khademian. 2008. Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review* 68 (2): 334–349. doi:10.1111/j.1540-6210.2007.00866.x.

Wickson, Fern, and Anna L. Carew. 2014. Quality criteria and indicators for responsible research and innovation: Learning from transdisciplinarity. *Journal of Responsible Innovation* 1 (3): 254–273. doi:10.1080/23299460.2014.963004.

Zwart, Hub, Laurens Landeweerd, and Arjan van Rooij. 2014. Adapt or perish? Assessing the recent shift in the european research funding arena from "ELSA" to "RRI". *Life Sciences, Society and Policy* 10 (1): 11. doi:10.1186/s40504-014-0011-x.

Chapter 12
Exploring Ethical Decision Making in Responsible Innovation: The Case of Innovations for Healthy Food

Vincent Blok, Tjidde Tempels, Edwin Pietersma, and Léon Jansen

Abstract In order to strengthen RI in the private sector, it is imperative to understand how companies organise this process, where it takes place (throughout the entire company or on specific levels), and what considerations and motivations are central in the innovation process. In this chapter, the questions of whether and where normative considerations play a role in the innovation process, and whether dimensions of RI are present in the innovation process, are addressed. In order answer these research questions, a theoretical framework is developed based on Jones's theory of ethical decision making and Cooper's stagegate model of innovation management. In order to answer the research questions, a specific case of innovations that contribute to public health is explored, namely, that of food companies that participate in a Front-of-Pack (FoP) logo for healthier food.

As the use of healthy food logos does not necessarily have a positive impact on sales and profits (Jansen LAM, De Vos S, Blok V. Motives of retailers for healthy food innovation and communication about healthy food choices. Conference paper at the MVI conference, 25–26 August 2015, The Hague, 2015), it is expected that in the decision-making process, as part of their innovation process, companies make several trade-offs between economic, technical and moral factors (Jahromi MJ, Manteghi N, Procedia Technol 1:490–495, 2012). As the social-ethical values at

V. Blok (✉)
Management Studies Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands

Philosophy Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands
e-mail: vincent.blok@wur.nl

T. Tempels
Philosophy Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands

E. Pietersma
Management Studies Group, School of Social Sciences, Wageningen University, Hollandseweg 1, 6707 KN Wageningen, The Netherlands

L. Jansen
Dutch Choices Foundation, Bronland 12-S, 6708 WG Wageningen, The Netherlands

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_12

stake in corporate innovation processes have remained to a large extent unexplored in research on innovation management, the aim of this chapter is to identify the motivations and barriers for companies embracing and continuing a FoP logo for healthier food, and to assess whether ethical considerations play a role in this innovation process. From the findings in this research, it will become clear that although the studied companies participated in a programme for healthy food and thus are responsive to the needs of society, and although the companies feel (partially) responsible for public health, ethical considerations do not play a central role in the operational innovation process. Instead, technical and economic considerations seem to prevail in the operational innovation process. Furthermore, none of the procedural dimensions of RI seems to be present at this level in the innovation trajectory. It is argued that this may be an indication that the ethical decision-making process for RI is not located at the level of the operational innovation process itself, but is something that might be located on a higher strategic level in the company. It is at this level that the moral decision is taken to adopt the FoP logo and to engage in the RI process. The findings cast a new light on the discourse on RI in general, and in the private sector in particular.

12.1 Introduction

In the wake of increasing lifestyle-related diseases like obesity, heart diseases and diabetes type 2, citizens, governments and civil society organisations are becoming increasingly concerned with the 'obesogenic' character of modern society. Over the past years, both governments and the general public have become increasingly aware of the impact that food consumption has on both public and individual health; a growing number of food consumers in western society no longer look only at the physical properties of food products, but are increasingly interested in the social, ethical, nutritional and environmental aspects of food (Van Loo et al. 2014).

In order to be responsive to the changed demands of society regarding healthy food, companies in the food sector are gradually taking responsibility for public health. Over the past years, the food industry has taken up a role in the prevention and mitigation of public health issues. These efforts move beyond general corporate social responsibility (CSR) practices, in which research and development (R&D) and innovation are often neglected, and primarily concern new product development[1]; a significant amount of the food sector's R&D budget is allocated to the reformulation of food products in order to reduce or substitute 'unhealthy' ingredients like sugar, saturated fatty acids and salt in food products (Roodenburg et al. 2011). These efforts can be understood as responsible innovation (RI), because, when innovating responsibly, corporate actors do not primarily try to achieve private economic goals, but rather to contribute to the solution of the grand challenge of lifestyle diseases (cf. Von Schomberg 2013).

[1] A comparison between CSR and RI is beyond the scope of this chapter. For this, see Pelle and Reber 2015.

Although numerous companies have joined in this innovation process for healthier food and take responsibility for societal problems, empirical research about RI in the private sector is scarce (Blok and Lemmens 2015; Blok et al. 2015). There is still little known about what drives companies to engage in the development of responsible innovations and whether these innovation processes can be characterised as responsible (cf. Stilgoe et al. 2013). In order to strengthen RI in the private sector, it is imperative to understand how companies organise this process, where it takes place (throughout the entire company or on specific levels), and what considerations and motivations are central in the innovation process.

In this chapter, the questions of whether and where normative considerations play a role in the innovation process, and whether dimensions of RI are present in the innovation process, are addressed. In order answer these research questions, a theoretical framework is developed based on Jones's theory of ethical decision making and Cooper's stage-gate model of innovation management (Jones 1991; Cooper 1990). The stage-gate model helps to elucidate how the innovation process is set up and where the key decision points are located, whereas Jones's theory can help to elucidate whether and where ethical considerations play a role in the decision-making process. Mapping the operational innovation process in this way makes it also possible to assess whether process dimensions of RI – anticipation, reflexivity, inclusion and responsiveness – are present in the innovation process (cf. Owen et al. 2013).

In order to answer the research questions, a specific case of innovations that contribute to public health is explored, namely, that of food companies that participate in a Front-of-Pack (FoP) logo for healthier food. FoP logos are used on food products to inform consumers about the healthier options in a product group. Food companies can only carry such logos when they meet a certain set of nutritional criteria, which are determined by the organisation behind the specific FoP logo (Jansen and Roodenburg 2015). When joining such a programme, or when existing criteria are tightened, companies are pushed to innovate for healthier food products in order to enable them to achieve or keep the logo.

As the use of healthy food logos does not necessarily have a positive impact on sales and profits (Jansen et al. 2015), it is expected that in the decision-making process, as part of their innovation process, companies make several trade-offs between economic, technical *and* moral factors (Jahromi and Manteghi 2012). As the social-ethical values at stake in corporate innovation processes have remained to a large extent unexplored in research on innovation management, the aim of this chapter is to identify the motivations and barriers for companies embracing and continuing a FoP logo for healthier food, and to assess whether ethical considerations play a role in this innovation process.

The remainder of this chapter is structured as follows: in Sect. 12.2, a theoretical framework is developed based on a literature review in the field of ethical decision making, RI and innovation management. In Sect. 12.3, the methodology is set out. The results are analysed in Sect. 12.4 and, in the final section, a conclusion is provided, as well directions for future research.

From the findings in this research, it will become clear that although the studied companies participated in a programme for healthy food and thus are responsive to the needs of society, and although the companies feel (partially) responsible for public health, ethical considerations do not play a central role in the operational innovation process. Instead, technical and economic considerations seem to prevail in the operational innovation process. Furthermore, none of the procedural dimensions of RI seems to be present at this level in the innovation trajectory. It is argued that this may be an indication that the ethical decision-making process for RI is not located at the level of the operational innovation process itself, but is something that might be located on a higher strategic level in the company. It is at this level that the moral decision is taken to adopt the FoP logo and to engage in the RI process. The findings cast a new light on the discourse on RI in general, and in the private sector in particular.

12.2 Literature Review

Because the exploration of the ethical decision-making process regarding RI for public health is the central goal of this chapter, the literature review starts with ethical decision making, followed by theories regarding RI and innovation management processes.

12.2.1 Ethical Decision Making

There is a wide variety of models of ethical decision making, but Jones's process-based four-stage model (1991) is considered to be one of the most inclusive and comprehensive (Crane and Matten 2010). According to Jones, ethical decision making takes place in four steps: (1) recognising moral issues; (2) making a moral judgement; (3) establishing moral intent and (4) engaging in moral behaviour (Jones 1991).

The process of ethical decision making starts with *the recognition of a moral issue*. A moral issue is present when a person freely engages in an action that could harm or benefit others. This means that many decisions have a moral dimension, but, in order to engage in ethical decision making, an actor has to recognise that he is dealing with a moral issue. An actor has to realise that his voluntary choice or action will affect other human beings. In the context of the development of food products, this can for instance refer to the awareness that certain ingredients can have a negative impact on consumer health. An ethical dilemma can arise when economic considerations of profit have to be weighed against societal interests (Nathan 2015). When a moral issue is not recognised, the decision-making process takes place according to other rationales, like for instance that of economic rationality (Jones 1991).

The second step in the model is *making a moral judgement*. It is in this stage that an actor takes a position regarding the moral issue at hand (e.g. whether the discrimination of a disabled co-worker is problematic, or whether investing in weapon industries that deliver weapons to dictatorships is a right decision). The outcome of this judgement is dependent on the stages of an individual's moral development, e.g. to what extent he holds the law in high regard, the kind of ethical principles to which he adheres and whether he is willing to uphold these principles. How this decision is made is also shaped by national, cultural and organisational characteristics (Jones 1991).

The third step is *the establishment of moral intent*. After making a moral judgement, an actor might know the 'right thing to do' and even have the intention of acting accordingly, but nevertheless decide not to act upon it when weighing the moral factors against other considerations (e.g. company interests, self-interest) (Jones 1991). A senior manager in a firm might know that it is morally wrong for his company to violate national environmental regulations and may have the intention of changing company policies, but may be hesitant to act upon it as he fears that doing so will negatively impact the profit of the firm and his position as a manager.

The fourth and final step is when an actor actually *engages in moral behaviour*, which can be understood as acting upon the established moral intent (Jones 1991). Assume that the senior manager in our earlier example did establish moral intent and proceeded to act upon the company's violation of environmental regulation. If he established a plan to stop the pollution and executed it, then he would be engaging in moral behaviour.

Apart from these four central steps, Jones identifies two central elements that can influence this process of decision making, namely, the moral intensity of an issue and organisational factors. Moral intensity can have an impact on each stage of the decision-making process and consists of six elements: magnitude of consequences, social consensus, probability of effect, temporal immediacy, proximity and concentration of effect (Jones 1991). The moral intensity of an issue can for instance be higher when the issue is perceived to affect a greater number of people, or if there is social consensus that certain moral wrongs are greater than others. In the context of food innovation, one can imagine that moral intensity would be higher if certain ingredients turned out to be toxic than if ingredients could have an unhealthy impact when consumed in high doses (salt, sugar, saturated fats). In addition, Jones posits that the last two stages of the decision-making process can be influenced by organisational factors, like group dynamics, authority and socialisation processes within the corporation. Job descriptions, reward systems and corporate culture can shape an inclination to engage in (un)ethical behaviour (Nathan 2015; Jones 1991).

Jones's model can be used to ascertain whether a moral issue is recognised and whether this ultimately leads to a company's engagement in moral behaviour. This idea of engaging in moral behaviour is closely connected to the framework of responsible innovation, which is often considered to be morally laudable behaviour (Von Schomberg 2013). If ethical decision making can be embedded in the RI framework, it becomes possible to assess where the normative junctures are located in the innovation process.

12.2.2 Responsible Innovation

Von Schomberg defines RI as "a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view to 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" (Von Schomberg 2013: 19). However, RI is not only achieved through interactive and transparent processes. In the debate on RI, it is possible to distinguish between two different approaches. The first leading approach is that of Von Schomberg, whose interpretation can be understood as a substantive approach to RI (Blok et al. 2017). Innovation should be an interactive and transparent process, and in line with the *normative goals* of society. In the context of the European Union, such *normative anchor points* are for instance: social justice, sustainable development, a competitive social market economy and a high quality of life. *Responsible* innovation should hence comprise these anchor points.

The other key approach to RI is that of Owen et al. (2013). Although Owen et al. (2013) consider the normative discussion on RI to be important, they argue for a deliberative and more *procedural* approach to RI. In the procedural approach, the RI norms are not set in stone, and the direction and shape of the innovation can still be influenced by public demand and changes in the environment (Stilgoe et al. 2013; Blok et al. 2017). Owen and colleagues' framework consists of four dimensions: anticipation, reflexivity, inclusion and responsiveness.

Anticipation in RI is about identifying potential intended or unintended impacts of innovation. These are not limited to the (in)direct impacts directly related to the function of the product itself, but also include the effects that an innovation could have on the economy, the environment or social relations. Anticipatory procedures are recognised if companies have activities in their innovation process to identify such impacts and subsequently use this knowledge in their decision-making process (Owen et al. 2013).

Reflexivity is focused on the role of the innovator. It requires the innovator to think about his broader role in society and reflect critically on his practices, activities, assumptions and knowledge. For companies, this also requires a reflection on values and motivations that drive their business activities and how this impacts society (Owen et al. 2013).

Inclusion is about including a wide range of stakeholders in the innovation process. This process goes beyond the inclusion of traditional stakeholders (shareholders, consumers and so on) and opens up innovation processes to a wide range of societal actors. If diverse actors are included, different perspectives enter the innovation process, making it possible to better understand the perceived risks and benefits of an innovation to society (Owen et al. 2013).

Responsiveness in RI refers to the ability of an innovating company to adapt or change its "shape or direction in response to stakeholder and public values and changing circumstances" (Stilgoe et al. 2013: 5). The innovating company should be willing to change the course of its innovation on the basis of societal concerns and public input.

This short overview of the RI concept shows that there should be many normative junctures in the RI process (cf. Owen et al. 2013). In this respect, the process can be understood as an ethical decision-making process, guided by Owen and colleagues' four dimensions. However, it remains unclear where these normative junctures are located in the course of the innovation process. In order to assess this, a clear insight is needed into how the innovation process is set up. Cooper's stage-gate model of innovation management (Cooper 1990, 2008) may help to map the decision-making process and to identify the procedural dimensions of RI in the operational innovation process.

12.2.3 The Stage-Gate Model of Innovation

There are many models in the field of innovation, but a process-based model is relevant for this study as its objective is to ascertain where in the innovation trajectory the ethical decision making takes place. An innovation management approach that is much used in both theory and practice is Cooper's stage-gate model (1990, 2008). In Cooper's model, the innovation process is divided into stages and gates, where each gate has a set of specified deliverables and criteria that the process has to meet before moving to the next (working) stage (Cooper 1990). Each stage consists of one or more activities where information (e.g. technical, financial, market, operational data) is collected and analysed. The results of this stage are used as input in a decision gate. A decision gate can be understood as a Go/Kill decision point, where it is decided whether or not to move a project on to the next gate. Screening criteria in such decision gates are often economic or technical in nature, based on costs, technical feasibility and consumer perception (Cooper 2008; Nathan 2015). Through this process, an innovation moves from an initial product idea to the eventual launch of the product on the market.

The stage-gate model can be used to assess how the decision-making and the innovation process take place in a company, and what criteria play a role in deciding to continue or stop an innovation process. The stage-gate model has also been identified as an approach that can be used within the RI context (cf. Macnaghten and Owen 2011). Instead of including only technical and economic criteria in the gates, the moral criteria from the four dimensions of RI can also be included. This means that possible societal risks and the impact of the product on social relations would be taken into account in the decision-making process (Macnaghten and Owen 2011; Nathan 2015).

12.2.4 Theoretical Framework: Innovation and Decision Making

Based on the above-discussed theoretical perspectives, the following theoretical framework is proposed that can be used to explore the ethical decision-making processes within an operational innovation process in the RI context (Fig. 12.1).

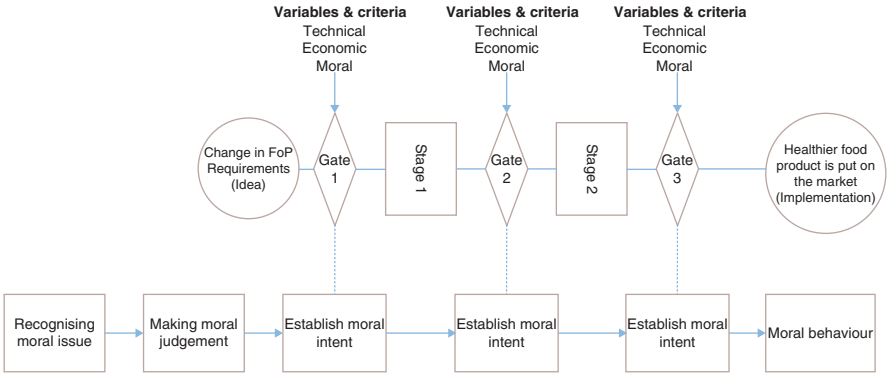

Fig. 12.1 Framework for ethical decision making in the operational innovation process (Note: *FoP* Front-of-Pack)

It is assumed that the ethical decision-making process starts in the early stage of the innovation process, for instance the ethical decision to innovate for healthy food. This generally starts with the idea for the development of a new product or the reformulation of an existing product. Before or synchronous with this process, it is assumed that the first two steps of ethical decision making – recognising a moral issue and making a moral judgement – take place. The next steps in the innovation process are linked to the process of establishing moral intent in the ethical decision-making process. It is here that various variables and criteria of an economic, technical and moral nature are used to decide whether or not to continue the innovation. The process of establishing moral intent is hence a process that can take place over multiple stages and gates. If in the end the responsible product is put on the market, the moral intent is translated into actual moral behaviour. This entire process can be influenced by both moral intensity and organisation factors. Although the moderating variables of moral intensity and organisational factors are likely to have an impact on this process, they do not determine the location of ethical decisions in the model, and consequently they have been left out of the model in this study.

12.3 Materials and Methods

In this research, an explorative qualitative approach is used to explore *whether* and *where* the ethical decision takes place in innovation processes for healthy food, and whether the procedural dimensions of RI are present in the operational innovation process. As there is limited research on this specific element of RI, an exploratory qualitative approach is warranted. In order to narrow down the scope of companies that are engaged in RI for healthy food, the companies selected for this study are involved in FoP nutrition logos. Products bearing such logos often contain less

saturated fat, sugar and/or salt and encourage consumers to make healthier choices. In order to meet the requirements for displaying such FoP nutrition logos on their products, companies have to invest in internal or external R&D and innovation for healthy food.

Given the wide variety of FoP logos, this study zooms in on the Dutch Choices Logo, which is a European FoP logo of the Choices International Foundation. Choices' logo criteria are based on national and international dietary guidelines. In order to display this FoP logo, products have to meet nutritional criteria of a specific food product group, which are set by a scientific committee. Every 4 years, the criteria for the logo are updated, and this can be a push for companies to innovate and make their products healthier (Roodenburg et al. 2011).

The subsequent case selection focused on larger companies (> small and medium enterprises) in the Dutch food industry that were participating in the Dutch Choices Foundation's FoP programme and marketing one or multiple products for which the criteria for the FoP logo had changed. The Netherlands was a useful location to collect data as it both has a mature food industry and is the country in which the Choices Foundation first started its FoP nutrition logo initiative. Six companies that fitted these criteria were approached to participate. Two companies declined because of time considerations. Therefore, the explorative case study consisted of three Dutch food innovating and producing companies and one Dutch retailer with private label products bearing the logo.

Based on the literature review, a questionnaire with open questions and closed questions was developed to structure the interviews. Audio-recorded interviews were conducted with key actors in the companies' operational innovation process, as this allows the participants to share experiences, explain the innovation process within their company and elaborate on the key considerations in the decision-making process (e.g. economic, technical and moral considerations). In order to ensure non-biased and structured data collection, an interview protocol was used. One of the researchers conducted four semi-structured interviews with one employee from each company. Employees were interviewed about their company's considerations in the innovation process, the barriers they encountered when confronted with new criteria for the FoP logo, the structure of the innovation process (e.g. inclusion of stakeholders), as well as their perceptions on corporate and personal responsibility for health and the impact of their products on public health. All interviews were transcribed. A summary of the transcription was sent back to the interviewees to enable them to make changes, adjustments or corrections. Subsequently, the transcriptions were analysed.

In addition, company documents (annual reports, CSR reports and company websites) were studied to obtain information on each company's perception of corporate responsibility for health and to get an overview of each company's CSR efforts (Boeije 2010). Combining the findings from the interviews with the data from the company documents triangulates the findings and helps to determine how the innovation process takes place and what considerations and barriers play a part in the decision-making process.

Table 12.1 Overview of case studies

	Company type	Number of product categories involved in the choices programme	Food and health in CSR strategy	Interviewee function
Company 1	Brand owner	Multiple	Yes	Nutrition manager
Company 2	Brand owner	Multiple	Yes	R&D manager
Company 3	Brand owner	Single	No	Marketing manager
Company 4	Private label (retail)	Multiple	Yes	Quality advisor

For reasons of data sensitivity, the data were coded. An overview of the cases is provided in Table 12.1. It shows the type of company, whether it produces food bearing the FoP logo in one or more product categories, whether a food and health-related CSR strategy is publicly available and the function of the respondents interviewed.

12.4 Results

In order to understand ethical decision making in the operational innovation process for healthy food products, first the companies' general position on CSR and public health is explored and subsequently the relation between their CSR strategy and their adoption of FoP logos for healthier food. Having identified the motivations and barriers for a company embracing the FoP logo for their products, the study focuses on the operational innovation process itself and possible moral considerations in the process.

12.4.1 CSR in Relation to FoP Logos for Healthier Food

12.4.1.1 Companies' Position on CSR and Public Health

An analysis of company documents (website, CSR reports, annual reports) provides general insights into the companies' position with regard to societal and environmental responsibilities, but shows as well how the companies view their particular responsibilities for public health.

Company 1 views sustainability and justice as core values in its business operations, which it aims to put into practice in both social and environmental programmes. As regards health, it is clear that the company is aware of the impact of its food and beverage products on society. It makes clear that it realises that, because of its market position and large consumer base, its products can have a significant

impact on consumer health, and therefore it takes responsibility for public health. It takes action by contributing to research on the relation between nutrition, health and disease, by informing consumers about the nutritional values of its products and by developing new healthier food products. In doing so, it seeks to promote a healthier lifestyle and combat lifestyle-related diseases. This was confirmed in the interview. Its CSR strategy is focused on sustainability in general, in which health and nutrition targets play a central role.

In *Company 2*'s CSR strategy, it is clear that the company seeks to contribute to the solution of global challenges such as the growing world population, food supply and scarcity of natural resources. With regard to public health, the company makes a clear commitment to tackling malnutrition and diseases like obesity. In order to take up this challenge, it wants to be transparent about the nutritional values of its products, develop healthier food products, and inform and educate people about food intake and physical exercise. This was confirmed in the interview. Nutrition/health is one of the main pillars of their CSR strategy.

Compared to the previous two cases, *Company 3* has a fairly limited CSR strategy. No CSR reports or annual reports were publicly available, as it is part of a larger MNE, so the company's website and the interview were the only sources of information. On the company website, only a short section is dedicated to the company's efforts on environmental sustainability and how the company tries to reduce its environmental impact by changing the production process. The website does not articulate a public health-related CSR strategy or a clear perspective on responsibility for consumer health. During the interview, a more nuanced picture appeared. According to the respondent, the company has a sustainability programme and it continuously challenges itself to produce in a more sustainable way, for instance by improving its packaging. The respondent further confirmed that nutrition and health are not part of its CSR programme. Nutrition and public health "is in the first instance part of our business strategy and is high on the agenda of the innovation strategy." One of the objectives of the company is to reduce the number of calories in its products and to communicate this to its consumers, but its innovations for healthier food are primarily part of the business case for healthier food, rather than a corporate *responsibility* for public health.

Lastly, *Company 4* has again a more inclusive CSR strategy. The company considers itself to have a responsibility for society and future generations. The CSR strategy touches not only upon both social and environmental themes like sustainable resources, connecting with the local community, reducing food waste, but also upon the theme of food and public health. It states that the company has a role to play in stimulating both consumers and employees to eat healthily and live a healthy life. This position on CSR and public health was also confirmed in the interview. In order to achieve its goals, the company participates in public–private projects that seek to improve healthy consumption and a healthy lifestyle, through information and education on health within the company, and by healthy innovating of its own food and beverage products. *Company 4*'s ambition is to enable grocery shopping without worries, implying that it wants to communicate healthier choices to its consumers (Table 12.2).

Table 12.2 Companies' positions on CSR and public health

Company	General CSR strategy	Public health-related CSR	Examples of public health efforts
Company 1	Wealth generation through business should go hand-in-hand with sustainability and justice. Thus, the company is committed to improving both the environment (e.g. reduction of carbon footprint) and sustainable practices of consumers, contractors and the company itself.	Clear recognition that food and beverage products can impact consumer health. Improving public health is a central part of the company's CSR strategy, as it aims to contribute to solving problems like obesity and malnutrition.	Reformulation of food products (salt, sugar, fat)
			Contribute to programmes that educate on healthy diets and lifestyle
			Contribute to research on the relation between nutrition, health and disease
			Food labelling
Company 2	Contribute to the solution of global challenges such as the growing world population, food supply and scarcity of natural resources.	Contribute to the solution of global health challenges such as food security and obesity.	Reformulation of food products (salt, sugar, fat)
			Contribute to programmes that educate on healthy diets and lifestyle
			Food labelling
Company 3	No inclusive CSR strategy. Only a short reference is made to improving environmental sustainability in the production process.	No food and public health-related topics or responsibilities are discussed as part of the CSR strategy.	Food labelling
Company 4	The company sees a clear responsibility for current society and future generations. This is translated into diverse CSR efforts that touch upon both societal and environmental themes.	Healthier consumption is one of the main topics in the CSR programme. Stimulating a healthy lifestyle and healthy eating habits among consumers and employees is considered to be important.	Reformulation of private label products (salt, sugar, fat)
			Contribute to programmes that educate on healthy diets and lifestyle
			Food labelling

Note: *CSR* corporate social responsibility

This short overview makes clear that, in three out of four cases, there seems to be a general perception of (partial) corporate responsibility for public health. Apart from *Company 3*, all other companies discuss the relation between their products and consumer health and recognise the (in)direct impact that food and beverage products can have on public health; they should provide healthier products and increase the opportunities for consumers to make a healthier choice. At the same time, the companies hold that consumers are themselves responsible for their own healthy food consumption. According to all companies, consumers are primarily

responsible for their food consumption, and the companies can help them by providing healthier options and by helping them to make the right choice.

12.4.1.2 CSR and FoP Logos

In order to identify the relation between their general position on CSR and public health on the one hand, and their use of FoP logos on the other, respondents were asked to reflect on this relation. According to *Company 1,* both its CSR strategy and the FoP logo for healthy food are congruent with each other. "I think that there are many similarities, that we both [company CSR strategy and FoP logo] just want to communicate in a transparent way to consumers what is in our products and what is the healthier choice." Although the company performs this ambition more in the background, the FoP logo does it in a more explicit way. At the same time, the respondent said that the company was not reactive to particular changes in the criteria of local FoP logos, as *Company 1* is an MNE operating in different countries with different markets. Different countries have different logos, different criteria and different time schedules for tightening the criteria, whereas the company proactively plans its innovations years ahead and for all products in a specific category. In its innovation strategy, it works therefore with its own targets regarding salt, sugar and fat reduction, which partly overlap with the criteria of the FoP logo for healthier food in the Netherlands. It considers that its own criteria for innovations for healthier food are much more important than those provided by local FoP logos, partly because of the local diversity among logos. Comparable forms of proactiveness can be recognised in the case of *Companies 2* and *3.*

Company 2's position on CSR and public health is explicitly linked with the FoP logo for healthier food, because the criteria of the FoP logo are part of its corporate nutritional criteria, and its ambition is to have the logo on as many of its products as possible. At the same time, it acknowledges that it also produces less healthy food products. According to *Company 2*, these 'pleasure products' have another goal and function: "We don't claim that these products are good for you; they are pleasure products that fulfil a particular function. You shouldn't eat these products every day and we also don't try to sell them that way." For *Company 2*, it is not necessary for all its products to be healthy; rather, it offers a healthier product next to the regular variant. *Company 4* also has the ambition to have the FoP logo on as many of its products as possible in order to communicate that its products meet the criteria. At the same time, *Company 4* argues that it focuses on products that are relatively easy to reformulate in order to receive the logo, the so-called low-hanging fruit. Only for *Company 3* is there no link between its position on CSR and public health and the FoP logo for healthier food. Because it sees that the market for healthy food products is increasing, its innovation strategy is to reduce the number of calories in some of its brands and product types, namely, the ones that consumers choose because of taste *and* low calories. If it meets the criteria of the FoP logo for healthier food and can use the logo, that is great, but it is not a necessary condition.

Table 12.3 Overview motivations and barriers for the adoption of the FoP logo

	Motivations	Barriers
Company 1	Consumer communication about healthier choices	Effects on product quality
		Confusion because of too many logos on one product (aesthetics)
	Company image	
	Company growth	Decision power of larger countries/ consumer markets
	Anticipating changes in rules and regulations	
Company 2	Company image	Effects on product quality
	Company growth	Effects on price
	Responsibility as a leading brand/ frontrunner in innovation	Negative impact on shelf life
Company 3	Consumer communication about healthier choices	Effects on product quality
		Effects on price
	Company image	Negative consumer perception of FoP logo
	Company growth	
		Negative impact on shelf life
Company 4	Consumer communication about healthier choices	Effects on product quality
		Effects on price
	Remain competitive with competitors that use the FoP logo	

12.4.1.3 Motivations and Barriers for Adopting the FoP Logo

In Table 12.3, an overview of the motivations and barriers for the adoption of the FoP logo for healthier food is provided.

The majority of the companies indicated that the adoption of the FoP logo was driven by the perceived positive contribution of the logo to the company's image, the positive impact on the company's competitive position and improved consumer communication. *Company 1* for instance stated that its main reason for displaying the FoP logo is "to help consumers to make the choice of healthier products easier." "We are involved in product improvements, but if consumers do not know it, then you don't help people to choose these products." A logo that is easy to understand "provides education to consumers because it makes clear what a healthier choice is." They saw several advantages in making the product healthier, as they stated that "it is important for our image to collaborate with the initiative" and consider it to be "the only way to grow as a company." *Company 2* believes it to be its "role as producer to use the logo as often as possible in order to provide responsible products for the consumer." However, market considerations seem to prevail, as making a healthier product could create an advantage over competitors; "in the end it is all about the consumer preferring our products." "If health plays a role in this – which it does for consumers – then we create a better business." In addition, it considered the collaboration with the FoP logo to be a tool to communicate a positive message to both consumers and NGOs. For *Company 3*, the main reason for participating was "to contribute to healthier choices for consumers," and it argued that "the logo helps to communicate the company's position on health to consumers." From a

market perspective, it also deemed it attractive to make a healthier product, "for where there are such demands, there are possibilities for commercial gain." *Company 4*, the retail company, indicated that it displays the logo on its private label products "to inform our customers about healthier choices." To do this in the right way, it opted for the FoP logo for healthier food. Another reason for adopting this particular logo was that its competitor also adopted it on its products. "If we develop a private label, we look at referent products. If they carry the logo and we don't, that isn't good." Although *Company 4* feels responsible for healthier food on the one hand (Sect. 12.4.1.3), it is on the other hand only involved in actual innovations for healthier food if its competitors do as well. In this respect, therefore, it seems to be rather reactive.

As regards barriers to the adoption of FoP logos for healthier food, these relate primarily to the technical feasibility/unfeasibility of complying with the criteria, the quality (taste, preservation) and the price of the product (*Companies 1, 2 and 3*). *Company 1* stated: "If a gain in health comes at the cost of taste, it simply does not work ... if products don't taste good or don't sell then there won't be a contribution to public health." In a case where it removed the logo, *Company 2* pointed out that this was because "the change in criteria was too large and the consumer would have noticed the huge change in taste." This relation between costs and taste as a barrier to making healthier products is also recognised by *Company 3*. It argued that a significant decline in taste is a reason for losing the logo for that particular product. As decline in taste would have a negative impact on sales, several companies would be inclined to remove the logo. The retailer (*Company 4*) saw similar barriers, but, as its strategy is to copy its competitors, its reason for refraining from using the logo is when its competitor's referent product stops carrying the logo. *Company 1* also indicated another barrier related to the fact that it produces for the European market. If the criteria in a small country/consumer market like the Netherlands are tightened, but not in larger countries/consumer markets, the larger countries take the decision on whether and to what extent recipes will be reformulated, and the smaller countries have to follow. This could be one reason for a product in the Netherlands losing its logo.

The general position regarding CSR and public health, the relation between CSR and FoP logos for healthier food, and motivations and barriers for the adoption of FoP logos for healthier food having been outlined, the next section focuses on the innovation process and ethical decision making in the operational innovation process.

12.4.2 Ethical Decision Making in the Innovation Process

12.4.2.1 Assessing the Innovation Process

The interviews made clear how the innovation process was set up within the different companies. Respondents were asked about all process steps taken to reformulate a product after a change in FoP logo criteria. In all four companies, a kind of

stage-gate innovation process is initiated, with central gates in which Go/Kill decisions are made. For the majority of the companies, this process really seems to be a closed and internal company process. NGOs like *Foodwatch* are informed by the companies, and these NGOs also influence the innovation agenda, but they are not involved in the operational innovation process. Apart from the FoP logo for healthier food that provides the initial criteria, no other stakeholders (NGOs, consumers and so on) are involved in the decision-making processes in the operational innovation process. Having outlined the structure of their company's innovation process, the respondents were asked whether ethical considerations played a role in the decision-making process during the operational innovation process.

12.4.2.2 Recognising the Moral Issue

The starting point of the ethical decision process is the recognition of a moral issue. Besides the general position on companies' corporate responsibility for public health (Sect. 12.4.1.1), information was collected during the interviews. Respondents were asked whether a FoP logo for healthier food can help or harm people and whether making a product healthier can help or harm people. They all agreed that both a FoP logo and making a product healthier can help people and have a positive impact on their lives. As Jones stated that a moral issue is present when help or harm can ensue for someone, it can be concluded that the respondents realise that they are dealing with a moral issue, namely, that food and specific nutrients can have a positive or a negative impact on public health.

12.4.2.3 Making a Moral Judgement

The second step in the ethical decision-making process is to make a moral judgement. Moral judgements were identified via statements during the interview, and responsibilities stated on the company website, CSR reports and annual reports were considered as well. In addition, the respondents were asked about their conceptions of responsibility for health. Most respondents agree that consumers are primarily responsible for their food consumption, yet also hold that the company has (partial) responsibility for public health by producing healthier products and by informing consumers about healthier choices (see Sect. 12.4.1.1).

Combining the findings of Sects. 12.4.1.1 and 12.4.1.3 about the motivations for adopting a FoP logo for healthier food, it becomes clear what moral judgements have been made by the companies. *Companies 1*, *2* and 4 take the position that responsibility for health is a personal responsibility of consumers, but they also recognise that their food products indirectly have an impact on public health. Consequently, they perceive their responsibility for public health mainly as a responsibility to be transparent about what they put in their products and to provide healthier products, in order to enable the consumer to make the right (healthier) choice. Based on these results, it can be concluded that *Companies 1*, *2 and 4* make

a moral judgement and proceed to the phase of establishing a moral intent. In the case of *Company 4*, there seems to be an ambiguous morality. On the one hand, it recognises the moral issue and does see its responsibility in providing healthier food, but, when it comes to using the FoP logo, it uses it only when its competitors do so. It argues explicitly that it does not use a FoP logo when its competitors do not use it, and thus only innovates for healthier food when its competitors do. In the case of *Company 3,* morality does not seem to play a major role, as it argues that its innovations for healthy food result mainly from increased consumer demand. Economic motivations seem to be more persuasive than moral considerations, in particular in this case.

12.4.2.4 Establishing Moral Intent

In the process of establishing a moral intent, stages and gates in the innovation process have to be identified where companies weigh moral factors with other factors. *Company 1* stated that "all projects are proposed in a project plan where we have to state what sustainability-related gains there are, such as health improvement or water reduction. This has to be approved in order to start the project." This makes clear that the company weighs several factors in this phase, including moral factors like the contribution to public health. The operational innovation process has several stages and gates after the project planning and approval phase, but there is not a juncture where moral criteria are weighed against other factors. According to the respondents, the criteria in each gate are project dependent. These criteria regarding taste and texture are established in the project planning phase. In general, it seems to be the case that the logo is dropped – and, with this, innovations for healthy food are *not* executed – when the taste, texture or price of the products have changed significantly. *Company 2* operates in a similar fashion. The company investigates the current formulation, the new criteria and the corresponding deviation. The respondent stated that "we start the project if the deviation is small and we expect that we can compensate without a significant cost price increase." "We will not start the project if we are sure that it affects the quality too much." Here, moral considerations do not seem to play a role. The same goes for *Companies 3* and *4*. *Company 3* starts with a plan in which it "considers the amount of calorie reduction and the influence on taste and price." Nevertheless, the respondent argued that lower calorie level is a boundary condition to proceed: "the starting point is that it has to be healthier." Given that *Company 3* focuses primarily on the business case for healthy food, moral intent does not seem to play a major role in this case. In the retail company (*Company 4*), the innovation itself is done by an external company, but the testing is done both internally and externally and here taste is the prime concern as well. Again, these considerations seem to be of a more technical and economic nature. Except for compliance with existing rules and regulations, regular risk assessments and food safety considerations, the four procedural dimensions of RI are not recognised in the operational innovation process.

12.4.2.5 Engaging in Moral Behaviour

In all cases, the companies engage in moral behaviour only when they develop a healthier food product that is actually brought to market. This only happens when economic and technical criteria have been met as well.

With this, it becomes clear that all four companies seem to recognise the moral issue (the impact of food on health) and make a moral judgement (that the company has a responsibility to develop healthier products and to inform consumers about healthier choices). Except for *Company 1*, which considers the contribution of the innovation to public health during the operational innovation process, in the following process of establishing moral intent, moral considerations or ethical criteria do not seem to play any role in the operational innovation process. In the other companies, the criteria for continuing the innovation are based mainly on economic criteria (impact on price and taste) and technical criteria (impact on shelf life, maintaining the structure of the product). There is no involvement of CSR departments in the actual decision-making processes during the operational innovation process. In any case, the eventual moral behaviour of the company – the development of actual healthier food products – is heavily influenced by non-moral criteria in operational innovation processes; economic and technical criteria are a necessary condition for the actual performance of moral behaviour. This may be explained by the fact that improved health – operationalised by the criteria of the healthy food logo – is the starting point of innovations for healthier food, whereas technical feasibility and economic performance are seen as necessary conditions to proceed with innovation.

12.5 Conclusion and Discussion

This chapter focused on the question of whether and where normative considerations play a role in the innovation process, and whether procedural dimensions of RI are present in the innovation process. In order to answer these questions, the position towards CSR and public health and the decision-making process within the operational innovation process of four Dutch food companies were explored. All four companies participated in the FoP nutrition logo of the Dutch Choices Foundation and have been confronted with changes in the criteria for the logo that create a push towards innovating specific food products.

From the findings in this research, it becomes clear that, although the companies in this study are participating in a programme that aims to tackle public health problems and thus are responsive to the needs of society, and although the companies feel (partial) corporate responsibility for public health, ethical considerations do not seem to play a central role in the operational innovation process itself. The results show that, in the innovation process, several factors such as price and expected impact on quality are weighed in the decision to continue an innovation project for

healthy food, but moral factors like the impact on public health are not mentioned as criteria or considerations for (dis)continuing the innovation.

This absence of moral criteria in the operational innovation processes runs contrary to the theoretical expectations developed in Sect. 12.2; it was expected that moral considerations would be applied throughout the entire innovation process. This can be considered to be a first indication that ethical decision making does not take place throughout the entire innovation process – this confirms earlier research (Blok et al. 2015) – but might be located at a higher or strategic level in the company. The hypothesis seems to be legitimate that in a first type of case, as in *Company 1,* the processes of recognising the moral issue, making a moral judgement and establishing moral intent have already been established at a higher decision level within the company. When the decision is made, the process of establishing moral intent is continued at the operational level. In a second type of case, as in *Company 2*, the moral intention to contribute to healthy food innovations is the starting point of its innovations as well, although technical feasibility and economic performance are seen as necessary conditions to actually proceed with the innovation. In a third type of case, as in *Company 3,* the moral intention seems to be absent; its innovations for healthy food are primarily part of the business case for healthy food, rather than a corporate *responsibility* for public health. In all three types however, technical feasibility and economic performance seem to be a necessary condition for the continuation or not of healthy food innovations.

To understand what kinds of considerations play a role in the strategic decision-making process, the decision-making process at company board level should be explored. Because displaying healthy food logos does not necessarily have a positive impact on sales and profits (Jansen et al., *working paper),* it is likely that moral considerations of (partial) responsibility for public health are weighed at the strategic level in the company against considerations of consumer demand, stakeholder pressure, competitor behaviour and legislation, resulting in the decision to adopt the logo and start the innovation process. This initial decision could provide the input for the operational innovation process.

This study's findings suggest the need for a revised model of ethical decision making in innovation processes for healthy food products (Fig. 12.2). The new conceptual model suggests how the recognition of a moral issue, the making of a moral judgement and the first step in establishing moral intent happen prior to the operational innovation process. It is at this stage that moral considerations are likely to play a central role, whereas in the operational innovation process itself pure technological and economic considerations like product quality and costs take precedence.

The findings also cast a new light on the discourse on RI. Although the food products that carry the logo can be considered to be responsible innovations because they contribute to the solution of the grand challenge of public health (cf. Von Schomberg 2013), the RI process dimensions do not seem to be present in the operational innovation process. As no moral criteria surface in the operational innovation process, there is little anticipation and reflection on possible impacts on public health; nor are there any other stakeholders involved and engaged in the operational

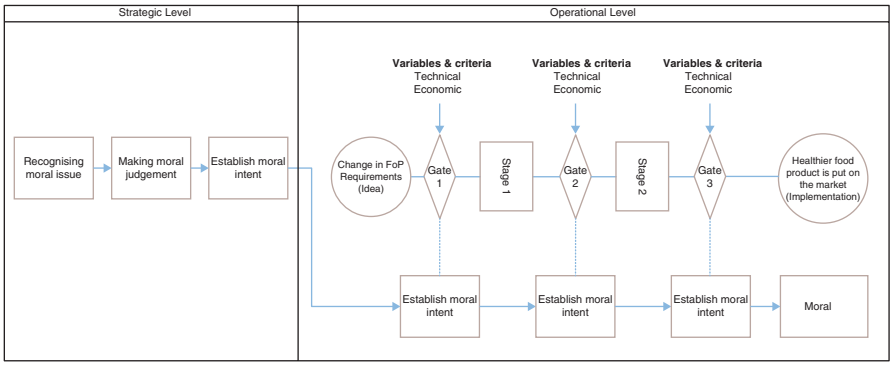

Fig. 12.2 Revised conceptual framework for ethical decision making in the innovation process (Note: *FoP* Front of Pack)

innovation process. This raises on the one hand the empirical question of whether the procedural dimensions of RI – anticipation, reflection, inclusion and responsiveness – are in fact present at the higher or strategic level in the company, and on the other hand the normative question of whether these RI dimensions *should* be present in strategic and operational innovation processes in order for such processes to be considered responsible innovation. We leave the question here of whether innovations that contribute to the solution of grand challenges but do not meet the RI process criteria should be considered as responsible innovations. This study's findings suggest that there might be a difference or even a discrepancy between considerations in decision making on the strategic and on the operational level. This has implications for the debate on RI in the private sector, as most current RI frameworks are about ethical governance during the entire innovation process (including both the strategic and the operational level).

There are several limitations to this study as well. First of all, an adapted version of Jones's theory is used in order to locate the junctures of ethical decision making. Including only the four core variables of Jones's model and omitting moderating variables like moral intensity and organisational factors may mean that factors that mitigate the considerations and motivations during the operational innovation process have been missed. Furthermore, the study may be only limitedly representative of the market because of the small number of cases and the specific focus on the Dutch food industry.

Although the current study is exploratory, it does provide first insights into the operational innovation process of innovations for healthier food in the Dutch food industry. Thus, it opens up new directions for future research. As the four case studies are exploratory rather than explanatory, the current findings can be used to guide new research. It will be relevant to find out (1) whether in other companies ethical considerations are absent from the operational innovation process, (2) whether a decision-making process for healthier food innovation takes place at the strategic level in a company and what considerations play a role at this level and (3) whether

the four RI dimensions surface at the strategic and/or the operational level of innovation processes in other sectors. Furthermore, (4) another question is whether the integration of the four RI dimensions in strategic and/or operational innovation processes can stimulate and guide future innovations for healthier food.

Conflicts of Interest Statement Leon Jansen is secretary of the Dutch Choices Foundation, which is responsible for the Dutch food logo. He was not involved in the data collection and primary analysis, but only in the further reflection on the findings.

References

Blok, V., and P. Lemmens. 2015. The emerging concepts of responsible innovation. Three reasons why it is questionable and calls for a radical transoformation of the concept of innovation. In *Responsible innovation 2: Concepts, approaches, and applications*, ed. B. Koops, I. Oosterlaken, J. van den Hoven, H. Romijn, and T. Swierstra, 19–35. Dordrecht: Springer.

Blok, V., L. Hoffmans, and E. Wubben. 2015. Stakeholder engagement for responsible innovation in the private sector: Critical issues and management practices. *Journal of Chain and Network Science* 15 (2): 147–164.

Blok, V., Lubberink, R., Van den Belt, H., Ritzer, S., Van der Kruk, H. and Danen, G. 2017. Challenging the ideal of transparency as a process and as an output variable of responsible innovation: The case of 'The Circle'. In *Responsible research and innovation. concepts and practices,* ed. R. Gianni, Pearson, J., B. Reber (forthcoming). Routledge: New York.

Boeije, H. 2010. *Analysis in qualitative research*. Los Angeles: Sage.

Cooper, R. 1990. Stage-gate systems: A new tool for managing new products. *Business Horizons* 33 (3): 44–54.

———. 2008. Perspective: The Stage-Gate® idea-to-launch process—Update, what's new, and nexgen systems. *The Journal of Product Innovation Management* 25: 213–232.

Crane, A., and D. Matten. 2010. *Business ethics*. New York: Oxford University Press.

Jahromi, M.J., and N. Manteghi. 2012. Innovation process in decision making model to choose an appropriate power system. *Procedia Technology* 1: 490–495.

Jansen, L.A.M., and A. Roodenburg. 2015. The use of food composition data in the choices international programme. *Food Chemistry* 193: 196–202.

Jansen, L.A.M., S. De Vos, and V. Blok. 2015. *Motives of retailers for healthy food innovation and communication about healthy food choices*. Conference paper at the MVI conference, 25–26 August 2015, The Hague.

Jones, T. 1991. Ethical decision making by individuals in organizations: An issue-contingent model. *Academy of Management Review* 16 (2): 366–395.

Macnaghten, P., and R. Owen. 2011. Environmental science: Good governance for geoengineering. *Nature* 479 (7373): 293–293.

Nathan, G. 2015. Innovation process and ethics in technology: An approach to ethical (responsible) innovation governance. *Journal on Chain and Network Science* 15 (2): 119–134.

Owen, R., J. Stilgoe, P. Macnaghten, M. Gorman, E. Fisher, and D. Gustion. 2013. A Framework for responsible innovation. In *Responsible innovation*, ed. R. Owen, J. Bessant, and M. Heintz, 27–50. London: Wiley.

Pelle, S., and B. Reber. 2015. Responsible innovation in the light of moral responsibility. *Journal on Chain and Network Science* 15: 107–118.

Roodenburg, A., B. Popkin, and J. Seidell. 2011. Development of international criteria for a front of package food labelling system: The international choices programme. *European Journal of Clinical Nutrition* 65: 1190–1200.

Stilgoe, J., R. Owen, and P. Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42: 1568–1580.
Van Loo, E.J., V. Caputo, R.M. Nayga Jr., and W. Verbeke. 2014. Consumers' valuation of sustainability labels on meat. *Food Policy* 49: 137–150.
Von Schomberg, R. 2013. A vision of responsible research and innovation. In *Responsible innovation*, ed. R. Owen, J. Bessant, and M. Heintz, 51–74. London: Wiley.

Chapter 13
Questioning the Normative Core of RI: The Challenges Posed to Stakeholder Engagement in a Corporate Setting

Merel Noorman, Tsjalling Swierstra, and Dorien Zandbergen

Abstract Responsible Innovation (RI) is a normative conception of technology development, which hopes to improve upon prevailing practices. One of its key principles is the active involvement of a broad range of stakeholders in deliberations in order to better embed innovations in society. In this paper, we examine the applicability of this principle in corporate settings and in smaller scale technological projects. We do so in the context of a case study focused on an innovation project of a start-up organisation with social aspirations. We describe our failed attempts to introduce RI-inspired stakeholder engagement approaches and articulate the 'reasonable reasons' why the organisation rejected these approaches. We then examine the methods that the organisation adopted to be responsive to various stakeholders' needs and values. Based on our analysis, we argue that there is a need for the field of RI to explore additional and alternative ways to address issues of stakeholder commitment and inclusion, in order to make RI's deliberative ideals more applicable to the rapid, fluid, partial, and provisional style of deliberation and decision making that we found in corporate contexts.

13.1 Introduction

Responsible Innovation (RI) is basically a normative conception of technology development, which hopes to improve upon prevailing practices. RI is both described in terms of substantive norms regarding the outcome (sustainability, etc.) or – more usually – in terms of procedural norms regarding the process. If the latter, some

M. Noorman (✉)
Maastricht University, Maastricht, The Netherlands
e-mail: m.noorman@maastrichtuniversity.nl

T. Swierstra
Department of Philosophy, Maastricht University, Maastricht, The Netherlands

D. Zandbergen
University of Amsterdam, Amsterdam, The Netherlands

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7_13

form of stakeholder involvement or participation in the development is invariably presented as being at the core of the process. At the same time, the notion of RI is unfortunately still fairly vague and general. Davies and Horst (2015) point out that RI is usually "situated at the macro-scale: its actors are policy organisations, countries, governments, or societies, and their field of action comprises entire innovation pathways, national regulatory systems, or 'the environment'" (p. 52). Many RI projects have a policy-oriented focus, looking for governance and regulatory mechanisms for technoscientific developments in fields such as bioengineering en nanotechnology. As yet, fewer RI research projects focus on developing ways to make innovators in corporate settings and in smaller scale technological projects take account of concerns about the embedding of technology in society (Van der Burg and Swierstra 2013).

Stakeholder engagement in RI does not come with a handbook of tools and methods and it can be difficult to achieve in practice. To enhance reflection on the possibilities and constraints regarding stakeholder involvement, we will describe our attempt to introduce a form of it into an innovation project of a start-up organisation with social aspirations. In Sect. 1 we explore the normative content of RI, and discuss literature that explores how these ideals are translated to the innovation practice. In Sect. 2 we describe our participation in a start-up organisation that aimed to develop a digital platform that would allow citizen-users extensive control over their data. We soon realised that our RI ideals were far removed from practice on the ground, since the organisation did not invite a wide range of stakeholders to the design table. Rather than simply deploring this situation, in Sect. 3. we try to explicate the 'reasonable reasons' not to do so. In Sect. 4 we follow how the project leaders developed their own approach to be responsive to various stakeholders' needs and values through tinkering and improvisation. In the concluding Sect. 5 we reflect on our findings. Should practice be lifted to theory; should RI adapt its theory to make it more practicable and realistic; or should we do both at the same time?

13.2 Responsible Innovation and Stakeholder Involvement

RI is only the latest sibling in a whole family of approaches that aim to improve technology and its chances to be successfully embedded in society, like constructive technology assessment (CTA), interactive Technology Assessment (TA), real time TA, participatory TA, participatory design, ESA/ELSI-research, value sensitive design, public engagement with science, socio-technical integration, anticipatory governance, etc. Compared to these and other approaches, RI probably stands out in the more prominent role it attributes to ethics (Van den Hoven et al. 2014, p. 5). For example, RI is explicitly aspirational in character: it is not geared primarily at avoiding negative consequences, but rather at realising "the right impacts" (Von Schomberg 2011, 2013; Owen et al. 2013). Another particularly prominent aspect in RI is the foregrounding of responsibility issues: who is responsible for what (Grunwald 2014), and then particularly with respect to an uncertain and open future (Owen et al. 2013). Furthermore, RI tends to provide more space to the morally

ambiguous aspects of technology, in comparison to other approaches that either do not pay much attention to normative uncertainty or tend to reduce value conflicts to stakeholder interests (Grunwald 2014). Finally, RI hopes to solve moral value conflicts through smart innovation (Van den Hoven et al. 2012).

These differences should not hide from sight the fundamental affinity between RI and the other approaches mentioned. They have all been developed as critical alternatives to more traditional, expert driven and *post hoc* forms of TA. They constitute a broader democratisation trend in science and technology governance and research (Delgado et al. 2010; Lövbrand et al. 2010). As such, they all advocate exploring in advance possible impacts of a technology together with the stakeholders. The term 'stakeholder' here refers to a broad range of parties affected by the innovation, including partner companies, governments, authorities, inspection agencies, research institutes, as well as non-governmental and civil society organisations, activists groups and the general public.

As technological innovation is political (Stilgoe et al. 2013), stakeholders should engage in dialogue and deliberate with the aim of arriving at mutual understandings and shared goals, values and expectations. These different parties, with their wide ranging interests and expectations, should be brought into the innovation process at an early stage, while the technology is still fluid and responsive to external influence. Because the future is essentially open and unpredictable, the decision process should be continuously reflexive so as to allow for informed incremental response to changing circumstances (Guston and Sarewitz 2002). The deliberations can either lead to a sociotechnical design that matches the norms and values resulting from this deliberation, or to a revaluation of existing values and interests in the light of the possibilities opened-up by the new technology. In either case, deliberation will increase legitimacy by helping to ensure a match between values, interests, and technologies.

Besides increasing legitimacy, deliberation offers additional benefits, according to RI advocates. It helps to anticipate and accommodate challenges as well as to develop better user-centred technologies, empower citizens, crowd-source design, and resolve value conflicts through smart design (Lund Declaration 2009; Siune et al. 2009; Von Schomberg 2011, 2014; Owen et al. 2013; Wickson and Carew 2014; Taebi et al. 2014). Moreover, according to Von Schomberg (2013) such deliberations also enable co-responsibility of stakeholders. Stilgoe et al. (2013) place a similar emphasis on stakeholder deliberation as "taking care of the future" (p. 3).[1]

Siune et al. (2009) explicitly tie RI to the ideal of deliberative democracy: "In policies and activities concerned with public participation in science and technology, the normative ideals of deliberative democracy have become highly influential" (p. 28, see also p. 35). Here they build on a previous EU report; *Taking European Knowledge Society Seriously*, which already pleaded for increased interactions between reflective citizens and scientists. (Felt and Wynne 2007) Owen et al. (2013, p. 35), Von Schomberg (2014, p. 40) and Chalmers et al. (2014) similarly link RI to the ideal of deliberative democracy. Without using the exact word, Wickson et al.

[1] Stakeholder involvement has by now been embedded in the Horizon 2020 program, and is formalised as one of the RI requirements in a number of recent EU policy documents (for instance recently in Strand et al. 2015 and Kuhlman et al. 2016).

(2010) do so too. The core idea behind deliberative democracy, harking back to Habermas' idea of communicative reason (1990; 1996), is that interests and preferences should not be treated as given and static, but instead approached as subject to rational reflection, deliberation, and revision. Deliberative democracy is essentially not about how to get what we want given that others have conflicting wants, but about what we *should* want (Elster 1998). Its core is thus not aggregating votes and negotiating amongst interest groups, but collective deliberation about the public good. For the outcomes to be binding and persuasive, deliberation has to meet certain normative standards, for example that no one is excluded from contributing, that all participants have equal access to relevant information, that everything can be put on the agenda, and that participants sincerely assess the validity of arguments (Habermas 1996; Gutmann and Thompson 1996).

The trend towards deliberative democracy within science and technology policy, has met with criticism from the start. It has especially been accused of being naive and unrealistic. Horst (2007) for instance questioned the optimistic belief, inherent to deliberative democracy, that rational consensus is always possible. Stirling (2008), although clearly in favour of deliberative democracy, highlighted the tension between 'opening up' a decision-making process by inviting stakeholders, and the need to 'close down' that process in order to achieve closure and to move on. More generally, Lövbrand et al. (2010) questioned whether the constructivist and contextualising assumptions of STS research are indeed compatible with the more rationalist and universalist assumptions of deliberative democracy, and whether increased deliberation indeed results in increased legitimacy. Pandza and Ellwood (2013) showed that technology actors have difficulty accepting responsibility for uncertain and coproduced impacts, which makes RI hard to achieve. More recently Van Oudheusden (2014) showed that neither on the EU level, nor on the Flemish level, RI conceptions allow for something like 'politics', understood as the constitution and contestation of power. And even Stilgoe et al. (2013) themselves pointed out that stakeholder inclusion is a problematic concept (p. 5). It should be about opening up, but what that entails is under discussion. Is it public dialogue, pure forms of engagement, or are there other approaches that are good enough? Are there power differences that frustrate deliberation? When and how should different voices be brought in and the process opened up? How should decisions be made in case of lingering disagreement? (For an excellent overview of these issues see Delgado et al. 2010).

Finding answers to these questions requires studying how the deliberative ideal fares in practice. There exists little RI research on what opening up to more voices and responsiveness means for companies or starts-ups that aspire to responsibly innovate to address particular societal challenges. Blok and Lemmens (2015) point out that such projects are very different from the RI projects on large-scale technoscientific developments. In the latter type of projects, RI is usually argued for from a scientific and research perspective, and there is room for consensus forums, citizen panels etc. Corporate practices, according to them, pose different constraints that complicate such deliberative methods. They critique the idea of RI for its often naïve conception of innovation. The call for transparent innovation processes, for

example, conflicts with common incentives for firms to exploit information asymmetries between stakeholders to gain a competitive advantage or to receive economic or societal support. Based on research in food innovation, Blok et al. (2015) argue that power imbalances, pressure to maintain a competitive advantage, time load, fear of loosing control, and conflicting interests stand in the way of sharing information and engaging in conversation on a regular basis with stakeholders. The idea of co-responsibility is also misplaced, according to Blok et al., as only the investor can be responsible as he or she makes the decision to fund the development and dissemination of the product. In some cases it is also not in the interest of non-economic stakeholders to be co-responsible or to engage in product development, because they would have to give up their critical stance.

This analysis seems to leave little room for RI related stakeholder involvement in the corporate sphere. To further explore this, we selected a small start-up company, but one working with an explicit social mission, i.e. to build digital products that allowed its users' to have control over the data they shared through and across these products with other parties. Given the centrality of the values of trust and empowerment through transparency and control, the social mission of this organisation is aligned with the ideals of RI. Yet, in line with the analysis of Blok and Lemmens, we found that this did not straightforwardly translate into inclusive stakeholder involvement, given several restrictions the innovators had to work with. However, we also found that within those limits they tried to be responsive to stakeholders by iterative mediated consultation.

13.3 RI in Practice

In this section we analyse some of our early findings of a case study of a start-up company, which we will call Datashare, in order to further explore the challenges posed to stakeholder involvement based on deliberative democracy ideals in a corporate setting. In this case study we employed various methods, including participatory observation and interviews and we intended to organise workshops, in order to explore how stakeholders are made part of the development of an innovative digital product and how the position of weaker stakeholders can be strengthened in this development.

13.3.1 A Tension to Be Resolved

Datashare is an organisation initiated in 2014 by an energy network operator. It is a small organisation with a relatively young staff, consisting mostly of contract employees. In the following, we will focus on the first 7 months of our involvement with the organisation. During this time, the number of people working for the organisation grew from about 10 to around 25. The staff had a diverse range of expertise from business development, marketing and management to information

architecture, interface design and software development. The company differed somewhat from other start-up companies that still have to look for investors or that have multiple investors in that it was entirely dependent on the energy network operator for its funding from the start. Nevertheless, like other start-up companies, Datashare was under continuous pressure to show its relevance to the energy network operator.

In our first conversations with Datashare's team leaders, whom we will call John and Chris, they explained that they were developing and implementing an online platform that would enable residents of neighbourhoods, government organisations and service providers to exchange local, neighbourhood-based information. We will call this product *Platform. Platform*, they explained, was a first step towards their ambition of creating a data-sharing platform or network that would allow citizens control over their data. An internal document states: "People don't trust third-parties with their data", particularly when it concerns data related to the private sphere of the home, such as data about energy usage. Datashare wanted to create a trusted environment in which people would want to share such data. To this end, Datashare was in the early stages of developing a second product or service, which we call *Own. Own* would enable users of data-dependent services to control the conditions for the exchange and sharing of this data. What this would look like was unclear at the time and subject to continuous negotiations. One possibility that was discussed was integrating *Own* in *Platform* such that it would become a platform for privacy-friendly data sharing. Another option was that *Own* would take a more distributed form of a network application that would pop up when consumers used certain services, as a sort of third party guarantee that these services were safe to use from a privacy perspective.

From the outset it was clear that a tension existed between the principles underlying *Platform* and *Own*. In order to make these products attractive to potential users in terms of services provided as well as commercially sustainable, they would have to attract businesses interested in accessing the personal data of the citizen-users. As one of the team members expressed it: "Datashare is not only about giving control to residents regarding their data, but it is also about giving businesses access to the data of residents". This conflicts with the principle underlying *Own*, i.e. the idea that data control might also lead to people restricting access to their data. Moreover, a project that both offers data control as well as data-gathering sits uneasy in the context of the societal debate where both these positions are situated at polar ends: on the one end of the debate are critics who envision data control to be incommensurable with commercial gathering and repurposing of data. They call for more education of the general public so as to be more critical towards commercial data-sharing platforms. On the other end of the debate are commercial entities whose profit models depend in large part on their ability to gather and repurpose data willingly provided by users. These entities, thus, depend on a public that trusts data-exchange platforms hosting commercial services. For Datashare, both these groups are important stakeholders, necessary to further and legitimate their cause. Yet, the tension between these visions poses a challenge to the Datashare team leaders seeking to integrate the seemingly conflicting values that these visions embody, such as pri-

vacy, personal control of data, autonomy and efficiency, ease of use and profit, into their project.

The fact that Datashare needs to balance the interests of the various different stakeholders and their values, makes it an interesting case to investigate the applicability of RI's focus on deliberative approaches to dealing with tensions between stakeholder positions. As we have seen in the previous section, RI suggests upstream stakeholder dialogue, resulting in binding collective goals. In the following section we further explore some of the reasons why this approach proved problematic in our case study.

13.3.2 Reluctance to Engage Non-commercial Stakeholders: Reasonable Reasons

Our research at Datashare combined observation and intervention: we explored Datashare's ways of engaging their stakeholders as well as the possibility to complement them with our own RI-inspired methods and tools for organising the innovation process. In the first months of our research, we discussed with team leaders the possibility of facilitating a series of workshops, designed to open their innovation processes to a wide variety of stakeholders. In these workshops we would focus on the articulation of implicit values, biases and interests of Datashare's proposed technical design. We would also explicate values, norms and viewpoints together with stakeholders of the system, including users, developers, privacy organisations and third-party business partners. The aim of these workshops was to start a conversation as a first step towards establishing a shared vision about the values that Datashare's products should embrace and how these should translate into its sociotechnical design. However, our proposals of upstream engagement with stakeholders were dismissed by the project leaders.

This turn of events rendered our proposal a useful heuristic device. It highlighted and explicated points of friction and mismatch between, on the one hand, our RI-informed approach and, on the other, the particular ways in which Datashare pursued stakeholdership in its own way. In the following, we discuss two reasons why Datashare was reluctant to participate in deliberative approaches to involving a wide range of stakeholders, in order to highlight some of the problematic aspects of RI.

13.3.2.1 The Difficulty of Securing and Maintaining Engagement of Stakeholders

In one workshop, we intended to invite a group of about 20 stakeholders from diverse backgrounds, including commercial companies, public institutions, and potential users, to explore expectations about their roles and responsibilities in a particular workflow design for *Own*. But when we proposed such a stakeholder workshop to two team leaders, they expressed several concerns.

The first concern regarded their own relationships to the stakeholders. Blok et al. (2015) in their overview of critical issues for RI in corporate contexts point out that it is a challenge for firms to engage stakeholders in a context where outcomes of innovations are uncertain. For Datashare engaging certain stakeholders also took considerable effort. First, a group of stakeholders with a high priority for Datashare were the commercial companies, civil society institutions, and governmental organisations. The Datashare members considered them as both their (potential) customers and the reason for citizens to use *Platform* and *Own*. In this latter sense, Datashare "wanted something from them", according to one of the team members working on business development. These parties were supposed to offer their services on *Platform* to registered users, and eventually offer their services through *Own*. Furthermore, some of these clients were also regarded as potential business partners in the further development of *Own*. They were, for instance, invited to collaborate in developing products or services that would integrate *Own*. As a new organisation, Datashare had to prove that it would be in the interest of these potential customers and partners to work with Datashare. This was not always easy: team members concerned with business development told us that most companies and organisations were only partly interested in privacy solutions, and even less in users' control of data. The relationship with clients and business partners was, thus, considered to be very fragile and in need of careful nurturing.

The second concern regarded Datashare's relation with potential users and non-users, another important stakeholder group that proved difficult to engage. Like many other start-ups, Datashare did not yet have a fully-defined vision of what it was developing. At the time we suggested the workshops, they were still developing the concept of *Own*: what it was for; who it was for; what it would do? They also did not have a design for an information or technical architecture for *Own* and were still reviewing and experimenting with various possible technical solutions that could provide a basis for further development. This made it difficult to explain what Datashare was to potential users or even to identify potential users or 'residents' to engage in conversation. Moreover, the project leaders were doubtful that at this stage the residents could be reliable reporters on their interests and values. The team members assumed that there was a "latent need" for privacy, but it would be difficult to ask people about it. Datashare's team leaders told us that they believed most people, or residents, would not be able to discuss concerns about privacy and data control because they lack the appropriate knowledge and understanding. It would take too much time, according to them, to explain to them the current state of Datashare.

Not only the relationship between Datashare and business partners and residents was perceived as fragile or problematic, the interaction between the various stakeholders was as well. For example, Datashare members also talked with privacy activists, who were considered to be future privacy-critical users as well as sources of potential resistance or expertise. These stakeholders were considered important conversation partners, because they could help Datashare further explore the problem of privacy and data control and test the robustness of their provisional solutions by critically reflecting on them. However, in order to maintain credibility amongst

their business partners, as one of the team leaders told one of the authors, Datashare did not want to associate itself too much with what they feared business customers or partners might consider to be "tin-foiled hats", as privacy activists are derogatorily referred to. At the same time, one team member told us that if Datashare wanted to become a leading voice in a movement focused on privacy and data control, they had to be careful associating themselves with the business of data gathering and trade, in order to remain credible amongst privacy activists. Getting these two stakeholders groups to participate in deliberations would thus present a significant risk from the perspective of Datashare.

The perceived fragility of both the relationships between Datashare and its different stakeholders, and of the relationships amongst the various stakeholders, thus stood in the way of inviting the different stakeholders to directly participate in collaborative or deliberative activities. The project leaders did not assume that stakeholders were eager to participate in deliberation, as RI does, but had to be enticed and persuaded. A direct confrontation between the different perspectives constituted a substantial risk, rather than a way to resolve tensions through explication of different perspectives and deliberation.

13.3.2.2 Time Pressure and the Need to Produce

Like Blok and Lemmens (2015), we found that time and financial constraints put pressure on the organisation and limited the opportunities for stakeholders engagement. Like many start-up organisations, Datashare was under pressure to show its potential to its investors. The organisation at the time was fully funded by the energy network operator that initiated the organisation. A steering group consisting of influential people within the energy network operator was responsible for making decisions regarding budget and continuation of the project. This stakeholder remained invisible to us and to most of the Datashare team members, yet it showed its presence and influence indirectly through the feedback they regularly gave to John and Chris. For them it was not always easy to explain to the steering group what their team was doing. The steering group, they told us, had difficulty grasping the concept, putting pressure on the team to quickly produce tangible and visible results, either in terms of an actual product or service or in terms of 'proof' that there was an actual demand for it.

The Datashare team was inspired by the *Lean startup method* in the way it worked on these 'proofs'. This method provides a strategy for developing businesses and products in a 'learn-by-doing' manner (Ries 2011; Blank 2013). It assumes that traditional ways of creating a business or product through developing a business plan and then building a product or service before releasing it – do not work for new start-up organisations. These organisations are not able to write a detailed plan, according to creators of the method, because they are still looking for a business model and cannot yet know who their customers will be. There are too many uncertainties to meticulously plan. Moreover, they usually do not have the financial resources or time to develop a detailed plan before creating the product and getting customers.

They have to move quickly and establish early on whether their idea fulfils a need or not. If an idea is able to quickly gain 'traction', i.e. interest from customers, then the organisation has to 'persevere' in pursuing the idea; if not then it has to 'pivot' and adjust its idea. The method therefore promotes experimentation through an iterative and incremental method of validated learning consisting of short cycles of building, testing, measuring and refining a 'minimal viable product' (MVP).

One consequence of this *lean* mind-set for the organisation of Datashare was that, in their day-to-day activities, the different divisions of Datashare worked separately on different parts of the products and on different short-term goals. The Technological Development (TD) team was focused on finishing and launching *Platform* and later on testing and fine-tuning it based on feedback from users and business partners. The business development (BD) team was concerned with getting companies and organisations to publish on *Platform* as well as with building relationships with potential business partners for collaborating on further developing *Own*. The Lab was a division set up to test products and explore the needs of users. Marketing and Communication (MC) was responsible for the communication about the products to users and business partners as well as finding a way to connect Datashare with the public debate on privacy and data-security.

To Datashare, this way of working accommodated the uncertainty regarding the relationship between *Own* and *Platform* as well as the ways in which these two products would be integrated both technically and in terms of governance structure. At the time we proposed our workshop, *Platform* was on the verge of being a 'MVP', having a clear concept and ready to be launched. With respect to *Own*, however, Datashare was still mostly in, what one member of the management team called: 'service design', or the hypothesis testing phase. At this stage, data control was more a matter of branding, business-development and generating 'proof of concept' than something that could actually be built. As John put it, they were still looking for "evidence for something that did not exist yet".

Given this context, the workshop we proposed, explicating the technical architectures and governance models of *Own*, was, according to John, "out of scope". Moreover, as both team leaders feared, trying to integrate all aspects of the project would merely produce confusion both for team members as well as for other stakeholders. This was further illustrated when we addressed questions of technical architecture with Jarell, head of the Lab and responsible for prototyping *Own*. At the time, Jarell was in the process of collecting 'evidence', or 'proof of concept' that data-sharing services offering some control over data to users would appeal to potential future users. When one of us enquired about the technical measures needed to guarantee 'trust', 'data-ownership' and 'privacy', Jarell frowned and admitted he had no idea. This led to some confusion for Jarell. In the absence of clear solutions regarding these issues, Jarell now worried that Datashare was only selling a promise without living up to it. To John and Chris, this instance proved why Datashare is not served well by open deliberations about the many different components of the project. As John told us:

> In the context of Datashare's day-to-day practices, team members need to work on the goals that have been clearly defined, in short iterations. In this setting, it is our job to protect them from unnecessary confusion and abstract questions.

The example illustrates the aim of the team leaders to carefully manage the flow of information through the organisation. They did want the staff to think about the larger visions, but only at particularly designated and carefully managed occasions, such as team days. Questions that would lead team members working on particular parts of the problem away from their immediate goals would endanger the innovation process of Datashare. Our workshops with a focus on technicalities, while connecting different aspects of the project, were considered to be too early, too 'upstream', and distracting from the process.

13.4 A Different Approach to Resolving the Tension

The reluctance of the team leaders to open up the innovation process did not mean that they were not responsive to a wide range of voices. Rather, they took a different approach to dealing with the perceived tension between the different stakeholder needs and values, focusing in particular on discursive strategies. In this section, we explore some of these strategies.

The *lean startup method* provided a basis for Datashare's approach to negotiating the perceived tension between the call for more data sharing and calls for data control. The method is not only about breaking up tasks according to short-term goals, but also about being responsive to stakeholders' needs: through continuously consulting with potential 'customers', a lean-start-up organisation aims to develop products or services in a way that is responsive to customer needs (Ries 2011). An early internal strategy document about the development of *Own* provided an example of what this entailed. The document described a stepped process, in which Datashare would first, in collaboration with business partners, develop a set of product and service ideas, called 'propositions'. Through interviews with potential users of these products and services they would then validate their assumptions about the problem or need that the proposition would address. If the latter step demonstrated that a proposition could potentially be successful, they would develop a prototype of this proposition in collaboration with business partners in the next step. The prototype would then be 'tested' with the existing users of *Platform* and the response measured. The final step in this sequence was to improve the propositions on the basis of the interpretation of the measurement data and the lessons learned.

So rather than directly deliberating with a wide range of stakeholders in order to plan the development of a data-sharing network or platform that enabled residents to have control over their data, Datashare adopted an incremental approach to test and refine 'hypotheses' through mediated interactions with stakeholders. These mediated interactions kept the external stakeholders separated, while the members of Datashare would act as translators. As part of this process, the team leaders considered it their task to carefully manage and "cultivate" the information, as John described it, that flows through the organisation. Moreover, they saw translating and negotiating between all the different stakeholders as part of their core activities. In the following we describe how these translating practices on the one hand meant

gathering information about perspectives and testing assumptions and hypotheses about the needs of various stakeholders; and on the other hand translating and adjusting concepts, visions and products to perspectives of different stakeholders.

13.4.1 Gather Information and Testing Assumptions

The way that Datashare approached residents provides an illustration of how the team members would gather information and test their assumptions about the needs and interests of particular stakeholders. Although Datashare's team did not directly engage 'residents' or potential users and non-users in deliberations, one of their aims was to explicate the 'latent' need for data control via multiple translation steps. In their bi-weekly team meetings, for example, individual team members would often reflect on what they, their neighbour, or a family member might think about the particular way that Datashare framed privacy in its designs. In these reflections, they explicitly framed themselves as 'average potential users' of the system. The Lab team, as part of their 'lean' way of working, took a more explicit approach to substantiate and refine their assumptions about people's needs regarding data sharing and their views on privacy. John explained to one of us that he believed that if you ask people directly whether they are concerned about their data they will say no, but if you ask them in the context of a concrete example you may get a different answer. That, according to him, was what they were doing in the Lab. In order to validate and refine their assumptions about people's needs regarding data sharing and their views on privacy, they would develop and test propositions with small focus groups and respondents recruited by an agency. Or as described in the previous section, they would use mock-ups of interfaces to gauge how people would respond to particular representations of Datashare's ideas. The team members framed this type of work in terms of 'translation' with the Lab providing the context in which users' preferences could be made explicit.

The insights about the needs of the user stakeholder group were further augmented through consultation of professionals and experts. During one of the team days an external company presented the result of a trend analysis, performed at the request of Datashare, of online sentiments and public opinions about personal data and privacy. The analysis was based on automated text analysis of blogs and news sites. It was intended to sketch the various ways that people thought about privacy. On the same team day, the authors were asked to give a presentation on what was happening in the field of privacy. Datashare also consulted other professionals about the perspectives of individuals in neighbourhoods. For instance, they talked on several occasions with the social innovation team, employed by the energy network operator. This team went into neighbourhoods to talk with people about what concerned them. Datashare also provided funding to a small agency that developed creative solutions for informal care in neighbourhoods. By remaining in contact with them and occasionally accompanying them on one of their visits to neighbourhoods John hoped to learn about the needs in such neighbourhoods. Conversations

with the various experts, professionals and academics were intended to help translate the needs of residents and the general public to the team members, and in particular team leaders and members of the business development team, who in turn would selectively use this information to further develop their propositions.

Some stakeholders were more directly involved as conversation partners in the development process. For example, an explicit goal of the BD team in their conversations with potential clients and business partners was to learn about their needs and expectations concerning privacy and data sharing. Moreover, Datashare was planning to develop a more collaborative relationship with a few chosen business partners. These particular partners were selected based on their interest in addressing privacy issues and data protection concerns, as well as their willingness to take part in more experimental development projects. The aim was to have the partners provide concrete examples of the kind of problems *Own* was to solve and to co-create a proposition based on the *Own* concept in the Lab. These business partners, thus, had a more direct influence on the decision-making process, being only mediated by the Lab and the BD team.

Through the mediated interactions with stakeholders Datashare intended to gather information about stakeholders' needs, interests and values without confronting them with the tension between the different perspectives on data sharing. Using mock-ups and prototypes as well as face-to-face conversations, they aimed to elicit and elaborate the different perspectives to inform the further development of Datashare's products and services.

13.4.2 Adjusting Concepts, Visions and Products

To negotiate the tensions between the different perspectives, Datashare team members would use the information gathered to continuously mould multiple propositions and visions. Illustrative of this moulding is that in the period described, the team continuously changed the terms used to refer to *Own.* At first they would emphasise privacy, but later they would shift the focus to trust, then to control and personal autonomy, to then emphasise privacy again. These terms mattered with regard to how different stakeholders perceived *Own,* as we mentioned above, but also in relation to the technological and governance design of *Own.* A recurring discussion, for example, was whether personal data would be stored in the business partners' databases, in residents' homes or elsewhere. If *Own* would offer a privacy solution than the first option was conceivable, but if it were to offer residents a high-level of control than the second would be preferable. At the time, Datashare did not commit to one option, but explored, adjusted and refined multiple possibilities through the use of stories, propositions, mock-ups and prototypes.

Datashare, thus, did not have a singular vision, but juggled multiple visions as part of its translation activities. It would present different stories to different stakeholders regarding their product, emphasising different parts of their products to these different stakeholders. All the while, these stories were constantly tweaked

and changed according to the context. To their business partners, Datashare offered the prospect of better, more intimate contact with residents, who act as customers, consumers and generators of data simultaneously. When Datashare brought up the topic of data control in this context, it did so in a way that responded to the specific legal and commercial needs of these parties: telling them they would offer privacy as a service; to unburden them from the complex legal requirements vis-a-vis privacy and data control as imposed by European legislation. At the same time, Datashare's story to privacy activists was about designing privacy, autonomy and data control into their products; showing them Datashare was aware of the political, social and ethical issues around data sharing, gathering and profiling. In this way, Datashare profiled itself as *the* alternative to other commercial data-sharing platforms. In its ambition to gain a monopoly position in this area, its social vision was that of protecting the public from 'data-hungry, immoral' data-sharing monopolies such as Google or Facebook. In the story they developed for residents, Datashare aimed to offer more insight and understanding of what happens in their neighbourhood. It proposed trusted services that would allow them more easy access to neighbourhood services and products and more intimate and responsive connections with neighbours, service providers, municipalities and health professionals. As part of these multiple stories, the Datashare team leaders drew on multiple conceptions of residents' attitudes towards sharing their data, as either privacy-prone, as disinterested in privacy-issues or as in need of Datashare's intervention as forerunner in a *movement* focused on education and raising awareness about privacy.

As such, a central strategy employed for dealing with the tensions intrinsic to the conceptualisation of Datashare, was to keep different stakeholder groups and different stories about Datashare, carefully separated from each other. This upholding of ambiguity is a common approach in innovation contexts with multiple different stakeholder values (Stark 2011). As Barta and Neff (2016) put it: "when multiple values are in play simultaneously […] then the work of innovators is to recognise how to keep these multiple values ambiguous in order to appeal to different kinds of people" (2016, p. 520). Rather than constructing one unifying story applicable to all settings, team leaders instead allowed for ambiguity and carefully managing the information flows that went back and forth between the different groups.

Resolving the tension between the different stakeholders was, thus mostly a discursive strategy, at the time, as part of an incremental process of translation, vision development, testing and adjusting. Team members took on a mediator role to translate the different perspectives and negotiate a solution in the form of a business concept/service design to be later elaborated in a technological application, which had yet to take place.

13.5 Discussion

In this paper we explored the feasibility of the ideals of RI in a small-scale corporate context by looking at the ways in which Datashare managed the tensions between different stakeholder interests and perspectives. In particular, we focused on the

ways in which Datashare mitigated the central tension of its project between its ambitions to provide 'data control' on the one hand and, on the other, its development of services that invite people to share more data with third parties. Each side of this tension represents a different stakeholder group with different ideas regarding the question how 'data control', 'trust', 'user-autonomy', 'privacy', etc. should be organised both technically and socially in our information society.

We argued that the applicability of RI methods proved problematic in the specific context of Datashare. For a start-up organisation there are many uncertainties about the feasibility and demand for the products or services it is developing. These uncertainties pose a challenge for stakeholder engagement. Moreover time and financial constraints require start-up organisations to quickly make progress and show results even when the long-term bigger picture is not clear. An RI approach, facilitating the exposure of tensions, conflicts and aiming to establish a common vision regarding the outcomes of the innovation process, does not account for these constraints. The question then is whether practice should be lifted to theory, i.e. whether Datashare should be urged to change its mode of innovating, or whether RI should adapt its theory to make it more practicable and realistic. We propose a combination of both.

RI researchers, we suggest, have a lot to learn from Datashare's approaches to stakeholder involvement. Datashare's team leaders had 'reasonable reasons' to reject our RI-informed workshop intervention and showed us different ways of involving stakeholder perspectives that they felt were more sustaining of the overall project in the long-run. Their careful negotiation of stakeholder perspectives and iterative and parallel development trajectories suggest an alternative to ideal-type RI approaches. This alternative approach is more fitting for a flexible and uncertain setting where antagonistic public debates, multiple opposing stakeholder groups and complex digital infrastructures seek to strike a balance in ways that cannot be predicted beforehand.

Yet, Datashare's strategies for dealing with stakeholder tensions also have features that sit uneasy with RI's goal of securing sustainable commitment to innovations by multiple stakeholders; *and* RI's premise that the success of this depends on the extent to which this involvement is genuinely inclusive. According to RI, careful nurturing of inclusive commitment generates the robustness necessary for dealing with strong antagonisms that can reasonably be expected to challenge sociotechnical systems in the future. This brings to light two issues for the RI-research agenda.

13.5.1 RI Without Strong Stakeholder Commitment

Datashare's strategy of 'controlled ambiguity' is successful at balancing multiple perspectives and engages both market players and privacy activists. The advantage for Datashare of this approach is that, for its continuation, it was not dependent on the continuous commitment of stakeholders to progress. In the absence of a unifying vision, Datashare as well as stakeholders have an opt-out possibility without the whole project collapsing. In addition, its flexible, *lean* way of innovation prevented Datashare from being tied to a pre-defined plan and enabled the team to more easily

divert from particular trajectories than if it had committed to a more specific goal or vision agreed upon by various stakeholders. The advantage for Datashare was that this enabled the organisation to demonstrate its potential to shareholders on a regular basis.

However, from an RI perspective this flexibility and ambiguity leave open the question of the sustainability of the moral orientation of the organisation. By not committing to a single vision and maintaining multiple representations of its vision i.e. by not making any promises Datashare has little accountability towards their stakeholders, other than their shareholders. This begs the question what happens if the tension cannot be resolved or if the team changes? Will the market players in the end determine the conditions for data-sharing or is it up to the privacy activists to define its trajectory? Both options have different consequences that will need to be decided upon: if *Own* were to live up to the radical visions of data control held by some of these activists it would have to rely on decentralised infrastructures of storage and control, demanding a lot of technical expertise both from residents and third-parties. If *Own*, on the other hand, were to live up to the commercial vision of data control held by businesses, its most important feature would be to merely create a 'sense of trust', without limiting data-sharing and repurposing. Due to this lack of stakeholder commitment and a shared unifying vision, the extent to which Datashare's products and services will be responsive to these questions primarily depends on the responsiveness and the ethics of the team. Whatever the team decides can only be challenged or averted by leaving the system, even if this poses financial or social disadvantages to resident-users. From an RI-perspective, this would be a less-democratic way of managing the relationship between users and sociotechnical systems. The challenge for the RI-research agenda is, thus, to cope with innovative contexts in which not all stakeholders are strongly committed: what tools and methods can enable companies to innovate responsibly in a sustainable way in the absence of this strong stakeholder commitment?

13.5.2 Balancing Inclusion

Datashare's strategy of mediating and translating stakeholder visions to each other, in a tightly controlled way, gives them several advantages: it protects the fragile stakeholder relationships by not exposing them to confusing and contradictory perspectives and by tuning into the specific needs and contexts of each of the stakeholders. From an RI perspective, however, the flipside of this approach is that not all stakeholders are granted an equally strong voice in the innovation process and that it narrows the scope for exploring ways of deliberating with stakeholders. The approach favours business partners with aligned interests as conversation and collaboration partners in the development of the business concept. They are enlisted to help frame the problem and its possible solutions, and to validate whether an idea solves a particular problem.

Residents or users, as we saw, play a different role: as part of the differing stories told to different stakeholders, they are objectified and made to represent different values. Where their opinion is solicited, it is done in a rather restricted way: they are allowed to pick one of a limited number of options, but not to contribute to the framing of the problem. This restricts their capacity to participate in deliberations as subjects capable of voicing their own concerns and setting their own terms for the discussion outside of pre-established interests. Datashare's team members, a few professional academics and critical activists are brought in as mediators for potential users, but there are few channels through which these stakeholders can speak for themselves about the definition of problems and their solutions. Moreover, their mediators are used as informants, but not as co-decision makers and there is minimal critical reflection on the representativeness of these mediators. As a result, powerful stakeholders have more weight in the decision-making process: only those with financial resources or exit threats, are really listened to.

As the literature on RI has stressed, the risk of excluding certain stakeholders from the decision-making process is that eventual products or services may not match well with the needs of more vulnerable groups. Another challenge for the RI research agenda is, therefore, to develop ways of making multiple voices, especially of weaker stakeholders, part of the decision-making processes, while enabling innovators to perform their balancing act between stakeholders. For instance, one avenue to explore is the role and position of mediators or translators in the decision-making process.

To conclude, while acknowledging the good reasons for not committing to the ideal-type setting of RI in the organisation of Datashare's innovation process, we observe a need to explore additional and alternative ways for addressing these issues of stakeholder commitment and inclusion. If only for the reason that granting data control to user-residents is a worthwhile goal from RI's perspective, and that it would be a shame if this project failed because of a lack of sustainable commitment or fall-out.

References

Barta, Kristen, and Gina Neff. 2016. Technologies for sharing: lessons from quantified Self about the political economy of platforms. *Information, Communication & Society* 19 (4): 518–531.

Blank, Steve. 2013. Why the lean start-up changes everything. *Havard Business Review* 91 (5): 64–73.

Blok, Vincent, and Pieter Lemmens. 2015. The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation. In *Responsible Innovation 2*, ed. Bert-Jaap Koops et al., 19–35. Heidelberg: Springer.

Blok, Vincent, L. Hoffmans, and E.F.M. Wubben. 2015. Stakeholder engagement for responsible innovation in the private sector: critical issues and management practices. *Journal on Chain and Network Science* 15(2): 147–164.

Chalmers, David, Rebekah E. McWhirter, Dianne Nicola, Tess Whitton, Margaret Otlowski, Michael M. Burgess, Simon J. Foote, Christine Critchley, and Joanne L. Dickinson. 2014. New avenues within community engagement: Addressing the ingenuity gap in our approach to

health research and future provision of health care. *Journal of Responsible Innovation* 1 (3): 321–328.

Davies, Sarah R. and Maja Horst. 2015. Responsible innovation in the US, UK and Denmark: governance landscapes. In *Responsible innovation, volume 2: Concepts, approaches, and applications*, ed. B.J. Koops, I. Oosterlaken, J. van den Hoven, H.A. Romijn, and T.E. Swierstra. Dordrecht: Springer.

Delgado, Anna, Kamilla Lein Kjølberg, and Fern Wickson. 2010. Public engagement coming of age: From theory to practice in STS encounters with nanotechnology. *Public Understanding of Science* 20 (6): 826–845. doi:10.1177/0963662510363054.

Elster, Jon, ed. 1998. *Deliberative democracy*. Cambridge, MA: Cambridge University Press.

Felt, Ulrike, and Brian Wynne. 2007. *Taking European knowledge society seriously*. Report of the Expert Group on Science and Governance to the Science, Economy and Society Directorate. Directorate General for Research, European Commission.

Grunwald, Armin. 2014. Technology assessment for responsible innovation. In *Responsible innovation 1*, 15–31. Dordrecht: Springer Netherlands.

Guston, David H., and Daniel Sarewitz. 2002. Real-time technology assessment. *Technology in Society* 24 (1): 93–109.

Gutmann, Amy, and Dennis Thompson. 1996. *Democracy and disagreement*. Cambridge, MA: Harvard University.

Habermas, Jürgen. 1990. *Moral consciousness and communicative action*. Trans. C. Lenhardt. Cambridge, MA: MIT.

———. 1996. *Between Facts and Norms*. Trans. W. Rehg. Cambridge, MA: Polity Press.

Horst, Maja. 2007. Public expectations of gene therapy scientific futures and their performative effects on scientific citizenship. *Science, Technology & Human Values* 32 (2): 150–171.

Kuhlmann, Stefan, Ordonez Matamoros, Hector Gonzalo, Bart Walhout, Dorbeck-Jung, R. Barbel, Jakob Edler, Sally Randles, Sally, Gee, Elena Pariotti, Guido Gorgoni, and Simone Arnaldi. 2016. Responsible research and innovation in a distributed anticipatory governance frame. A Constructive Socio-normative Approach. Deliverable D4.8. Interim design requirement report. ResAGorA.

Lövbrand, Eva, Roger Pielke, and Silke Beck. 2010. A democracy paradox in studies of science and technology. *Science, Technology & Human Values* 36 (4): 474–496.

Lund Declaration. 2009. Europe must focus on the grand challenges of our time. Swedish Presidency Research. https://era.gv.at/object/document/130. Accessed 3 Aug 2016.

Owen, Richard, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and Dave Guston. 2013. A framework for responsible innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. R. Owen, J. Bessant, and M. Heinz, 27–50. Chichester: Wiley.

Pandza, Krsto, and Paul Ellwood. 2013. Strategic and ethical foundations for responsible innovation. *Research Policy* 42 (5): 1112–1125.

Ries, Eric. 2011. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. New York: Crown Business.

Siune, Karin, Eszter, Markus, Marina, Calloni, Ulrike, Felt, Andrzej Gorski, Armin Grunwald, Arie Rip, Vladimir de Semir, Sally Wyatt. 2009. Challenging Futures of science in society. *Emerging trends and cutting-edge issues*. Brussels: MASiS Expert Group, European Commission. https://ec.europa.eu/research/science-society/document_library/pdf_06/the-masis-report_en.pdf. Accessed 3 Aug 2016.

Stark, David. 2011. *The sense of dissonance: Accounts of worth in economic life*. Princeton: Princeton University Press.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42 (9): 1568–1580.

Stirling, Andy. 2008. "Opening up" and "closing down" power, participation, and pluralism in the social appraisal of technology. *Science, Technology & Human Values* 33 (2): 262–294.

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.

Van der Burg, Simone, and Tjsalling Swierstra, eds. 2013. *Ethics on the laboratory floor*. Basingstoke: Palgrave Macmillan.

Van den Hoven, Jeroen, Gert-jan Lokhorst, and Ibo van de Poel. 2012. Engineering and the problem of moral overload. *Science and Engineering Ethics* 18 (1): 143–155.

den Hoven, Van, Neelke Doorn Jeroen, Tsjalling Swierstra, Bert-Jaap Koops, and Henny Romijn, eds. 2014. *Responsible innovation 1: Innovative solutions for global issues*. Dordrecht: Springer Netherlands.

Van Oudheusden, Michiel. 2014. Where are the politics in responsible innovation? European governance, technology assessments, and beyond. *Journal of Responsible Innovation* 1 (1): 67–86.

Von Schomberg, Rene. 2011. Prospects for technology assessment in a framework of responsible research and innovation. In *Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden*, ed. M. Dusseldorp and R. Beecroft, 39–61. Wiesbaden: VS Verlag für Sozialwissenschaften.

———. 2013. A vision of responsible research and innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society*, ed. R. Owen, M. Heintz, and J. Bessant, 51–74. Chichester: Wiley.

———. 2014. The quest for the 'right' impacts of science and technology: a framework for responsible research and innovation. In *Responsible innovation 1*, ed. J. van den Hoven et al., 33–50. Dordrecht: Springer Netherlands.

Wickson, Fern, Ana Delgado, and Kamilla Kjølberg. 2010. Who or what is 'the public'? *Nature Nanotechnology* 5 (11): 757–758.

Wickson, Fern, and Ana L. Carew. 2014. Quality criteria and indicators for responsible research and innovation: Learning from transdisciplinarity. *Journal of Responsible Innovation* 1 (3): 254–273.

About the Authors

Huib Aldewereld is lecturer at the University for Applied Sciences in Utrecht. Before that, he worked as an assistant professor at Delft University of Technology, in the field of value sensitive software design.

Lotte Asveld is an assistant professor at Delft University of Technology studying the societal aspects of biotechnology. Her main research interests concern responsible innovation in the field of biotechnology and synthetic biology: how can the societal debate on biotechnology and synthetic biology be integrated in innovation trajectories? Lotte has worked as a researcher in the department of Philosophy at DUT, where she also received her PhD. Lotte also worked as a researcher at the Rathenau Insituut, focusing on the bioeconomy, and as a freelance researcher in China.

Vincent Blok is associate professor in Business Ethics and Responsible Innovation at the Management Studies Group, and associate professor in Philosophy of Management, Technology & Innovation at the Philosophy Group, Wageningen University (The Netherlands). Blok's research group is involved in several (European) research projects at the crossroads of business, philosophy and innovation. Blok's work has appeared in *Journal of Business Ethics, Journal of Cleaner Production, Journal of Agricultural and Environmental Ethics* and *Journal of Responsible Innovation* and other journals. See www.vincentblok.nl for more information about his current research.

Jacqueline Broerse is full professor of Innovation and Communication in the Health and Life Sciences (in particular addressing issues of diversity and social inclusion) and director of the Athena Institute, Vrije Universiteit Amsterdam. She has a master in Biomedical Sciences and a PhD in Science, Technology and Society studies. Her research interests focus on designing and studying multi-stakeholder innovation processes. Jacqueline has widely published on multi-stakeholder participation in health (system) innovations.

L. Asveld et al. (eds.), *Responsible Innovation 3*,
DOI 10.1007/978-3-319-64834-7

Aad Correljé is attached to the Economics of Infrastructures section of the Faculty Technology, Policy and Management, TU Delft, as an associate professor. He is a research fellow with the Clingendael International Energy Programme (CIEP) of the Netherlands institute for International Affairs Clingendael. He is a Member of the Editorial Board of the academic journal *Energy Policy* and an instructor at the Florence School of Regulation.

Eefje Cuppen is Associate Professor at the department of Multi Actor Systems in the Faculty of Technology, Policy and Management at Delft University of Technology. Her research focuses on public engagement, participatory governance and responsible innovation in the field of energy and sustainability. She studied Innovation Sciences at Eindhoven University of Technology, and did her PhD at the Institute for Environmental Studies (VU Amsterdam) on methodology for stakeholder participation.

Virginia Dignum is Associate Professor on Social Artificial Intelligence at the Faculty of Technology Policy and Management at TU Delft. Her research focuses on value-sensitive design of intelligent systems and multi-agent organisations, in particular on the formalisation of ethical and normative behaviours and of social interactions. In 2006, she was awarded the prestigious Veni grant from NWO (Dutch Organization for Scientific Research) for her work on agent-based organizational frameworks. She has wide experience with obtaining research funding both at national as international level, and has reviewed several EU and national proposals and projects.

Neelke Doorn is associate professor Ethics and Philosophy of Technology at the Technical University Delft, the Netherlands. She holds master degrees in civil engineering, philosophy and administrative law. Her current research concentrates on moral and distributive issues in water governance. In 2013, Neelke was awarded a prestigious Veni-grant for outstanding researchers from the Netherlands Organization for Scientific Research (NWO) for a project on the ethics of flood risk management. She was shortlisted for the engineer of the year award 2014 from the Dutch professional engineering organization KIVI for her work on the interface of ethics and engineering.

Ulrike Felt is Professor of Science and Technology Studies, Dean of the Faculty of Social Sciences, Head of the interfaculty research platform "Responsible research and innovation in academic practice" and founding director of the STS Master at the University of Vienna. Her research focuses on governance, democracy and public participation, on shifting research cultures as well as on the role of time and future at the interface of research and society. Her areas of study cover life science and (bio)medicine, nanotechnology and material sciences, new technologies and sustainability research. From 2002–2007 she was editor-in-chief of *Science, Technology, & Human Values*. She was the leading the editorial team of the most recent Handbook of Science and Technology Studies (MIT Press 2017).

Elisabeth van de Grift holds a BSc in journalism and an MSc in cultural anthropology. Her ethnographic research focused on social and ideological dynamics among members of a local sustainability initiative in the United Kingdom. Before Elisabeth started her PhD on controversial energy technologies, she worked as a communications consultant for a variety of organizations, like UNEP-WCMC and the CGIAR's Climate Change, Agriculture and Food Security programme.

Léon Jansen obtained a MSc in food technology, a PhD in toxicology from Wageningen University (the Netherlands) and performed post-doc research at the WHO International Agency for Research on Cancer (Lyon, France) and at the Dutch Cancer Institute (Amsterdam, The Netherlands). For 5 years, he was the toxicologist of the Netherlands Nutrition Centre. In 2006 he joined the Schuttelaar & partners and worked on the launch of the Dutch Choices Foundation. Since then, he is scientific coordinator of Choices International Foundation and the Dutch Choices foundation.

Pim Klaassen is assistant professor of Policy, Ethics and Communication in the Health and Life Sciences at Vrije Universiteit Amsterdam. He obtained his PhD in Philosophy of Science at the University of Amsterdam and his MPhil in History, Philosophy and Sociology of Science, Technology and Medicine at the University of Cambridge (UK). Currently he works at the Dutch National Institute for Public Health and the Environment on an EC-funded project aimed at developing a roadmap for implementing RRI in industry (PRISMA).

Frank Kupper is assistant professor in Science Communication at the Athena Institute, Vrije Universiteit Amsterdam. He is interested in sense-making processes and dialogue at the interface of science and society. His research focuses on the development of creative methodologies and tools to build reflexive conversations about science in society. He participated in various responsible innovation projects, such as VOICES, SYNENERGENE, RRI Tools and the MVI project Neurosciences in dialogue. He has much experience as a facilitator of learning and change in social innovation projects. Currently, he assists the Dutch parliamentary TA office, the Rathenau Institute, innovating their methods for citizen and stakeholder engagement.

Saskia Lavrijssen is Professor of Economic Regulation and Market Governance of Network Industries at Tilburg University. Lavrijssen specializes in the regulation of the energy and water sectors. Prof. Lavrijssen's research, which is embedded in TILEC (Tilburg Law and Economics Center), addresses the role of domestic and non-domestic consumers in the network sectors, including the water and energy sectors. Prof. Saskia Lavrijssen (b. 1976, Waalre) was (part-time) professor of Consumer and Energy Law at the University of Amsterdam after having worked as an associate professor of Public Economic Law at Utrecht University. She has been teaching at Tilburg Law School since April 2014 as associate professor.

Dr. C.A. (Kees)Linse is Chairman of the Netherlands Commission for Environmental Assessment. He holds Supervisory Board positions in TNO (Netherlands Organisation for Applied Physics), the Leiden University Medical Centre, and in industry. He graduated in Chemistry in Utrecht in 1972, got his PhD in Physics in Leiden in 1978 and worked for Shell for 32 years.

Rob Lubberink is a PhD candidate at the Management Studies Department of the Wageningen University. His PhD research focuses on Responsible Innovation in the context of Social Entrepreneurship. Rob started his PhD after graduating in Management, Economics and Consumer studies at the Wageningen University in 2013.

Merel Noorman is a post-doc researcher at Maastricht University. She has a background in artificial intelligence and science and technology studies. In her PhD dissertation she analysed the role of metaphorical concepts, like autonomy and social interaction, in research on artificial agents. Since completing her PhD, her research interests center on the ethics of intelligent technologies, and in particular on the distribution of responsibility around these technologies. Her current research project focuses on long-term equal stakeholdership in smart city technological innovation.

Onno Omta graduated in biology in 1978 and after a management career he defended his PhD thesis on the management of innovation in the pharmaceutical industry in 1995 (both at the University of Groningen). In 2000 he was appointed as chaired professor in Business Administration at Wageningen University. He is (co-) author of many scientific articles on innovation management. His current research interest encompasses entrepreneurship and innovation in chains and networks in the life sciences.

Johan van Ophem is associate professor at Wageningen University. His teaching and research relate to the field of consumer behaviour, lifestyles, health economics and household economics. He has published in these fields and he has edited several publications in the field of economics and sociology of the household. He is (co)-organizer of various scientific conferences and an active member of European university networks.

Edwin Pietersma currently works in the Food Industry in a R&D related position. He has a master degree in Food Technology with a specialisation in Innovation and Management from Wageningen University. His interest in understanding the drivers of certain decisions helped him during his master thesis in investigating ethical decision making in food companies.

Klara Pigmans is a PhD student at the faculty of Technology, Policy and Management at TU Delft, focusing on the role of social and moral values in complex decision-making processes. She has a Master's degree in Business Informatics and a Bachelor's degree in Information Science. Previously, she was employed at the International Institute for Communication and Development (IICD) in The Hague. During her PhD research, she also initiated and founded the 'KNAW Hendrik Muller Network' for PhD students who like to reflect on broader, societal issues in academia (such as integrity, academic freedom, diversity).

Eugen Popa studied Communication at the National University of Political Sciences and Public Administration in Bucharest and Argumentation Theory at the University of Amsterdam. He obtained his doctoral title in 2016 with a dissertation titled Thought experiments in academic communication. He currently holds a position as postdoctoral researcher and lecturer at the department of Science Communication in the Athena Institute, Vrije Universiteit Amsterdam.

Steve Rayner is James Martin Professor of Science and Civilization and Director of the Institute for Science, Innovation and Society at Oxford University. He previously held senior research positions in two US National Laboratories and has taught at leading US universities. He has served on various US, UK, and international bodies addressing science, technology and the environment, including Britain's Royal Commission on Environmental Pollution, the Intergovernmental Panel on Climate Change and the Royal Society's Working Group on Climate Geoengineering and the Lead Expert Group for the British Government's Foresight Programme on Future Cities. Until 2008 he also directed the national Science in Society Research Programme of the UK Economic and Social Research Council. He was included in the 2008 Smart List by Wired Magazine as "one of the 15 people the next US President should listen to".

Michelle Rijnen earned her bachelor degree in biomedical sciences, after which she studied bioethics and healthcare policy in her MSc in Management, Policy analysis and Entrepreneurship in the Health and Life Sciences. She worked as a junior researcher at the Athena Institute, Vrije Universiteit Amsterdam, in assignment of the FP7-funded RRI Tools project. Currently, she works as a policy advisor in patient participation at the Dutch Cancer Society.

Tsjalling Swierstra is head of the Philosophy Department at the University of Maastricht and adjunct professor at NTNU in Trondheim. He is initiating co-founder of the *Journal for Responsible Innovation*. He has published widely on the ethics of new and emerging science and technology (NEST-ethics), on the 'soft impacts' of technology, and on technomoral change, i.e. the mutual shaping of science and technology and morals.

Behnam Taebi is associate professor in ethics of technology at Delft University of Technology, and associate with the Harvard Kennedy School's Belfer Center for Science and International Affairs. His research interests are in energy ethics, nuclear ethics, responsible innovation and engineering ethics. He studied Material Science and Engineering (2006) and received his Ph.D. in Philosophy of Technology (2010). His research interests are in energy ethics, nuclear ethics, responsible innovation and engineering ethics. Taebi is a member of The Young Academy of the Royal Netherlands Academy of Arts and Sciences.

Tjidde Tempels is a Ph. D. candidate at the Philosophy group of Wageningen University, The Netherlands. He holds a BSc and MSc degree in Political Science from Radboud University Nijmegen and Master's degree in Philosophy from Utrecht University. His research interests span business ethics, political philosophy, public health ethics and (political) corporate responsibility. In his doctoral project he explores the grounds for corporate responsibility for public health within the context of the food and beverage industry.

Job Timmermans, MSc/MA is a PhD candidate on Responsible Research and Innovation in ICT at the Centre for Computing and Social Responsibility at De Montfort University, Leicester, UK. His interests cover ethical issues arising from

the intersection of technology and society. Job has worked within the Framework for Responsible Research & Innovation in ICT (FRRIICT) project, the FP7 Governance for Responsible innovation (GREAT) project and Ethical Issues of Emerging ICT Applications (ETICA) project.

Federico Vasen is a postdoctoral fellow at the Institute for Social Research, National Autonomous University of Mexico, Mexico City (IIS-UNAM). He holds a BA in Philosophy from the University of Buenos Aires and a Ph.D. in Social Sciences from the National University of Quilmes. His research interests include the history of science and technology policy in Argentina and Latin America and the evolution of the conceptual frameworks attached to it. He has also published on the risk and reguation of emerging technologies (stem cell therapies and agricultural biotechnology) in Argentina. He has taught philosophy of science and technology to undergraduate engineering students at the Argentine Catholic University in Buenos Aires and STS in the graduate program on Philosophy of Science at UNAM, Mexico.

Rietje van Dam-Mieras is Professor emeritus at Leiden University. She graduated in chemistry and obtained her PhD in Biochemistry at Utrecht University. She subsequently worked at Maastricht University, the Open University of the Netherlands and Leiden University. She had a UNESCO Chair at the Open University and was visiting professor at the United Nations University Institute of Advanced Studies. She is a founding member of RCE Rhine-Meuse and has been actively involved in the RCE-initiative of United Nations University (UNU) in the context of the UN Decade Education for Sustainable development (UN DESD 2005–2014).

Jeroen van den Hoven is full professor of Ethics and Technology at Delft University of Technology and editor in chief of Ethics and Information Technology. He was the founding scientific director of 3TU. Centre for Ethics and Technology (2007–2013). In 2009, he won the World Technology Award for Ethics as well as the IFIP prize for ICT and Society for his work in Ethics and ICT. Jeroen van den Hoven was founder, and until 2016 Programme Chair, of the Dutch Research Council on Responsible Innovation. He is also member of the TU Delft Blockchain Lab.

Sara Vermeulen MA is a researcher and lecturer in the department Communication and Information Studies at Radboud University, with specialisation in Responsible Research and Innovation. She previously worked as a researcher and lecturer at Vrije Universiteit Amsterdam and as a consultant for Technopolis Group Amsterdam. With an academic background in Political and Social Philosophy and a strong interest in Science and Technology studies, her work examines the implications of various concepts of responsibility in research and innovation practices.

Dorien Zandbergen is a comparative researcher of digital history, culture and politics at the University of Amsterdam. She did studies on the countercultural background of Silicon Valley digital culture, the Quantified Self movement, and on two collaborative air quality sensing projects in Amsterdam and London. She also co-produced a documentary on the aspiration of the city of Amsterdam to become a Smart City and is founder of Gr1p, a non-profit aiming to find ways to help diverse publics come to grips with the politics and materiality of complex information technologies.