Research Policy 47 (2018) 61–69

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Research paper

Collingridge and the dilemma of control: Towards responsible and accountable innovation

Audley Genus[a,*], Andy Stirling[b]

[a] Kingston Business School, Kingston University, Kingston upon Thames, KT2 7LB, United Kingdom
[b] Science Policy Research Unit, University of Sussex, Falmer, Brighton, BN1 9RH, United Kingdom

ARTICLE INFO

Keywords:
Accountability
David Collingridge
Dilemma of control
Governance of innovation
Incrementalism
Responsible innovation

ABSTRACT

The paper critically reviews the work of David Collingridge in the light of contemporary concerns about responsibility and accountability in innovation, public engagement with science and technology, and the role of scientific expertise in technology policy. Given continued interest in his thoughts on the 'social control of technology', and the 'dilemma of control', this attention is both timely and overdue. The paper illuminates a mismatch between the prevalence of citations to Collingridge's work on the dilemma of control in the literature on responsible innovation, and the depth of engagement with his arguments. By considering neglected aspects of Collingridge's substantive, methodological and philosophical analysis, important implications can be drawn for theory and practice relating to the governance of innovation and co-evolution between technology and society. The paper helps to improve understandings of wider political contexts for responsible innovation, especially in relation to anticipatory, participatory and institutional aspects of governance.

1. Introduction

David Collingridge was an important contributor to the field of Science and Technology Studies. An active researcher from the late-1970s to the early-1990s, he developed a distinctive and substantive line of thinking concerning the 'social control of technology' – the scope for conventionally excluded people and interests to shape the forms and orientations of innovation trajectories in particular sectors. Collingridge's work was concerned with increasing social agency over technology – away from incumbent interests in what have come to be called 'innovation systems'. It drew on a critique of what he called the "justificationist school" in decision-making, philosophy, and political science. Language fashions change, and now more is spoken of the 'governance' of science, technology or innovation than of 'social control'. The notion of control itself can be viewed as being as problematic as it is helpful to Collingridge's underlying aims (Smith and Stirling, 2010; Stirling, 2016a). And a succession of vocabularies that have burgeoned in this field largely since Collingridge's work (for instance, around 'precaution', 'participation' and 'engagement'), are now being substituted in some quarters by the language of responsible (research and) innovation (Guston et al., 2014; Owen, 2014; Owen et al., 2012, 2013a,b; Stahl, 2012; Stilgoe et al., 2013, 2014; Von Schomberg, 2011, 2015; see: https://www.epsrc.ac.uk; Stirling, 2016b).

Responsible innovation is not a new concern. Academic and policy discussions over responsibilities, risks and 'control' in the governance of science and technology go back many years (Donnelley, 1989; Jonas, 1984). Over 70 years ago landmark contributions to understandings of the history of science and relationships among science and society were produced by Bernal (1939, 1954) and Haldane (1939). In the 1960s, significant contributions emphasised the social responsibility of science and greater public understanding of science (e.g. Rose and Rose, 1969; Ziman, 1968). The British Society for Social Responsibility in Science ran from 1969 to 1991 whilst its US equivalent was active between 1949 and 1976. These societies were concerned with the transparency and openness of science policy-making as well as the environmental and health consequences associated with the operation of new technologies implicated in scientific discoveries and inventions.

There have been longstanding concerns regarding the framing and promotion of scientific citizenship or 'citizen science' (Irwin, 1995, 2001), which now figure in discussions of responsible research and innovation (see the special issue of Public Understanding of Science, edited by Stilgoe et al., 2014). An important development was the emergence of the ELSI approach towards the end of the 1980s, connected with the ethical, legal and societal implications of, for instance, the Human Genome project and the proper governance of scientific research and technological innovation. More recently, there have been concerns to address the limitations of the ELSI approach and to elicit more active or responsive public engagement with science and

* Corresponding author.
E-mail addresses: a.genus@kingston.ac.uk (A. Genus), A.C.Stirling@sussex.ac.uk (A. Stirling).

http://dx.doi.org/10.1016/j.respol.2017.09.012
Received 25 February 2016; Received in revised form 10 February 2017; Accepted 25 September 2017
Available online 05 October 2017

technology (Rose, 2012).

In the UK, a landmark parliamentary select committee report recognised citizens' distrust of conventionally-institutionalised science, following a number of high profile technological controversies and crises which brought into question the accountability and autonomy of science and the role of society therein (House of Lords, 2000). Moreover, it represented a moment of reflection on what it meant for science and innovation to be conducted 'responsibly'. As subsequent discussion has shown, this 'responsibility' may be understood in relation to what may be considered proper behaviour – or 'responsiveness' – in defined contexts (i.e. the ethical dimension of responsible innovation), and in terms of who is to be held 'responsible' for what ('role responsibility').

Temporally, consequentialist notions of accountability on the part of science/scientists and innovation/innovators pertain to whether and how they should be held to account for the consequences of actions taken in the past (Grinbaum and Groves, 2013; Stilgoe et al., 2013). But they also concern more immediate democratic accountabilities for the contemporary driving motivations of innovation processes (Lee and Petts, 2013) – with respect to impacts which may not have been reasonably foreseeable with any particularity, but which nonetheless arise from the prioritisation in any given setting of some kinds of general orientation for innovation rather than others (Wynne, 1993; 2002). While some analyses of responsible research and innovation (RRI) include consideration of accountability (Lee and Petts, 2013; Sutcliffe, 2011), others are explicitly critical of what is portrayed as the "traumatic" and "infantilising" effects on processes of innovation (Grinbaum and Groves, 2013). For the most part, the RRI literature tends to neglect the quality of accountability. The associated relevance of democracy to the orientation of technology is also thereby downplayed. Collingridge's emphasis on Popperian qualities of openness and his development of specific qualities around which to structure accountability offer crucial but under-realised values for understanding responsible research and innovation.

Putting problems of accountability to the fore, this paper seeks to identify aspects of Collingridge's work that have not been especially well assimilated or further developed in attempts to implement a framework for responsible governance of research innovation. More fruitful analysis would draw on Collingridge's contribution to engage more strongly with accountability in debates bearing on key elements of responsible innovation, such as anticipatory technological decision-making and inclusive deliberation. Within this discussion, Collingridge's ideas are themselves subject to productive critique, which may also helpfully inform understanding and practice of RRI.

The paper has the following structure. Section 2 outlines developments in practice and research connected with building and institutionalising frameworks for responsible innovation. Section 3 discusses how Collingridge's work addresses core concerns of RRI, focusing initially on the 'Collingridge dilemma' and the 'corrigibility' of technology. Section 4 critically reviews the contribution of Collingridge's work to RRI. Section 5 considers revisions and extensions that may be made to Collingridge's contribution which have the potential to improve understandings of responsible innovation. Section 6 concludes the paper, reflecting on implications of the foregoing sections for future research on – and practice of – more responsible and accountable innovation.

2. The practice and a framework of responsible innovation

Multiple labels, approaches and genealogies alluded to in different areas of the responsible research and innovation literature make this subject difficult to characterise in any definitive way. What is clear is that an expanding body of analysis and policy practice has built up over the last decade around notions of "responsibility" in this area. This may be identified with a number of key journal articles such as by Guston (2006), Guston and Sarewitz (2002), Roco et al. (2011) and Stilgoe et al. (2013), the book by Owen et al. (2013a), and the launch of the Journal of Responsible Innovation (Guston et al., 2014). Reading across multiple definitions Wickson and Carew (2014: 255) identify the following core characteristics of RRI: a focus on addressing 'significant socio-ecological needs and challenges'; a commitment to actively engage a range of stakeholders for the purpose of substantively improving decision-making and mutual learning; a dedicated attempt to anticipate potential problems; and a willingness among all participants to act and adapt according to these ideas.

The EU employs RRI as a cross-cutting theme within its Horizon 2020 funding framework (alongside 'science with and for society'). Horizon 2020 is itself a core element of the flagship Innovation Union programme, which in turn is a central aspect of the EU2020 strategy. Taken as a whole, EU initiatives and policies tend to characterise innovation in an undifferentiated way – as a self-evidently generally "good thing" irrespective of the specific kind of innovation involved or the alternatives that might thereby be foreclosed (Stirling, 2014). Thus "pro-innovation" policies are prized as a means to 'smart growth', which is in turn seen simply in terms of the gross numbers of jobs involved – rather than in terms of net comparisons with numbers and kinds of jobs that would be created by the same investments by other means (Stirling, 2010a).

In this view, innovation is whatever happens to emerge from incumbent structures of interest, privilege and power in prevailing innovation systems (Stirling, 2008). Justification is provided by reference to variously direct or indirect engagement with societal challenges, such as those connected with promoting green and secure energy, food security, climate action, and 'smart' transport. But these "solutions" are typically addressed by starting first with the incumbent innovation trajectory and simply highlighting those problems that it may promise to address. Far less attention is given to any analysis starting with the challenges themselves, in order to decide which innovation trajectories might be most appropriate (Felt et al., 2008).

Of course, this highly pressurized and expedient approach needs to be understood in relation to EU concerns about the need to close the "competitiveness gap" with other global economic blocs and countries, notionally by increasing R & D (Felt et al., 2013). And it is here that it is relevant that 'responsible innovation' has also been strongly invoked (under similar dynamics) in US policy on research and innovation – notably in the governance of nanotechnology. Examples here are the National Nanotechnology Initiative's strategic plan (NNI, 2011) and the National Science Foundation's Nanotechnology in Society network (Roco et al., 2011).

Such initiatives raise a key point regarding the emergence of responsible innovation. This concerns the potential for understandings of RRI to become unduly attenuated or instrumentalised, resulting in more attention being devoted to deciding on how to implement an incumbent innovation pathway, than on choosing which pathway to follow (STEPS Centre, 2010). To address the shortcomings of these more instrumentalised approaches, it is often urged that responsible innovation move beyond preoccupation with research and development and economic benefits of individual technologies to address the innovation process more fully, including social as well as technical and other aspects (Blok and Lemmens, 2015; Von Schomberg, 2015).

Whilst orientations and emphases vary, four resulting interacting dimensions are highlighted as means by which RRI might mitigate such criticisms. First, RRI aims to be anticipatory in the sense of exploring possibilities (not making predictions) and analysing 'intended and potentially unintended impacts that might arise'. It aims to be 'deliberative' – 'inclusively…inviting and listening to wider perspectives from publics and diverse stakeholders'. It prioritises 'reflectiveness' regarding 'underlying purposes, motivations and potential impacts'. And finally, RRI is argued to be 'responsive', 'using this collective process of reflexivity to both set the direction and influence the subsequent trajectory and pace of innovation' (Owen et al., 2013b: 38; see also Stilgoe et al., 2013). Among other issues, these processual aspirations of RRI raise a number of implications for the accountability of research and

innovation.

First, proponents of RRI are concerned about the limitations of risk-oriented approaches in providing reliable 'early warnings' of potentially deleterious effects of new technologies (Harremoës et al., 2000). Here, they emphasise instead anticipatory approaches involving expert studies of diverging futures and socio-technical imaginaries. In this view, anticipation in responsible innovation involves an intermediate position between the 'closing down' of governance through the making of predictions and promises and the full 'opening up' of spaces for more direct forms of public accountability in substantive citizen participation. (Stirling, 2008).

The timing of accountabilities is also an important aspect of anticipation, with RRI enjoining effective early instigation of "upstream" governance processes. Whether referred to as deliberation (Owen et al., 2013b) or inclusion (Stilgoe et al., 2013), responsible innovation prioritises the admission of 'new voices' to the governance of science and innovation. A plethora of divergent forms of public engagement have emerged. Yet these have been criticised for remaining marginal to governance of science and innovation, with the key commitments undertaken elsewhere, often even further 'upstream' in innovation processes (Conrad et al., 2011). A further risk of erosion in accountability arises here in tendencies to favour designs for "invited" forms of public engagement (Wynne, 2007) that reinforce (rather than fully interrogate) political closures (Chilvers, 2008).

For its part, reflexivity is a key quality in inclusive deliberation. Here, RRI literatures are informed by Beck's (1992; 1999) seminal work on reflexive modernisation, in which he argues that increasing difficulties in calculating scientific and technological risks will lead scientists themselves to become more reflexive – heralding what he calls a 'second modernity'. In this vein, Stilgoe et al. (2013) draw on work by Wynne (1993, 2002) to argue for 'institutional reflexivity' among funders, regulators and users of scientific research concerning the assumptions and practices implicated in science and innovation and their governance. But in order to contribute to the 'opening up' of accountabilities, reflexivity must be more than a private process of self-questioning regarding values and interests in science and innovation. Rather than a quality located in individual social actors, reflexivity needs to be recognised as a plural and distributed social capability (Stirling, 2016a). In this sense, reflexivity is a public practice, capacities for which may be enabled by application of standards and codes of conduct (Voss et al., 2006).

In all these ways and more, RRI involves mutual accommodation and adjustment in the needs, interests and values of contending "stakeholders". In particular, the quality of 'responsiveness' – for instance according to Owen et al. (2013b: 38) – involves an 'open process' of 'adaptive learning'. In such ways, interwoven principles of anticipatory governance and inclusive deliberation help confer mutual responsiveness on the part of participating interests – and enjoin reflexivity over the positions and values that they themselves and others hold. But again, there tends to be relatively little explicit, direct, or substantive provision for addressing power dynamics in these processes of adaptation, learning and responsiveness. It is these kinds of power dynamic that are addressed in wider notions of political accountability.

With these dimensions of RRI in mind, the following section outlines how Collingridge's work has been and could be employed in the elaboration of RRI. In particular, it considers how his ideas on the dilemma of control and corrigibility of decisions about technology might address issues of accountability referred to above.

3. Collingridge's contribution to RRI

When RRI literatures draw on Collingridge, the most explicitly invoked theme is the 'dilemma of control' (or 'Collingridge dilemma') (Asante et al., 2014: 13–14; Fonseca and Pereira, 2013: 51; 2014: 17; Kiran, 2012: 220; Owen, 2014; Owen et al., 2013b: 34; Stilgoe et al., 2013: 1569; Lee and Petts, 2013: 145; Rose, 2012: 9; Van den Hoven, 2013; Van den Hoven, 2013: 80). The dilemma runs thus: 'attempting to control a technology is difficult…because during its early stages, when it can be controlled, not enough can be known about its harmful social consequences to warrant controlling its development; but by the time these consequences are apparent, control has become costly and slow' (Collingridge, 1980: 19). This leads to the importance of a second theme in more detailed accounts, concerning the corrigibility of innovation trajectories (see: Blok, 2014; Owen et al., 2013b; Lee and Petts, 2013). And alongside these explicit references to his work, Collingridge's thinking also has implicit influence on RRI – for instance in the recognition of the significance of corrigibility in the form of 'responsiveness'. Stilgoe et al. (2013: 1572) describe this as the 'capacity to change shape or direction in response to stakeholder and public values and changing circumstances'.

In general Collingridge's work is used as fundamental grounding for framing: (a) the problem agenda of RRI; and (b) particular strategies for steering technology-society more effectively. The dilemma of control has been invoked in a general sense to underpin discussions of how to govern uncertain or potentially undesirable innovations in contexts where knowledge is unavailable or contested (Asante et al., 2014). In this sense, RRI emerges as a direct response to the Collingridge dilemma, in which respect a number of other approaches for governing emerging technologies have been found wanting (Rose, 2012). In essence, the argument is that the Collingridge dilemma can be overcome when responsibility is embedded in emerging technologies in the form of enhanced reflexivity among researchers alongside wider provision for 'upstream' engagement (Fonseca and Pereira, 2013, 2014).

Reference to Collingridge is especially prominent where RRI seeks to address the 'first horn' of the Collingridge dilemma (concerning the dearth of necessary early information about technological implications). Eschewing simplistic instrumental approaches, attention focuses on exposing developments to a range of 'mid-stream, multidisciplinary perspectives' of kinds that were undeveloped when Collingridge first developed his 'dilemma' (Kiran, 2012). Other particular dimensions of RRI that involve elaboration of core ideas from Collingridge, include recognition for the importance of continuing dialogue processes as a means to enhance the corrigibility of decisions. In this regard, 'dialogical responsiveness' involves destruction and reconstituting of the identities of those participating in deliberations about scientific research and innovation (Blok, 2014).

Alongside these well-recognised themes, other aspects of Collingridge's work are under-acknowledged in RRI, and offer significant potential for useful influence. For instance, Collingridge's approach to seeking social control over technology is often interpreted simply to involve monitoring and continual search for error – with remediation best facilitated by ensuring that those pathways that are pursued are as corrigible as possible. However, this is only part of the story. Other themes in Collingridge's work that are undervalued in the RRI literature involve other strategies than corrigibility and address the afore-mentioned dimensions of RRI. For instance, in relation to anticipatory technological decision-making Collingridge describes a series of 'equivalent' ways of 'overcoming obstacles to the control of technology', including: keeping options open; increasing the insensitivity of performance of technology to error; escaping the 'hedging circle'; enhancing controllability; managing entrenchment; reducing dogmatism of experts; and minimising the diseconomies of scale.

Collingridge (1980) argues that keeping future options open facilitates the *social control* of technology by enhancing the flexibility of decisions. Having a range of technical options available avoids reliance on any one technology. For Collingridge, the choice of which nascent innovation pathways to pursue (or not) is a matter of societal and technological choice, implicated with competing visions of the purposes, benefits and limitations of technology and more or less effective processes for decision-making. In relation to the first horn of his dilemma, the knowledge required to avoid 'big mistakes' may be knowledge pertaining to a class of similar decisions about technology,

albeit that operational knowledge about the technology in question may not yet be available. In Nordmann's (2014) terms, RRI should emphasise the search for alternative scenarios and technological options, rather than comprehensive 'knowledge of the future'. This increases freedom of manoeuvre, opening up a wider range of possible future actions. In this way, a system composed of many small units of operation or production presents far wider options than dependence on a few very large units which rely on highly specialised and capital intensive infrastructure. The trade-off is between flexibility of decisions on one hand and the loss of the economies of scale from the more inflexible technology on the other. Public accountability may also be more practicable in cases when technology is insensitive to error, that is when the time taken to discover and remedy a 'mistaken' decision is short and the costs of the mistake compared with those of following the 'correct' option "make little difference to the final pay-off" (Collingridge, 1980: 40).

A particular problem for anticipatory decision–making is manifest in what Collingridge (1980: chapter 5) calls the 'hedging circle'. This refers to a process in which liberal 'just-in-case' assumptions (e.g. in the energy sector about future energy demand growth and GDP) interlock with existing 'low variety' supply technology to create a vicious circle in which supply is increased in the expectation of growing demand, and demand growth materialises as consumers adjust to supply increases. Decision-makers understand the cost of error in terms of failure to supply (energy) according to expected demand as generously forecasted. In this way, it appears as if expansion of the prevailing low-variety system (in this case based on centralised generation of non-renewable electricity), has a lower cost of error, than investment in a more decentralised supply system or energy efficiency. Subsequent thinking and experience has illuminated this fallacy (GEA, 2012).

Here, Collingridge distinguishes between 'controllability' and 'corrigibility' in that the former relates to the efficacy, cost and timeliness with which wider social agency can be asserted over the orientation of a technological trajectory, and not just to the ease with which specific errors may be corrected. Decisions that are easily controlled will have what Collingridge calls low 'control cost', which he defines as the costs of applying a remedy to a mistaken decision (Collingridge, 1980: 33). Where such costs are unknown in advance, then options with low fixed costs are preferable. If these decisions are mistaken then the losses of sunk costs associated with highly capital intensive options may be kept low.

Assuring responsiveness to changed circumstances or values is difficult in situations where technologies are 'entrenched' (Collingridge, 1979; 1980: chapter 3; 1983; c.f. Arthur, 1989 for the related notion of technology 'lock in' referred to in RRI literature). Entrenchment may be thought of in terms of the 'second horn' of the Collingridge dilemma, extending beyond problems of information regarding the future performance of technology to challenges of insufficient agency – a developed technology is more clear in its implications but more entrenched in the face of efforts to reshape it (Owen et al., 2012, 2013b), though continued monitoring may be possible. In addition to entrenchment, however, Collingridge identifies further impediments to responsiveness, which tend not to be considered in RRI literatures.

First, responsiveness may become difficult due to competition (which may for Collingridge be between firms, nation-states or regional blocs). This is due to the ways in which competition can serve to restrict the number of options while dulling the reflexivity, or critical ability, of influential stakeholders. This occurs, for example, when the technology in question becomes so intrinsically embedded in imaginations of the future that competition is based entirely around optimising this trajectory, rather than exploring alternatives.

Secondly, responsiveness is impaired by dogmatism, which serves to reduce societal reflexivity and inclusive deliberation. This involves the many ways in which incumbent interests that are associated with a particular entrenched technology can avoid criticism of their favoured commitments. Dogmatism flourishes in the absence of transparent processes of scrutiny or where such monitoring is conducted too lightly or by groups unable to exercise independent views of the technology concerned (Collingridge, 1980: chapter 9). Scientific experts and the use of scientific expertise are central to dogmatism. Collingridge argues that scientific experts distort the 'proper task' of technology which is supposed to be to meet societal needs and fail to address more human aspects of technology. Instead, they and their work tend to be unduly optimistic, overly technical and serve the narrow needs of large organisations and government, by and for which their expertise is commissioned (Collingridge, 1992: 180–182).

Thirdly, responsiveness is inhibited by diseconomies of scale. Collingridge (1980, 1992) identified a number of indicators of scale, and applied these to a variety of case studies, including system-built high-rise buildings, nuclear power, large irrigation schemes in developing countries, and the US space shuttle. The indicators thereby implicated include long lead time, large unit size, capital intensity and dependence on specialised infrastructure. Collingridge shows how these characteristics of large inflexible technologies may bring economies of scale once they are deployed but impose severe diseconomies of scale in terms of response time and control costs on society if decisions are subsequently found to be mistaken.

In all these ways, more fulsome institutionalisation of Collingridge's ideas to RRI might additionally emphasise and combine attention to: (a) flexible, incremental decisions, which are more likely to be taken by (b) incremental decision-making processes accountable to stakeholders who are usually left out of such mechanisms (c.f. Lindblom, 1959, 1990); and (c) increased criticism (and relaxed reliance on) the worldviews of the kinds of professional experts that tend to be most implicated in RRI – again a key focus of 'accountability' as distinct from 'responsibility' (c.f. Lindblom, 1990). But Collingridge's work is of course also susceptible to critique – not least in relation to other emerging insights in RRI. In particular, this discussion relates to a series of analyses that have emerged since Collingridge ceased active research, including research on: constructive technology assessment and discourse; the dynamics of expectations; and socio-technical scenarios and imaginaries. It is to this that attention will now turn.

4. A critique of Collingridge's work in relation to RRI

Over the last three decades, Collingridge's own approach has been subject to much useful criticism. Sometimes – despite its age – prompting fresh research questions even now, this also offers significant insights for developing RRI agendas. For instance, some critics of Collingridge have read his work as an "externalist" view, unduly separating of relations between 'technology' and 'society' (e.g. Johnston, 1984; Kiran 2012). Collingridge's (1985: 380) (somewhat defensive) response is that (to him) the 'story of technology… obviously involves its social and institutional aspects as well'. But it is important here to understand connections between the content of inflexible technologies and the closed processes and interactions (i.e. between policy-makers and large firms) through which they are promoted. In Kiran's (2012) view RRI approaches should not seek to 'appropriate' this position. In other words, governing technologies responsibly requires a more nuanced approach to the relatively simple distinction between upstream or downstream 'design strategies' implied in the Collingridge dilemma. Instead contemporary 'mid-stream' multidisciplinary perspectives recognise the mutual interdependence of the 'technical' and the 'social', and offer the potential to allow participants to work out questions of function and meaning in the fray of sociotechnical development.

A second kind of criticism emanates from contributions to RRI literature which emphasise (and problematise) the responsiveness of actors to each other (e.g. Blok, 2014; De Bakker et al., 2014; Lövbrand et al., 2011; Sykes and Macnaghten, 2013; Van Oudheusden, 2014). These juxtapose ideals of deliberative democracy working towards 'the common good' with corresponding critiques from the field of science

and technology studies (STS) which stress the plurality of knowledges and assumptions which can 'inform collective action' (Lövbrand et al., 2011). Here however, there is actually a strong consistency with Collingridge's position on the role of scientific expertise in technological controversies, which emphasises how scientific experts called upon to present knowledge for opposing sides in technical controversies, will obviously disagree over interpretations and implications (Collingridge, 1980; Collingridge and Reeve, 1986). Yet Collingridge does envisage that mutual rapprochement of experts may be secured if appeals are successful to 'background values' (Collingridge, 1980: chapter 12). For its part, STS recognises that these values themselves are typically contestable (Felt et al., 2008, 2013).

Here again, however, Collingridge (1992: 186) does nonetheless still converge with pluralist STS, in that his notion of associated trial and error learning 'requires mutual co-ordination between disparate interest groups sharing power, each having a veto'. It is from this (in STS terms) more reflexive process, that more flexible sociotechnical configurations may be selected. These may be far from perfect, but they may be expected to better serve the people affected by them – as well as proving better adapted to realities of human ignorance about the future. In any case, this kind of appeal to shared or 'background values' can also to be found in the STS-informed RRI literature – for instance in von Schomberg's description of the EU approach to RRI (Von Schomberg, 2013). The extent and depth in which underlying social values must be differentiated, or may be assumed to be shared, could usefully be problematized and subject to further scrutiny in future RRI research.

Two further key points here arise in the extent to which processes of technology development embody or enable democratic influence of those who stand to be affected. Relating to the accountability issues mentioned above, these concern the requisite diversity of stakeholders involved in 'alignment' processes, as well as the capacities of these processes for reflection and responsiveness to plural values and interests. Relevant here is Collingridge's concern that debates about technologies be seen as continuing processes, rather than one-off exercises. Accordingly, his work highlights the need to conceive of accountabilities not consequentially in relation to postulated future outcomes, but rather pragmatically in terms of more immediately apprehensible contemporary qualities of innovation processes themselves: like inclusion, openness, incrementalism, flexibility and reversibility.

Also, in Collingridge's terms, the focus should not only be on emerging technologies but also on monitoring those which have become entrenched. Recognising that judgements regarding the flexibility or robustness of technologies are always provisional, Collingridge argues that scrutiny should persist for as long as it is possible for debate to continue. However, he does not say how the category of 'affected' parties is to be determined in practice *ex ante*. Nor does he offer guidance about how democratic governance should be practised so as to admit such potential participants, let alone how to ensure in the presence of encompassing power gradients, that they exert influence over decisions. Instead, Collingridge simply analyses cases where such actors are 'missing' (Winner, 1993) and cautions about potentially infinite regress in criticism in (or of) science. Subsequent contributions, including contemporary approaches to RRI, still struggle to address these issues effectively. Here in particular, there seems much further scope for further critical research and action.

One area where exactly these issues where this challenge was especially addressed is Constructive Technology Assessment (CTA). Also in danger of neglect in subsequent RRI research – and especially in parallel fields like transition management (Meadowcroft, 2009; Shove and Walker, 2007; Smith and Stirling, 2010) – CTA focuses directly on the open and inherently political nature of alignment processes. Focussing on the dynamics of negotiating between promotion and control, a range of strategies emerge on the part of regulators, marketing or environmental departments in firms and their various protagonists in wider political debate (Rip and te Kulve, 2008; Schot, 1992; Schot and Rip, 1997; c.f. Genus, 2006; Genus and Coles, 2005). In Jasanoff's work on sociotechnical imaginaries, as well, visions and understandings on the part of non-specialists are afforded equal attention and significance to expert perspectives, with systematic contrasts observed between circumstances in contrasting national settings (Jasanoff and Kim, 2009). Here, RRI shares in common with many other subsequent academic contributions to technology governance – including transition management and CTA – a tendency to be most preoccupied with interactions between social scientists, scientists, research funders, policy makers and entrepreneurial or innovating firms. They have been less directly concerned with relations (like accountability) extending to NGOs and citizens more widely.

For example Owen et al. (2013b: 46) refer to the important role of 'universities, institutes, and research funders' who enjoy 'co-responsibility' for defining responsible innovation and to 'institutionally embed' the RRI framework they advocate. In CTA reference is made to 'the relevant institutions and networks that are directly involved [in technology development], but also to "third parties" who can provide or withhold credibility and legitimation (for example insurance companies, NGOs and critical or activist groups' (Rip and te Kulve, 2008). Thus citizen perspectives and democratic control of technology can appear secondary (even 'tertiary') considerations. This is despite declared aspirations and recognition that the institutionalising of new technologies is an inherently social process in the widest sense, implicating broad societal and environmental concerns and unintended effects. Again, Jasanoff's (2003) discussion of 'technologies of humility' offers some especially salient principles for 'upstream' citizen engagement, of kinds that remain to be institutionalised in any effective way in RRI. Indeed, tendencies discussed here in some RRI practices and structures towards relatively instrumental orientation, narrow scope and circumscribed participation, mean these may sometimes more accurately be referred to in Jasanoff's terms as 'technologies of hubris'.

For his part, however, Collingridge may in these terms also be criticised for lack of attention to wider patterns of institutional organisation and practice around technological decision-making. Albeit still under-developed, this is an important area of emerging attention in contemporary work in RRI (Stilgoe et al., 2013). Seen, after Giddens, as recursive rules and resources through which social practices are made and reproduced (Giddens, 1984: 24), institutions are 'by definition' the most enduring features of social life in modern societies. They are thus crucial to the dynamics of emerging sociotechnical configurations.

This said, though he does not explicitly use the term, it is clear that Collingridge is strongly aware of institutional dynamics. For example, Collingridge (1982: preface) refers to how his approach can enable more effective criticism of those features of existing social institutions that give rise to inflexible technologies. It could be argued, then, that Collingridge's views have implications for the formal or regulative institutional arrangements which could (or should, in his view) be adopted to promote the responsible governance of innovation. These involve the state avoiding gold plating favoured technologies through financial subsidies or taxation allowances. They also highlight the limitations of a 'picking winners' approach to technology and innovation policy, with incrementalism more closely fitting a policy of generating and preserving alternatives. Thus, to adapt Collingridge's approach, analysis of relevant policies should seek to investigate not just the governmental institutional arrangements, but also wider societal and cultural contexts, which might bear on the shaping of more open forms of research or more flexible configurations for innovation capable of being implemented by smaller and more diverse kinds of organisations (Voss and Freeman, 2015).

In relation to institutionalised practices for monitoring decisions about technology, Collingridge notes that conventional 'administrative rules' serve to impede discovery of expert bias due to a series of factors. First, there is unequal funding to expertise aligned with different interests, making expertise unfair both in its availability and orientation. Second, there is the tendency to domination of research fields and agendas by a few experts – the so-called 'Kehoe' problem. This refers to

the leading scientific authority on environmental lead in the 1930s, whose findings came to determine an erroneous threshold limit for safe exposure to lead. Third, there are bureaucratic rules which impose secrecy or protect key "facts" from criticism – as repeatedly documented in public inquiries (Millstone and van Zwanenberg, 2005; Wynne, 2010).

These existing institutional arrangements may be contrasted with another model of the role of scientific experts, 'which can accommodate the fact that experts can be expected to disagree' (1980: 191; see also Collingridge and Reeve, 1986). This is also one in which it is recognised that 'there is nothing wrong' in scientists being 'advocates' for particular perspectives or policies. As explored further by Pielke (2007), however, this holds only for as long as scientific practices and norms are adhered to in a broader sense – as well as wider values of reasoned deliberation (Dryzek, 2006).

Collingridge's work is informed by the contribution of Charles Lindblom on the prevalence and reduction of 'professional impairment' (Lindblom, 1990). A contemporary development of such thinking directs attention to 'normative' institutional arrangements which govern the training of scientists and social scientists – for example, in relation to public engagement and ethics, and in terms of their capacity for self-criticism, reflection and proactive accountability. But what does not feature so directly in Collingridge's thinking, are institutionalised rules implicated with understandings, beliefs and expectations concerning participation of publics *per se*. This includes such rules concerning the responsibilities of (and requirements for) citizens to play active roles in technology development, to help guard against unfair or unwise innovations. But again, this gap in Collingridge's thinking also seems to persist equally in much contemporary academic thought.

The following section reflects on how RRI might usefully engage more deeply with Collingridge's ideas and the implications of doing so for developing both understanding and practice of RRI. This discussion also benefits from appreciation for how Collingridge's thinking is informed by Charles Lindblom's work on political incrementalism and the associated 'Popperian' fallibilist approach to decision-making under conditions of uncertainty or ignorance. The discussion offers for RRI a sceptical view of anticipatory decision-making. It also offers insights into the constituting of reflexivity and responsiveness and to some extent inclusive deliberation of kinds that are emergent in Collingridge's work but not referenced within RRI literatures, even amongst those contributions which cite him.

5. Building RRI: revisiting Collingridge?

Collingridge's approach emphasises active processes of learning from a particular class of past decisions in order to inform future decision-making about technology development, scientific research and innovation. This contrasts with the more hubristic techno-scientific approaches to 'futures' which form one branch of the technology assessment literature (Selin, 2014). Collingridge is concerned pragmatically with the qualities of emerging innovations, rather than consequentially with their outcomes. This should not be confused with questions about the availability or not of knowledge regarding a focal, emerging technology of a kind that might otherwise be the central framing adopted in current approaches to responsible innovation. Whereas his contributions are often thought of as being preoccupied with lock-in and the closing down of governance processes, it is the incrementalist and Popperian underpinning of Collingridge's thinking that most emphasise the need for open processes (c.f. Stirling, 2008). These themes are developed in the paragraphs below.

5.1. Incrementalism

Some contributions to RRI literature which refer to Collingridge's work, neglect to recognise its immersion in a wider pool of thinking around incremental policy, strategy and decision-making. The writings of Charles Lindblom, for instance, are a particularly strong influence on Collingridge's thinking about what he called the 'management of scale' and flexible decision-making about technology.

For Lindblom (1959) incrementalism involves two aspects: a) an emphasis on relatively small changes from some pre-existing state of affairs (which may nonetheless be large in their cumulative effects); and b) the participation of likely-affected 'partisans' whose continual proactive 'mutual adjustments' help to prevent the more egregious kinds of mistake associated with centralised planning. Later, Lindblom (1990) emphasised the role and need for wider citizen engagements, allowing the probing of past and prospective decisions. So too, in Collingridge's (1980) view, the post-hoc monitoring of past technological decisions should be encouraged in the form of public debate, with recognition that this may periodically entail the reversal of deep commitments. In ways that chime with Van de Poel's (2000) emphasis on the need to include 'outsiders' and Winner's (1993) concern with 'missing actors', both Lindblom and Collingridge see this in terms of potentially empowering and involving those who are often left out of technological decision-making. To the extent that some of the more instrumental applications of RRI practice are seen (as discussed above) in conservative quarters as substitutes for waves of policy enthusiasm for 'public engagement', this work of Collingridge's serves as a reminder that wider inclusion and empowerment are central to – and constitutive of – responsibility.

The resulting issues extend very widely. Collingridge identified as one of the 'most pressing problems of our time', the question of: 'can we control our [sic] technology'? (Collingridge, 1980: preface). In summarising his overall response, Collingridge notes that what is required extends well beyond specific bolt-on methods, tools and practices, to encompass the entire 'normal machinery of politics'(1983: preface). For him, it is this 'normal decision-making' that needs to be much more strongly conditioned by incrementalism. Thus he states that 'we' ought to avoid those decisions that can't be taken incrementally – a clear, normative commitment with extensive repercussions not just for policy but for the politics of technology more widely. It is in this way that Collingridge seeks to escape the excessively deterministic connotations of "control" (Collingridge, 1983). And it is in these senses that he offers what are arguably some of his most salient messages for RRI.

5.2. Ignorance and fallibility

Reaching back especially to the work of Popper (1979), Collingridge's incremental approach to the social control of technology is based on the philosophy of fallibilism. Here, Collingridge considers error to be an unavoidable part of being human. He contrasts his emphasis on human fallibility, with then (and still) prevailing conditions under which incumbent governance cultures aspire – and claim – synoptically rational decision making. With Bayesian probabilistic methods for risk-based decision-making even more prominent today than they were in Collingridge's time, such attitudes are also exemplified in the current emphasis on 'sound science' and 'evidence based decision-making' – as if these permitted single self-evidently definitive prescriptions (Stirling, 2010b).

This continuing 'dominant tradition' rests on what Collingridge refers to as the justificationist model, in which only those decisions which can be fully justified are seen as rational. As a result, strong pressures are formed to suppress recognition for uncertainties, ambiguities and irreducible ignorance. In "Bayesian" and kindred approaches critiqued by Collingridge, these kinds of indeterminacy and intractability are instead treated with aggregated probabilities. It is central to Collingridge's insights that these kinds of 'reductive aggregative' tools are deeply misleading. Yet many parts of the responsible innovation literature that otherwise draw on his work, are not so clear concerning these kinds of problems with what continue to be dominant methods (Martin, 2013; Walport, 2014).

There is a key parallel here between the Popperian process in which

scientific progress rests on effective falsification of conjectures, and ways in which wider capacities for reasoned societal interrogation can expose flaws in technology development and operation. It is arguably here that the importance of accountability becomes most significant – as an element of responsibility relating not to anticipated consequences, but to the appropriate prioritisation of 'Collingridge qualities' in emerging technologies themselves. It is in qualities of inclusion, openness, incrementalism, flexibility and reversibility that the parallels are strongest between Popper and Collingridge. Conditions favouring robust knowledge and robust technology are not so different. Again, the implications for RRI are quite direct.

Taken together, there arise in the foregoing discussions a number of searching questions with quite practical implications: 1). What are the different meanings and forms of RRI and how do these relate to concrete qualities identified by Collingridge? 2). To what extent is RRI becoming institutionally embedded as a means to assert, or alternatively to challenge, processes of justification? 3). What visions and styles of scientific expertise and discourse are most promoted by RRI – openly contending or apparently harmonious? 4) What are the relations between public participation and stakeholder engagement with RRI – as a source of challenge and substantive orientation, or a resource for securing legitimacy? 5) What are the implications of practices and structures of RRI for wider qualities of democratic accountability – do these tend to be emphasised and reinforced, or suppressed and substituted?

6. Conclusion

This paper seeks to make a three-fold contribution. First it offers a relatively full and systematic critical analysis of the work of David Collingridge – which is acknowledged to have exercised an important influence on the currently burgeoning field of responsible (research and) innovation (RRI), but of which some of the wider and deeper implications have arguably been neglected. Second, the paper has explored some specific implications of this relative neglect, for some central themes in RRI – for instance arising in the importance of pragmatic incrementalism rather than consequentialist justification. The third contribution has been to analyse some significant limitations in Collingridge's approach as illuminated in the light of subsequent developments. This also raises practical issues for RRI, particularly with regard to the strengthening of 'Collingridge qualities' – including inclusion, openness, diversity, incrementalism, flexibility and reversibility.

Of course, some problems are presented for aspects of Collingridge's analysis by the advent of contemporary 'grand challenges' for innovation policy, such as climate change and social justice. In particular, serious questions appear to be raised for his vision of incrementalism, by imperatives to achieve urgent and broad-scale transformations. These may partly be alleviated by pointing to the potentially radical cumulative effects of incrementalism in achieving occasionally transformative emergent cultural "murmurations" (Stirling, 2016a). But the tension still remains.

Likewise, there are also criticisms that Collingridge presents a decisionistic and 'machinery view' of technology in society, more consistent with 'risk regulation' than the 'innovation governance' of RRI approaches. This links to the paradoxically deterministic connotations of the 'control' metaphor that Collingridge uses to communicate his analysis. Again, Collingridge's choice of metaphor is of its time. His prescriptions of inclusion, openness, diversity, incrementalism, flexibility and reversibility might all now be better expressed in terms of qualities other than 'control' – including care, solidarity, mutualism, non-consequentialist notions of accountability and responsibility itself.

Nonetheless, whilst the questions raised for RRI still stand, these limitations in Collingridge's approach do detract from its contemporary value. And, though he highlights the problem well, Collingridge's work also fails effectively to resolve how to achieve the necessary kinds of wider and deeper democratic deliberation, including by vulnerable marginalised communities who are repeatedly deeply affected by incumbent patterns of decision-making, but who remain perennially 'missing actors' (Winner, 1993). It is in this area that work in subsequent decades around modalities for participatory deliberation, marginalised interests and responsiveness to 'uninvited' collective action arguably have most to offer (Wynne, 2007).

This reinforces a point that is already quite well appreciated in particular areas of RRI but which may fruitfully be restated in terms of three implications of the paper for future RRI research and practice and the directions that these should take. First there is a need to develop and invigorate more concrete and assertive frameworks for enabling practice of critical citizen engagement and participatory deliberation (see Macnaghten and Chilvers, 2014 for a recent example). Attention to the full scope of Collingridge's analysis would fortify such moves. But it would also have the effect of raising particular questions and pointers.

Secondly, and in relation to the previous point, RRI needs to have due regard to Collingridge's emphasis on fallibility and the ever-present intractabilities of ignorance. In terms of the direction of ensuing RRI research and practice this highlights the value of processes and discourses that illuminate, rather than suppress, contention among specialists and wider societal interests. Particular challenges are raised by this for the search for "right impacts" in RRI (Von Schomberg, 2013). Here, an implication of the paper is that responsibility lies not in engineering consensus, but in exploring dissensus (Genus, 2006). And it is in this regard that the 'Collingridge qualities' – around inclusion, openness, diversity, incrementalism, flexibility and reversibility – offer concrete constituting (albeit sometimes contending) axes meriting further deliberation in RRI.

Thirdly, to the extent that RRI approaches can fully embrace Collingridge's contributions, they will need to grapple not only with contending qualities and principles for rationalistic decision making but also with the fundamental realities (foundational for Collingridge) that the governance of research and innovation are fundamentally about 'muddling through' in the presence of steep power gradients and strongly asserted interests. In this sense, it is a core feature of responsibility that it is often better engaged with as a struggle against incumbent power, than as an instrumental facilitation (Stirling, 2016a). Thus understood, one of the most important properties of responsibility lies in the reinforcing – rather than the attenuation – of accountabilities. In the end, the kinds of humility and pluralism urged by Collingridge in the face of ignorance and contending interests, underscore the point that the most responsible way to govern innovation is by democracy itself. The institutions and practices of RRI are arguably only progressive insofar as they helps to strengthen, rather than weaken, this general aim.

References

Arthur, W.B., 1989. Competing technologies, increasing returns, and lock-in by historical events. The Econ. J. 99 (394), 116–131.
Asante, K., Owen, R., Williamson, G., 2014. Governance of new product development and perceptions of responsible innovation in the financial sector: insights from an ethnographic case study. J. Responsible Innovation 1 (1), 9–30.
Beck, U., 1992. Risk Society: Towards a New Modernity. Sage, London.
Beck, U., 1999. World Risk Society. Polity, London.
Bernal, J.D., 1939. The Social Function of Science. Routledge, London.
Bernal, J.D., 1954. Science in History. Faber and Faber, London (4 volumes).
Blok, V., Lemmens, P., 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: In: Koops, B.-J., van den Hoven, H., Swierstra, T., Oosterlaken, I. (Eds.), Responsible Innovation, vol. 2. Springer, Dordrecht, pp. 19–35 Concepts, Approaches, and Applications.
Blok, V., 2014. Look who's talking: responsible innovation, the paradox of dialogue and the voice of the other in communication and negotiation process. J. Responsible Innovation 1 (2), 171–190.
Chilvers, J., 2008. Delivering competence: theoretical and practitioner perspectives and effective. Sci. Technol. Human Values 33, 155–185.
Collingridge, D., Reeve, C., 1986. Science Speaks to Power: the Role of Experts in Policymaking. Frances Pinter, London.
Collingridge, D., 1979. The entrenchment of technology. Sci. Public Policy 332–338.

Collingridge, D., 1980. The Social Control of Technology. Pinter, London.
Collingridge, D., 1982. Critical Decision Making. Frances Pinter, London.
Collingridge, D., 1983. Technology in the Policy Process: the Control of Nuclear Power. Frances Pinter, London.
Collingridge, D., 1985. Controlling technology (response to Johnston). Soc. Stud. Sci. 15 (2), 373–380.
Collingridge, D., 1992. The Management of Scale. Routledge, London.
Conrad, E., Cassar, L.F., Christie, M., Fazey, I., 2011. Hearing but not listening? A participatory assessment of public participation in planning. Environ. Plann. C: Gov. Policy 29, 761–782.
De Bakker, E., de Lauwere, C., Hoes, A.-C., Beekman, V., 2014. Responsible research and innovation in miniature: information asymmetries hindering a more inclusive 'nanofood' development. Sci. Public Policy 41, 294–305.
Donnelley, S., 1989. Hans Jonas, the philosophy of nature, and the ethics of responsibility. Social Res. 56 (3), 1–13.
Dryzek, J.S., 2006. Deliberative Global Politics: Discourse and Democracy in a Divided World. Polity Press, Cambridge.
Felt, U., Wynne, B., Callon, M., Gonçalves, M.E., Jasanoff, S., Jepsen, M., Tallacchini, M., 2008. 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. European Commission, Brussels.
Felt, U., Barben, D., Irwin, A., Pierre-Benoit-Joly, Rip, A., Stirling, A., Stockelova, T., 2013. Science in Society: Caring for Our Futures in Turbulent Times. Strasbourg.
Fonseca, P., Pereira, T.S., 2013. Emerging responsibilities: brazilian nanoscientists' conceptions of responsible governance and social technology practices. In: Konrad, C., Coenen, A., Dijkstra, C., Milburn, H. (Eds.), Shaping Emerging Technologies: Governance, Innovation, Discourse. IOS Press/AKA, Berlin, pp. 49–65.
Fonseca, P.F.C., Pereira, T.S., 2014. The governance of nanotechnology in the Brazilian context: entangling approaches. Technol. Soc. 37, 16–27.
GEA, 2012. In: Davis, G., Goldemberg, J. (Eds.), Global Energy Assessment Toward a Sustainable Future. Cambridge University Press, Cambridge, UK.
Genus, A., Coles, A.-M., 2005. On constructive technology assessment and limitations on public participation in technology assessment. Technol. Anal. Strategic Manage. 17 (4), 433–443.
Genus, A., 2006. Rethinking constructive technology assessment as democratic, reflective, discourse. Technol. Forecasting Social Change 73 (1), 13–26.
Giddens, A., 1984. The Constitution of Society: Outline of the Theory of Structuration Polity. Cambridge.
Grinbaum, A., Groves, C., 2013. What is 'Responsible' about responsible innovation? understanding the ethical issues. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, Chichester, pp. 119–142.
Guston, D.H., Sarewitz, D., 2002. Real-time technology assessment. Technol. Soc. 24, 93–109.
Guston, D.H., Fisher, E., Grunwald, A., Owen, R., Swierstra, T., van der Burg, S., 2014. Responsible innovation: motivations for a new journal. J. Responsible Innovation 1, 1–8.
Guston, D.H., 2006. Responsible knowledge-based innovation. Society 19–21.
Haldane, J.B.S., 1939. Science and Everyday Life. Pelican, London (reprinted 1941) (1939).
Harremoës, P., Gee, D., MacGarvin, M., Stirling, A., Keys, J., Vaz, S.G., Wynne, B. (Eds.), 2000. Late Lessons from Early Warnings: the Precautionary Principle 1896–2000. European Environment Agency, Copenhagen.
House of Lords, 2000. Select Committee on Science and Technology. Science and Society, London.
Irwin, A., 1995. Citizen Science: A Study of People, Expertise and Sustainable Development. Routledge, London.
Irwin, A., 2001. Constructing the scientific citizen: science and democracy in the biosciences. Public Underst. Sci. 10, 1–18.
Jasanoff, S., Kim, S.-H., 2009. Containing the atom: sociotechnical imaginaries and nuclear power in the United States and South Korea. Minerva 47, 119–146.
Jasanoff, S., 2003. Technologies of humility: citizen participation in governing science. Minerva 41, 223–244.
Johnston, R., 1984. Controlling technology: an issue for the social studies of science. Soc. Stud. Sci. 14, 97–113.
Jonas, H., 1984. The Imperative of Responsibility: in Search of an Ethics for the Technological Age. Chicago University Press, Chicago.
Kiran, A., 2012. Does responsible innovation presuppose design instrumentalism?: Examining the case of telecare at home in the Netherlands. Technol. Society 34, 216–226.
Lövbrand, E., Pielke, R., Beck, S., 2011. A democracy paradox in studies of science and technology. Sci. Technol. Human Values 36, 474–496.
Lee, R.G., Petts, J., 2013. Adaptive governance for responsible innovation. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, London, pp. 143–164.
Lindblom, C.E., 1959. The science of muddling through. Public Adm. Rev. 19, 79–88.
Lindblom, C.E., 1990. Inquiry and Change: The Troubled Attempt to Understand and Shape Society. Yale University Press, New Haven.
Macnaghten, P., Chilvers, J., 2014. The future of science governance: publics, policies, practices. Environ. Plann. C: Gov. Policy 32, 530–548.
Martin, B., 2013. Innovation Studies: an emerging agenda. In: Fagerberg, J., Martin, B., Andersen, E.S. (Eds.), Innovation Studies: Evolution and Future Challenges. Oxford University Press, Oxford.
Meadowcroft, J., 2009. What about the politics? Sustainable development, transition management, and long term energy transitions. Policy Sci. 42 (4), 323–340.
Millstone, E., van Zwanenberg, P., 2005. BSE: Risk, Science, and Governance. Oxford University Press, Oxford NNI (The National Nanotechnology Initiative), Strategic Plan 2011. http://www.nano.gov/sites/default/files/pub_resource/2011_strategic_plan.pdf.
NNI (The National Nanotechnology Initiative), 2011. Strategic Plan. http://www.nano.gov/sites/default/files/pub_resource/2011_strategic_plan.pdf.
Nordmann, A., 2014. Responsible innovation, the art and craft of anticipation. J. Responsible Innovation 1 (1), 87–98.
Owen, R., Macnaghten, P., Stilgoe, J., 2012. Responsible research and innovation: from science in society to science for society, with society. Sci. Public Policy 39, 751–760.
Owen, R., Bessant, J., Heintz, M. (Eds.), 2013. Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, London.
Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., Guston, D., 2013b. A framework for responsible innovation. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, London, pp. 27–50.
Owen, R., 2014. Responsible Research and Innovation: Options for Research and Innovation Policy in the EU Report for ERIAB..
Pielke Jr., R.S., 2007. The Honest Broker: Making Sense of Science in Policy and Politics. Cambridge University Press, Cambridge.
Popper, K., 1979. Objective Knowledge: an Evolutionary Approach. Oxford University Press, Oxford Revised paperback edition.
Rip, A., te Kulve, H., 2008. Constructive technology assessment and sociotechnical scenarios. In: In: Fisher, E., Selin, C., Wetmore, J.M. (Eds.), The Yearbook of Nanotechnology in Society, vol. I. Springer, Berlin, pp. 49–70 (Presenting Futures).
Roco, R.C., Harthorn, B., Guston, D., Shapira, P., 2011. Innovative and responsible governance of nanotechnology for societal development. J. Nanopart. Res. 13, 3557–3590.
Rose, H., Rose, S., 1969. Science and Society. Penguin, Harmondsworth.
Rose, N., 2012. Democracy in the contemporary life sciences. BioSocieties 7, 459–472.
STEPS Centre, 2010. Innovation, Sustainability Development: a New Manifesto. STEPS Centre, Brighton.
Schot, J., Rip, A., 1997. The past and future of constructive technology assessment. Technol. Forecast. Social Change 54 (2/3), 251–268.
Schot, J., 1992. Constructive technology assessment and technology dynamics: the case of clean technologies. Sci. Technol. Human Values 17, 36–56.
Selin, C., 2014. On not forgetting futures. J. Responsible Innovation 1 (1), 103–108.
Shove, E., Walker, G., 2007. CAUTION! Transitions ahead: politics, practice, and sustainable transition management. Environ. Plann. A 39 (4), 763–770.
Smith, A., Stirling, A., 2010. The politics of social-ecological resilience and sustainable socio-technical transitions. Ecol. Soc. 15 (1), 1–13.
Stahl, B.C., 2012. Responsible research and innovation in information systems. Eur. J. Inf. Syst. 21, 207–211.
Stilgoe, J., Owen, R., Macnaghten, P., 2013. Developing a framework for responsible innovation. Res. Policy 42, 1568–1580.
Stilgoe, J., Lock, S., Wilsdon, J., 2014. Why should we promote public engagement with science? Public Underst. Sci. 23, 4–15.
Stirling, A., 2008. Opening up and closing down: power, participation, and pluralism in the social appraisal of technology. Sci. Technol. Human Values 33 (2), 262–294.
Stirling, A., 2010a. Direction, Distribution, Diversity! Pluralising Progress in Innovation, Sustainability and Development, STEPS Working Paper 32. pp. 1–45 Brighton.
Stirling, A., 2010b. Keep it complex. Nature 468, 1029–1031.
Stirling, A., 2014. Towards Innovation Democracy: participation, responsibility and precaution in innovation governance. In: Walport, M. (Ed.), Annual Report of the Government Chief Scientific Adviser 2014, Innovation: Managing Risk, Not Avoiding It. Evidence and Case Studies. Government Office of Science, London, pp. 49–62. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/376505/14.
Stirling, A., 2016a. Knowing doing governing: realising heterodyne democracies. In: Voss, J.-P., Freeman, R. (Eds.), Knowing: Governance: the Epistemic Construction of Political Order. Palgrave MacMillan, London Chapter 12.
Stirling, A., 2016b. Addressing scarcities in responsible innovation. J. Responsible Innovation 3 (3), 274–281.
Sutcliffe, H., 2011. A Report on Responsible Research and Innovation. MATTER, London.
Sykes, K., Macnaghten, P., 2013. Responsible innovation – opening up dialogue and debate. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Emergence of Science and Innovation in Society. Wiley, Chichester, pp. 85–108.
Van Oudheusden, M., 2014. Where are the politics in responsible innovation? European governance, technology assessments, and beyond. J. Responsible Innovation 1 (1), 67–86.
Van de Poel, I., 2000. On the role of outsiders in technical development. Technol. Anal. Strategic Manage. 12, 383–387.
Van den Hoven, J., 2013. Value sensitive design and responsible innovation. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, London, pp. 75–83.
Von Schomberg, R., 2011. Prospects for technology assessment in a framework of responsible research and innovation. In: Dusseldorp, M., Beecroft, R. (Eds.), Technikfolgen abschätzen lehren: Bildungspotenziale transdisziplinärer Methoden. Springer, Fachmedien, Wiesbaden, pp. 39–62.
Von Schomberg, R., 2013. A vision of responsible research and innovation. In: Owen, R., Bessant, J., Heintz, M. (Eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society. Wiley, London, pp. 51–74.
Von Schomberg, R., 2015. From the responsible development of new technologies towards Responsible Innovation. In: Holbrook, J.B., Mitcham, C. (Eds.), Ethics, Science, Technology, and Engineering: An International Resource, 2nd ed. Cengage, London Available at: http://renevonschomberg.wordpress.com/from-responsible-

development-of-technologies-to-responsible-innovation/.
Knowing Governance: The Epistemic Construction of Political Order. In: Voss, J.-P., Freeman, R. (Eds.), Palgrave MacMillan, London.
Reflexive Governance for Sustainable Development. In: Voss, J.-P., Bauknecht, D., Kemp, R. (Eds.), Edward Elgar, Cheltenham.
Walport, M., 2014. Innovation: Managing Risk, Not Avoiding It − Evidence and Case Studies Annual Report of the Government Chief Scientific Adviser. Government Office of Science, London.
Wickson, F., Carew, A.L., 2014. Quality criteria and indicators for responsible research and innovation: learning from transdisciplinarity. J. Responsible Innovation 1 (3), 254–273.
Winner, L., 1993. Autonomous Technology: Technics Out of Control as a Theme in Political Thought. MIT, Cambridge, MA.
Wynne, B., 1993. Public uptake of science: a case for institutional reflexivity. Public Underst. Sci. 2, 321–337.
Wynne, B., 2002. Risk and environment as legitimatory discourses of science and technology: reflexivity inside-out? Curr. Sociol. 50, 459–477.
Wynne, B., 2007. Public participation in science and technology: performing and obscuring a political?conceptual category mistake. East Asian Sci. Technol. Soci.: Int. J. 1 (1), 99–110.
Wynne, B., 2010. Rationality and Ritual: Participation and Exclusion in Nuclear Decision-Making, 2nd ed. Earthscan, London.
Ziman, J., 1968. Science Is Public Knowledge. Cambridge University Press, Cambridge.