|  |  |  |  |
| --- | --- | --- | --- |
| **Responsible Innovation Models Report** | | | |
| Deliverable | | **D.2.4** | |
| Workpackage No. | **WP2** | Workpackage Title | **Current Theory and Practice** |
| Task No. | **2.4.** | Task Title | **Determination of responsible innovation models** |
| Start Date: |  | Revision Date: |  |
| Authors | | **Sophie Pellé, Bernard Reber** | |
| Contributors | |  | |
| Status (F: final; D: draft; RD: revised draft) | | **F** | |
| Distribution | |  | |
| Document ID / File Name | |  | |

Table of content

[Introduction 3](#_Toc392104737)

[Chapter 1: Current RRI conceptions through the frame of the Analytical Grid 7](#_Toc392104738)

[Chapter 2: Presuppositions of norms in context in current RRI theories 25](#_Toc392104739)

[Chapter 3: Governance models in RRI theories 33](#_Toc392104740)

[3.1. The four models of governance 33](#_Toc392104741)

[3.2. RRI theories between consultation and co-construction 36](#_Toc392104742)

[3.3. A typology of governance models of RRI conceptions 40](#_Toc392104743)

[3.3.1. Model 1 – responsibility grounded in social acceptability 42](#_Toc392104744)

[3.3.2. Model 2 – responsibility through responsiveness and deliberation 48](#_Toc392104745)

[Chapter 4: In search of a reflexive governance theory for RRI 57](#_Toc392104746)

[4.1. RRI models of responsibility in light of the SIM presuppositions. 57](#_Toc392104747)

[4.2. Maesschalck and Lenoble’s conception of reflexivity 58](#_Toc392104748)

[4.3. Reflexivity in RRI theories 60](#_Toc392104749)

[4.4. The missing reflexive theory of governance 63](#_Toc392104750)

[Conclusion 66](#_Toc392104751)

[References 68](#_Toc392104752)

# Introduction

In previous deliverables[[1]](#footnote-1), we established an overview of RRI conceptions pointing out to the various meanings of responsibility which inspired them and to the understanding of responsible research and innovation, i.e. the main “ingredients”, to which such conceptions adhered in order to promote RRI (del 2.2 - *Theoretical Landscape*). We initiated a reflection on the possible blind spots of current RRI conceptions based upon a specific outlook of the context, its role as well as the correlated degree of reflexivity embedded in the social determination of norms.

In a second step, the *Analytical Grid* (AG), presented in del 2.3., provided a general framework of the modes of governance on which RRI conceptions are grounded. It highlighted the central role of first order and, more prominently, second order reflexivity which designates the capacity, not only to reflect on the ethical issues of technology and innovation, but, most importantly, the will and the ability to reconsider the framing according to which an issue is made intelligible. The AG defined a preliminary list of parameters driven by our analysis of the context and the role of reflexivity. In other words, several categories were established to analyze theories (RRI conceptions) and practices (innovation and research projects). These categories were inferred from the theoretical framework drawn in del 2.2. and del 2.3, which focused on the governance tools suggested by RRI approaches in order to implement a specific conception of responsibility. These categories represent a first step in the translation of our theoretical framework into practical tools of assessment. They aim at analyzing the institutional answers RRI theories and practices build in order to implement a sound process of reflexivity that is likely to allow innovation and research for being shaped by dynamics of responsibility.

What we need now is to analyze and challenge, in a more systematic way, the various conceptions of RRI presented in the *Theoretical Landscape* using the parameters from the grid. In other words, we will scrutinize the afore-mentioned RRI conceptions in the light of the grid parameters in an attempt to distinguish typologies of responsibility. This effort will include an investigation into the degree and nature of reflexivity RRI conceptions offer. What kind of institutional instruments are thought of (and eventually implemented) to allow for a reflexive, dynamic and adaptive process of research and innovation? What role does reflexivity play within a given conception? What are the theoretical links established between knowledge and uncertainty approaches and the corresponding process of elaborating norms? What type of synthetic approach of responsibility do current RRI approaches rely on? Establishing such a cartography of the processes by which norms are constructed within RRI theories will then enable us to extract possible lines of theoretical improvement.

Compared to the theoretical landscape (del 2.2.), our analysis is here conducted in a more systematic way, applying the grid parameters to RRI conceptions. To cast a critical glance at RRI approaches, we need to define what are their blind spots and their common presuppositions. This is made possible with the help of a fine and detailed investigation of what answers RRI conceptions provide to issues that have been identified and translated into a non-exhaustive list of parameters.

Two remarks must be noted here. First, the empirical work conducted in WP3 and WP4 will help to refine this preliminary list of parameters, adding some elements, specifying others. The drawing of the final list will proceed from both a top-down and a bottom-up determination, where *a priori* deduction and empirically grounded induction together will contribute to the building of a relevant analysis grid. Second, most of the RRI approaches we scrutinize here were already presented in deliverable 2.2. They include the approaches of von Schomberg (2011, a, b; 2013), Grunwald (2011, 2012), an EC Report (2013), Owen *et al*. (2012, 2013a), Sutcliff (2011), Groves and Grinbaum (2013), to which we added Guston (2013), from outside Europe. This set of approaches brings together what seemed to be prominent texts and perspectives delivered focusing on RRI.

The final step of our analysis results in a typology of RRI conceptions, i.e. in distinguishing patterns in the models of governance that are defended by the main approaches of RRI. Echoing the typology of governance models inspired by the work of Pierre-Benoît Joly and Michel Callon, which has been presented in del 2.3., our typology defines two models that articulate in a coherent way a) a definition of responsibility with b) ethical and institutional tools to foster reflexivity towards the context. In the end, this methodology allows us to provide a diagnosis of RRI “pathologies”, and, to spin the medical metaphor, to suggest possible pathways to look for remedies.

The deliverable is structured as follows. **Chapter 1** scrutinizes RRI conceptions with an eye on the grid parameters: what tools do they propose to deal with ethical thinking? What are their processes of reflexivity, the relation they establish with norms and law, their conceptions of responsibility, etc.? Then, following the theoretical steps defined in the AG, we examine each theory according to the type of presuppositions (Schematizing, Intentionalist, Mentalist, SIM hereafter) they rely on (**Chapter 2**). This will allow us to determine the model of governance under which our RRI conceptions fall (**Chapter 3)**. We will show that three of the four Joly models are exemplified in at least one of the current RRI conceptions. Then, from the systematic analysis of current RRI conceptions, we draw a typology of RRI models of governance.We show that current theories of RRI fall into two types of governance models. We present and discuss these models and the differences they hold with Joly’s four-fold distinction. We conclude that, above all, RRI approaches do not fit into Joly’s typology because they introduce a shift in the “center of gravity” of the way in which the democratic governance of science and technology is theoretically conceived: participation is now included in the core – in the DNA, could we say – of every conception of technology and innovation management, with a focus on a co-construction process that goes beyond mere consultation. The old-fashioned model of top-down approaches led by experts who have no consideration for allegedly irrational social actors (i.e. the Standard Model), seems to have come to its end – at least at the level of RRI conceptions(compared to that of practices[[2]](#footnote-2)). Finally, focusing on the three elements that ground the actual AG (the relation between norm and context, the possibility of institutional learning, and the relation to ethics) we highlight the main blind spots of current RRI conceptions (**Chapter 4**). Proposing a general interpretation of the sources of inefficiency of common approaches of responsibility in research an innovation, we suggest possible paths that might be followed to define possible answers.

# Chapter 1: Current RRI conceptions through the frame of the Analytical Grid

Following the methodology defined in the AG, the first step of our study consists in analyzing the set of RRI conceptions we identified in light of the grid parameters. For each conception, we identified 1) what is the *Weltanschauung* behind it; 2) what is the approach of ethics – what are the “ethical tools” (ELSI approaches, already existing set of norms, explorative philosophy, narratives virtue ethics, *etc.*); 3) on what kind of products it focuses (all types of products or emerging technology for instance); 4) the process of reflexivity promoted by the conception (only seeking to identify ethical conflicts and problems or consciously tackling the issue of framing and its possible reflexive construction); 5) the epistemic tool favored (principle of precaution or risk assessment); 6) the process of assessment (i.e. the way in which responsibility is assessed; 7) if approaches are sensitive to cultural differences (which would be another manner of taking the context into account); 8) the relation established with already existing norms and laws (compliance, construction of new items); 9) the role given to inclusion (participation and deliberation) and finally 10) the conception of responsibility[[3]](#footnote-3).

These parameters allow us to characterize the governance models RRI theories are promoting to generate responsible research and innovation. In pinpointing how each conception works out reflexivity but also the ethical tools they rely on, the role given to inclusion or the relation established with already existing norms, we end up with an original picture of different ways of implementing and generating RRI, also of use to draw some comparisons.

Results are summarized in tables 2. 1 to 2.5., while their interpretation belongs to chapter 3 and 4.

**Table 2.1. René von Schomberg’s and Armin Grunwald’s approaches**

|  |  |  |  |
| --- | --- | --- | --- |
| **Approach**  **Parameter** | | **von Schomberg (2011: a, b ; 2013)** | **Grunwald (2011, 2012)** |
| **Anticipation** | Vision of the world (*Weltanshauung)* | * Promotion of sustainable & ethically acceptable R&I. * Focus on social desirability. * Attempt to articulate economic efficiency with societal and ethical concern. | * Constructivist framework: RRI as shaping technology (or innovation) according to social values. * Importance given to the articulation between ethics, actions and decisions. * Ethics embedded in technological development, not distant from it. |
| Relationship with the future | Need of technology foresight to anticipate research and innovation outcomes. | Role of explorative philosophy: considering possible and rational futures. |
| **Tools** | | * Rightness of the process (transparent, interactive, responsiveness of the actors, * Grounded on EU norms (EU Treaties, Human Rights declaration, etc.) * Concrete tools are:   + Multi-stakeholder involvement.   + Codes of conducts.   + Standards and self-regulation. | * Explorative philosophy * Interdisciplinarity (natural science + social science work hand in hand). * Participation. |
| **Product** | | Applicable to every kind of product. | Focus on emerging technologies (mainly nano-biotechnology). |
| **Process (of reflexivity)** | Transparency | Required during the process. | * Scientists must be transparent as regards their knowledge. * Transparency in the relation between science, politics, and the public through deliberation. * Distribution of responsibility among actors must be transparent. |
| Reflexivity | First order reflexivity:   * Ethics helps to avoid public reluctance. * Ethics by design: It plays a role in the shaping of technology. | In the case of incrementalism: “cautious” use of the PP = Constant calling into question the norms ruling technology (confronting norms with new knowledge).  First order reflexivity (raising the important ethical issues).  Towards second order reflexivity (c.f. assessment). |
| **Epistemic Tools** | Precautionary Principle (PP) | ×   * Embedded in EU laws (e.g. REACH, Nanocodes of conduct). * Positive driver for research on risk methodologies. | ×  “Incrementalism” (step-by-step, case by case laws and norms elaboration in the case of nanotechnologies, for instance).  Avoiding a too radical application of the PP. |
| Risk assessment | ×   * Anticipating negative and positive impacts. * Identifying socially desirable products. |  |
| **Assessment** | | * Technology assessment and technology foresight. * Public participation. | Not only technology but the very process by which norms are constructed is subject to assessment: does it help to shape technology or not? If not, the process, at least, results in a reconfiguration of the dialogue between science, politics & the public (e.g. the nano-debate). |
| **Cultural differences** | | No specific mention. | No specific mention. |
| **Norm/law relation** | | * No production of new norms. * Compliance with existing EU norms, codes of conducts and laws. | Prominent role of norms, codes of conducts and then laws in implementing RRI. |
| **Participation/deliberation**  **Collective/inclusive/interactive** | | * Public debate, monitoring of public opinion. * Deliberation   Issue of the implementation of norms not raised. | Participation and deliberation needed:   * to enhance trust between science & society * because public debate can shape, if not technology, at least policy frameworks (nuclear energy in Germany, GMO’s in EU). |
| **Responsibility** | Liability/blameworthiness |  |  |
| Care |  |  |
| Responsiveness | ×  Going beyond one’s own traditional roles (i.e. firms assessing benefits, and NGOs risks). | ×  Norms, laws, principle regulating technology must adapt according to new knowledge. |
| Accountability |  | Close to accountability.  Responsibility strongly attached to an actor and its actions. |

**Table 2.2. Richard Owen *et al.* and Alexeï** **Grinbaum and Christopher Groves’ approaches**

|  |  |  |  |
| --- | --- | --- | --- |
| **Approach**    **Parameter** | | **Owen *et al*. (2012, 2013a)** | **Grinbaum and Groves (2013)** |
| **Anticipation** | Vision of the world (*Weltanshauung*) | * Importance of the purposes and the ends of science (not only its means). * Against binary conception of risks/economic & social benefits. * Towards an “informed incremental response” with the help of responsiveness. | Importance of the political and social significance of what scientists intend to do.  RI means taking responsibility in ways:   * quasi parental (implying a limited kind of individual responsibility that focuses upon:   + the duty of care for the malleability of technical artifacts,   + the vulnerability of their future users, * collectively political in nature. |
| Relationship with the future | Beyond pure rationalism.  Uncertainty must be dealt beyond moral luck, beyond traditional risk assessment governance (control, regulation), beyond knowledge based models (as knowledge of consequences and of potential transgression of norms is low) and beyond the “ethics of neighborhood”.  Anticipation is an only “*entry point*” for reflecting on the purposes, promises and possible impacts of innovation. | Foresight is limited. So are our capacities of knowledge. Therefore, foregoing pure consequentialist frameworks is needed. |
| **Tools** | | **Procedural** view of responsibility based on a constant adaptation of means to ends, in order to achieve “informed incremental response.” | Virtue ethics and narratives.  List of virtues scientists should cultivate:  - the duty of care for the vulnerability of future people,  - to bind one’s desire,  - to check ambition by humility,  - to maintain both internal interrogation and external dialogue about the meaning of one’s actions. |
| **Product** | | Applicable to every kind of product.  Question of privacy, safety, justice, usefulness of the product, etc. (c.f. Stilgoe). | Applicable to every kind of product. |
| **Process (of reflexivity)** | Transparency | No specific mention. | No specific mention. |
|  | First order reflexivity: “reflecting on underlying purposes, motivations and potential impacts, what is known … and what is not known.”  No reflection about the “framing”. | Possibly sparked off by the relation of care for others’ vulnerability: ideal of modifying one’s own action (that of scientists or innovators) according to others’ weaknesses. But no clear mention of reflexivity. |
| **Epistemic Tools** | Precautionary Principle (PP) | Need to go beyond mere precaution to include a forward-looking understanding of responsibility. | Neither (because PP is unjustifiable due to limited knowledge). |
| Risk assessment | ×   * Through anticipation of intended and unintended outcomes. * Use of foresight, technology assessment and scenario development. | Nor (also constrained by our limited knowledge. *Cf*. limits of consequentialism) |
| **Assessment** | | Dynamic process of responsiveness. Adaptive way of innovating. | Analogy of the parent-child relationship when parents care for their child’s capabilities and to provide children with the best opportunities without being able to foresee everything what could happen.  Teaching and encoding virtues in created artifacts.  In addition, power of narratives. |
| **Cultural differences** | | Role of the context in elaborating RRI rules. Not a process rule-based in its implementation. Context here = pure exteriority. | No specific mention. |
| **Norm/law relation** | | Norms and law compliance. | Norm production through the relation of care  No specific mention to law. |
| **Participation/deliberation**  **Collective/inclusive/interactive** | | “Collective commitment of care for the future.” Deliberation and public engagement = pillar of RRI. Involving multi- stakeholders. To ensure plurality of perspectives. | Need for responsive dialogue and political deliberation. |
| Responsibility | Liability/blameworthiness |  |  |
| Care | × | Ideal type of the relationship between parents and their children. |
| Responsiveness | Dynamic and flexible process of adaptation. Against rule based application of principles. Care & responsiveness (prospective, way to adapt means with ends) + deliberative democracy (embedded in the concept of responsiveness = respond to others). |  |
| Accountability |  |  |
| Miscellaneous |  | Importance of collective responsibility (H. Arendt like):   * Responsibility for what we have *not* done. * Reason for her Responsibility is her membership in a group which no voluntary act of hers can dissolve. |

**Table 2.3. David Guston’s approach**

|  |  |  |
| --- | --- | --- |
| **Approach**    **Parameter** | | **Guston (2013)** |
| **Anticipation** | Vision of the world (*Weltanshauung*) | * Main pillar of RRI is active anticipation. Against Nordmann’s criticism of “speculative ethics”: “ambiguous priming is preferable to unambiguous surprise”. * Anticipating possible futures helps lead to RRI. * Anticipation = “exercise, practice, and rehearsal, oriented in a non-predictive way toward an undefined future.” * Anticipation = responsible innovation for techno-science governance. |
| Relationship with the future | Technology assessment based on anticipation (both normative and descriptive) and foresight. |
| **Tools** | | * *Foresight* as “a methodologically pluralist approach to plausible futures”. * With an emphasis on scenario development that provide “a more diverse and normative vision compared with other methods that seek to identify a single, most likely future”.   Helps to:   * elaborate possible futures, * make scientists & engineers more reflexive about their research. |
| **Product** | | Focus on emerging technology (above all, nano-biotechnoloy). |
| **Process (of reflexivity)** | Transparency | No specific mention. |
|  | * First and second order reflexivity = reflecting on the ethical issues of technology but also on the framing of those issues. * Anticipation helps researchers & decision-makers to be reflexive: they can focus on the normative dimension of their enterprise. * Embedding social scientists in the making of technology enhances reflexivity. |
| **Epistemic Tools** | Precautionary Principle (PP) | ×  Can be nuanced by anticipatory activities. |
| Risk assessment |  |
| **Assessment** | | * Technology assessment through anticipation. * Elaborating future scenarios leads to define and reveal the normative conceptions attached to technology. * Future scenarios are evaluated today = towards a co- construction of desirable ends. |
| **Cultural differences** | | No specific mention. |
| **Norm/law relation** | | Production of norms through anticipatory activities. |
| **Participation/deliberation**  **Collective/inclusive/interactive** | | * Diversity of participation activities (faithful or not to anticipatory governance) helps to create dialogical spaces. * Need for *upstream* public engagement and dialogue. |
| Responsibility | Liability/blameworthiness |  |
| Care |  |
| Responsiveness | ×  Responsiveness is favored by anticipation. |
| Accountability |  |
| Responsibility | Miscellaneous | Responsibility of the scientist and the engineer today that are consequential for the future. |

**Table 2.4. EC report’s and Hilary Sutcliffe’s approach**

|  |  |  |  |
| --- | --- | --- | --- |
| **Approach**    **Parameter** | | **EC Report (2013)** | **Sutcliffe (2011)** |
| **Anticipation** | Vision of the world (*Weltanshauung*) | * Three-fold RRI definition: knowledge about the consequences, ethical assessment, functional design of technology. * Importance of social acceptability. Meeting social needs. * Economic constraints have a prominent role. | RRI both a process (inclusive) and an outcome. |
| Relationship with the future | * Forecast: importance of acquiring knowledge about consequences. | “RRI is about trying to get better at anticipating problems”.  Defense of an anticipatory governance.  Foresight + scenarios. |
| **Tools** | | * Traditional Risk assessment. * Traditional Ethical, Legal, Societal Implications (ELSI) evaluation. * Creating incentive for ethical concern at the level of the firm (*ethics as a driver, not an impediment*) and for research (*reward for RRI research*). * Economic drivers such as competition should also be concerned by meeting ethical achievements. * The normative content is given by an anchoring in EU values. * Codes of conduct, standards, coordination between Member states, voluntary adoption of norms and standards (flexibility). | * Risk assessment. * Relying on EU norms. * Cost benefit analysis. * Ethical assessment. * Different governance mechanisms:   + Standards, (ISO), CSR, Risk management initiatives, Intellectual property, regimes, scientific committees, *etc*. |
| **Product** | | Applicable to every kind of product. | Focus on emerging technologies: nanotechnologies, genomics, synthetic biology and geoengineering but also “financial instruments, ICT, public policy or community innovations”. |
| **Process (of reflexivity)** | Transparency | No mention. | × |
|  |  | Promoting self-reflection initiatives (where individual freely reflect on the impact of their actions) through codes of conducts, (e.g.“hypocratic oath” of scientists). |
| **Epistemic Tools** | Precautionary Principle |  |  |
| Risk assessment | × | ×  For instance, Real time Technology Assessment + Impact assessment. |
| **Assessment** | | Purely consequentialist. Risk Benefits Analysis. | No lexicographic order between values (ethics *vs* economic interests).  Early involvement is supposed to help solving value conflicts. |
| **Cultural differences** | | No mention |  |
| **Norm/law relation** | | Difficulty to translate RRI into norms.  No production of new norms.  Ethical acceptability includes compliance with existing EU norms. |  |
| **Participation/deliberation**  **Collective/inclusive/interactive** | | Participation has to be promoted because it:   * helps meeting the Grand Challenges of EU, * ensures that societal needs are met, * ensures economic efficiency of products (meeting the target), * creates a dynamic process calling for creativity: e.g. innovation can be a way of overcoming clash between different societal concerns.   Only few mention of public dialogue. | * Involvement of society from the beginning to the end, including the public and environmental groups. * Involvement of citizens (stakeholders) is mostly consumer oriented (firms shouldn’t miss their targets). * Aim: people consider they have been listened to (social acceptability) * Quality of the engagement = need of feedback to participants. * Different devices = Citizen’s juries, brainstorm events, focus groups, partnerships, co-creation or crowd-sourcing initiatives.   Priority given to inclusion, not necessarily to dialogue and deliberation |
| **Responsibility** | Liability/blameworthiness |  | × |
| Care |  |  |
| Responsiveness | Forward looking conception of responsibility. Inspired by Owen *et al*. (2012, 2013a) |  |
| Accountability |  | × |

**Table 2.5. INDECT Project[[4]](#footnote-4)**

In del 2.2., we analyzed the case of the INDECT project aimed at developing technologies related to an enhancement of security for European citizens. To allow for a brief comparison between theories’ analysis and real projects’ analysis, here are the synthetic results of INDECT assessment, using the grid parameters.

|  |  |  |
| --- | --- | --- |
| **Approach**  **Parameter** | | **INDECT** |
| **Anticipation** | Vision of the world (*Weltanshauung*) | Security and control paradigm. |
| Relationship with the future | Traditional foresight of consequences. |
| **Tools** | | Ethical Board (EB) identifying the ethical issues. |
| **Product** | | Security devices. |
| **Process (of reflexivity)** | Transparency | Conflict with secrecy. |
|  | No process of reflexivity related to the way in which ethical norms are produced during the project. |
| **Epistemic Tools** | Precautionary Principle (PP) |  |
| Risk assessment | × |
| **Assessment** | | Through the EB. |
| **Cultural differences** | | No specific mention. |
| **Norm/law relation** | | Compliance with existing norms and laws. |
| **Participation/deliberation**  **Collective/inclusive/interactive** | | Neither participation nor deliberation. |
| **Responsibility** | Liability/blameworthiness | × |
| Care |  |
| Responsiveness | × |
| Accountability |  |

# Chapter 2: Presuppositions of norms in context in current RRI theories

Having decomposed RRI theories along parameters that helped us to characterize the governance approach of responsibility these theories were promoting, we need now to focus more specifically on the way in which they relate norms (of responsibility) with their context. This requires us to identify how each theory conceives the process of norms production and norms application.

In del 2.3. it has been argued that the work of Maesschalk and Lenoble (Maesschalck, 2001; of Maesschalk and Lenoble, 2003) addresses a powerful criticism to most of the contemporaneous theories of governance – whether stemming from sociology, philosophy or economics – that have sought to conceive of the elaboration of social norms in different contexts. One salient point of their argument was that most theories (from neo-institutionalism to Habermassian proceduralism) only focus on the rationality of norms and neglect the issue of their “effectuation” – their “normativity” in Jean Ladrière’s words –, i.e. the concrete ways by which social actors modify their actions according to a proposition carrying a prescriptive value. Procedural theories relying on deliberation (Habermas, 1981; Rawls, 1971, 1993, for instance), suggest that the communicational capacities of social actors and the process of deliberating itself ensure the selection of valid norms. If procedural theories of norms production aim to avoid an objectivation of the world that would lead to predetermined understandings and conceptions of “reality”, the dialogical process by which prescriptive content emerge focuses on the rational acceptability of norms and not on their implementation. Against this view, Maesschalck and Lenoble (2003) constantly argue that the rationality of norms does not exhaust the issue of applicability. Yet, most governance theories assume that once the validity and the rational relevance of norms have been demonstrated, social actors are naturally willing to follow their prescriptive content. Hence, it is needed to open the “black box” of the implementation of norms, i.e. to unfold the concrete processes by which a norm can be implemented, and finally initiate a new research program focusing on a better epistemological representation of the “action of norms”[[5]](#footnote-5).

What is needed, then, as we have seen in del 2.3., is to move beyond the epistemological fallacy according to which a rule or a norm can, by itself, transform the social context that it has to regulate. Such an approach fails because it supposes an “intentionalist, mentalist and schematizing” operation of norms application:

“Intentionalist: because the norm’s effects are supposed to be deductible from the simple intention directing its adoption. Mentalist and schematizing, because what enables the determination of the effect of a norm is supposed to be linked to rules (or schemes to use Kant’s term) located in every mind and therefore a function of mental capacities which do not at all depend on a thinking subject’s exterior context. It is for this same reason that the operation of application can be considered as a simple formal operation of deduction on the basis of the rule itself.” (Maesschalck and Lenoble, 2003, p. 5).

Let us recall the more extensive definition of the Schematizing, Intentionalist, Mentalist (SIM) presupposition provided in del 2.3.

**Table 2.1. The SIM presuppositions**

|  |  |  |
| --- | --- | --- |
| **Schematizing Presupposition**: | **Intentionalist Presupposition**: | **Mentalist Presupposition**: |
| The norms’ application is a simple formal deductive reasoning on the basis of rules themselves. The determination of the norm is linked to these rules, such as ethical guidelines, laws, or other schemes, and are deemed functional to predetermine the effect and therefore the application of a norm. External constraints are not taken into account. | The norms’ effects are supposed to be deducible from the simple intention to adopt the norm. Additionally, we find the implicit presupposition that an actor will have full capacity and intention to contribute to the discussion when involved in a participatory approach. | The norms’ application derive from an imaginary set of rules that the mind is supposed to have. Also, the context does not play any active role and a process ‘interruption’ is considered as an expression of irrational attitudes or behaviors. |

These three presuppositions taken together or not, are probably necessary steps for an adequate theory of norms application: it seems reasonable to suppose, from a cognitivist and a moral point of view, that social actors are willing to act responsibly (at least in some cases) or that they have a capacity of schematization in order to apply rules. The limitation pointed out by Maesschalck and Lenoble challenging most governance theories, consists in reducing the whole process of norms application to one or all of these presuppositions. Governance theories, in the end, close the theoretical inquiry of the effect of norms, in delegating to this set of tools the task to ensure the application of norms. Yet, good intentions, and the capacity of schematization for instance, are probably necessary conditions for the effectuation of norms. However, by no mean can they provide a full theory of norms’ implementation.

**Analysis of current RRI conceptions**

Furthering our analytical understanding of RRI theories and the identification of their blind spots, we need now to uncover the different presuppositions grounding them, i.e. to understand which type of relation to the context they build for the construction of norms. We will follow a systematic analysis of four of the parameters highlighted in the grid (Participation and/or Deliberation, Ethical tool, Norm/Law compliance and the definition of responsibility). Focusing on these specific dimensions of each conception’s normativity and reflexivity, will help us to emphasize how they fall into one of the three SIM presuppositions and to construct the typologies presented in chapter 3.

It must be noted that if all RRI theories rely on the three presuppositions, they differ as for the weight and the emphasis they give to each presupposition.

**Table 3.1. René von Schomberg (2011, a, b; 2013)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| RVS | Participation/deliberation | × | × |  |
| Ethical Tools (EU norms) | × |  | × |
| Norms/Law relation |  | × | × |
| Responsibility | × |  |  |

**Participation and public dialogue** are invoked as if they could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The concrete conditions by which norms would be implemented are not investigated (schematizing).

**The normative content** of this conception is given by an anchoring into EU norms which implies 1) the intention of social actors to follow them and 2) that this set of rules has an intrinsic power of action (intentionalist, mentalist).

**Compliance with law and norms** presupposes that the application of a rule or a norm derives from the rule itself and that what determines the application of the rule already exists in the mind of social actors (schematizing and mentalist).

**Responsibility** is as responsiveness supposes that “technical innovators become responsive to societal needs and societal actors become co-responsible for the innovation process” (von Schomberg, 2013, p. 21). Here, the ability to adapt technological design to societal needs presupposes the good intentions of the actors (intentionalist).

**Table 3.2. Armin Grunwald (2011, 2012)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Grunwald | Participation/deliberation | × | × |  |
| Ethical Tools (explorative philo) |  | × |  |
| Norms/Law relation |  | × |  |
| Responsibility | × |  |  |

**Participation** and deliberation are invoked as if they could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The concrete conditions by which norms would be implemented are not investigated (schematizing).

**The normative content** of this conception is produced by explorative philosophy (as in Guston’s approach) and the effect of norms are considered as a formal outcome of this process (schematizing).

**Compliance with law and norms** presupposes that the application of a rule or a norm derives from the rule itself and that what determines the application of the rule already exists in the mind of social actors (schematizing and mentalist).

**Responsibility** as responsiveness (norms and laws must adapt according to new knowledge) and accountability (actors need to give accounts of their actions) presupposes the good intentions of the actors (intentionalist).

**Table 3.3. Richard Owen et al. (2012, 2013a)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Owen et al. | Participation/deliberation | × | × |  |
| Ethical Tools (incrementalism) | × | × |  |
| Norms/Law relation |  | × |  |
| Responsibility | × |  |  |

Here, **participation** and **deliberation** are both intentionalist and schematizing because it is assumed that the inclusion of stakeholders and public dialogue will, as such (and owing to the good intentions of stakeholders), lead to more responsibility.

**The normative content** of this conception rests on incrementalism (constant adaptation to a changing environment) and implies the good will of actors to adapt and that the application of a rule or a norm derives from the rule itself (intentionalist, schematizing).

**Compliance with law and norms** presupposes that the application of a rule or a norm is logically deduced from the rule itself (schematizing).

**Responsibility** as care for the future and as responsiveness (i.e. the flexibility and capacity to change research and innovation processes according to public values) presupposes the good intentions of the actors (intentionalist).

**Table 3.4. Alexeï Grinbaum and Christopher Groves (2013)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Grinbaum & Groves | Participation/deliberation | × | × |  |
| Ethical Tools (virtue ethics) | × |  | × |
| Norms/Law relation |  | × | × |
| Responsibility | × |  | × |

**Participation and public dialogue** are invoked as if they could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The concrete conditions by which norms would be implemented are not investigated (schematizing).

**The normative content** of this conception materializes as a set of virtues that scientists and innovators should show. This implies their good will to act virtuously and that they have a relevant set of rules in order to apply the chosen set of norms (intentionalist, mentalist).

**Compliance with law and norms** presupposes that the application of a rule or a norm is logically deduced from the rule itself and that what determines the application of the rule already exists in the mind of social actors (schematizing and mentalist).

**Responsibility** as care for the future and care for others’ vulnerability presupposes the good intentions of the actors and their mental capabilities to apply the rules derived from a reflection in terms of virtue (intentionalist, mentalist).

**Table 3.5. David Guston (2013)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Guston | Participation/deliberation | × | × |  |
| Ethical Tools (explorative philo) | × | × |  |
| Norms/Law relation |  |  |  |
| Responsibility | × |  |  |

**Participation** **and public dialogue** are invoked as if they could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The concrete conditions by which norms would be implemented are not investigated (schematizing).

**The normative content** of this conception is produced by explorative philosophy and the effects of the norms are considered as a formal deduction of this process (schematizing). The norms built through explorative activities are meant to be complied with by individuals (intentionalist).

There is no mention of **the role of norms and laws**.

**Responsibility** as responsiveness and accountability (scientists and engineers’ choices of today are consequential for the future) presupposes the good intentions of the actors (intentionalist).

**Table 3.6. EC Report (2013)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| EC Report | Participation/deliberation | × | × |  |
| Ethical Tools (ELSI) | × |  | × |
| Norms/Law relation | × | × |  |
| Responsibility | × |  |  |

**Participation** is invoked as if it could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The effects of norms produced throughout the process is a formal deduction from the process of participation (schematizing). No real focus on **deliberation**.

**The normative content** of this conception is given by traditional ELSI approaches and by an anchoring into EU norms implying that 1) social actors are willing to comply with them and 2) that social actors already have the relevant set of rules to apply these norms (intentionalist, mentalist).

**Compliance with law and norms** presupposes that the application of a rule or a norm is logically driven by the rule itself and that what determines the application of the rule already exists in the mind of social actors (schematizing and mentalist).

**Responsibility** as responsiveness (the ability of innovators and scientists to adapt to social values) presupposes the good intentions of the actors (intentionalist).

**Table 3.7. Hilary Sutcliffe (2011)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Sutcliffe | Participation/deliberation | × | × |  |
| Ethical Tools (ELSI) | × |  | × |
| Norms/Law relation | × | × |  |
| Responsibility | × | × |  |

**Participation** is invoked as if it could, as such, modify the trajectory of the innovation projects at stake (intentionalist). The effects of norms produced throughout the process is a formal deduction from the process of participation (schematizing). No real focus on **deliberation**.

**The normative content** of this conception is anchored into EU norms and rely on existing sets of norms and laws (ISO CSR, intellectual property regimes) implying that 1) social actors are willing to comply with them and 2) that social actors already have the relevant set of rules to apply these norms (intentionalist, mentalist).

No mention of the **role of norms and laws.**

**Responsibility** as accountability (innovators and researchers need to give accounts of their actions) and liability presupposes the good intentions of the actors to act responsibly (intentionalist) and means that the application of norms is a simple formal deductive reasoning based on the conception of responsibility (schematizing).

**Table 3.8. INDECT project**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Presuppositions/ parameters | | Intentionalist | Schematizing | Mentalist |
| Sutcliffe | Participation/deliberation |  |  |  |
| Ethical Tools (ELSI) | × | × |  |
| Norms/Law relation | × | × |  |
| Responsibility | × | × |  |

There is no attempt to include neither **participation** nor **deliberation** in the INDECT project.

The project involves an ethical committee and an ELSI approach **of ethical issues** (intentionalist, schematizing).

**Compliance with law and norms** presupposes that the application of a rule or a norm is logically deduced from the rule itself and that what determines the application of the rule already exists in the mind of social actors (schematizing and mentalist).

**Responsibility** as liability presupposes the good intentions of the actors to act responsibly (intentionalist) and that the application of norms is a simple formal deductive reasoning on the basis of the conception of responsibility (schematizing).

# Chapter 3: Governance models in RRI theories

Now that we have a detailed picture of how each RRI conception addresses the need of a dynamic process of reflexivity and builds the normativity of its possible prescriptions, it is possible to identify the type of governance models they rely on.

## 3.1. The four models of governance

First, we use the four-fold typology of Joly (2001) inspired by the work of Callon (1998) concerning models of technology governance. As ideal types, they are suitable to be applied in concrete situations of technology management (nanotechnology, synthetic biology, etc.) to compare the reality with the model. Here, we will apply these models to abstract, theoretical conceptions of RRI. This implies that some of the characteristics of the four models are not relevant for RRI conceptions (especially what relates with the efficiency of institutions at a practical level). Yet, the Callon-Joly typology is a fruitful tool because it allows us to highlight the degree to which the voice of the public and the necessity of dialogue are taken into account within a given approach of RRI. This will eventually help us to discriminate between different RRI conceptions according to the weight they give to participation and deliberation. It will also be useful to build our own typology of RRI conceptions. Let us recall the typology (cf. Joly (2001) and del 2.3.)

***Standard Model*:** In this model, the disagreements between the experts and the public are perceived as irrational due to the public’s lack of knowledge. There are various reasons for the public to be considered irrational, such as cognitive bias, the lack of comprehension of technical subjects, and aversion to novelties and risk. This model fits perfectly into the classical distinctions between facts and values[[6]](#footnote-6). Experts have an objective approach to risk whereas the risks perceived by the public are marked by a greater degree of subjectivity. In this positivist model, it is necessary to preserve the purity of expertise by not combining facts and value judgments meaning that science and technology are often seen as neutral. Expertise is generally independent from political, economic and social influence. Trust is a central element for the functioning of the system. It is the condition of the delegation of a decision to institutions. Different mechanisms can contribute to the construction of trust. In this model, it is considered that trust gives credibility to institutions and that it is better to contain problems rather than to draw attention to them. The difference in perceptions between experts and the public can be reduced by means of education. It is supposed that people who have more advanced knowledge, especially in ethical disciplines, understand better and adopt experts’ arguments. In this model, risk communication plays an important role. It is related to a one-way method of communication since the experts have little to learn from the public. The objective is to reassure the public to perceive the benefits concealed behind the risks.

***Consultation Model*:**

This model brings into question the fundamental thesis of the standard model, namely the opposition between an irrational public and experts’ rationality. The distance between experts and non-experts is not due to the level of knowledge, but to the difference in risk perception. The public asks wider questions with regard to risk because they are no longer confronted with abstract scientific theoretical risk, but with real risk[[7]](#footnote-7).

It is no longer correct to consider that only experts are rational. Moreover, the experts’ perception of risk encompasses their connections with industry and commercial interests etc. The solutions provided with this model are different from those of the Standard Model. Risk communications and risk management are based on a two-way process. Both the experts and the public have valid views and opinions to contribute. Each side respects the opposition’s insights. In this model, trust is incompatible with a closed, confined, or secretive attitude. To establish trust the public needs to participate in the decision process. Only by engaging the public can regulatory institutions gain legitimacy. In practice, there is a clear distinction between public opinion and the ethical opinion of experts. The public, still seen as irrational, is engaged only in risk management but not in responsibility definition.

***Revised Standard Model*:** this model rests on a greater institutionalization of the management of risk issues. The first two models are based on the atomistic perception of society. Public opinion is seen as a kind of data – the aggregation of individual opinions – and not as a social product, i.e, the outcome of a confrontation between various social groups in the public arena. In this way, the social construction of a problem, as an issue, is omitted. In this model, which is the extension of the standard model, the emphasis is placed on the interaction between the regulation process, social groups and media. Breyer’s model of vicious circle of risk regulation is a good illustration of this model. For him, the legislative process is caught in a vicious circle with the source of the problem being the public attitude towards risk and uncertainty created by the media. Breyer claims that public perception of risk is usually inadequate. Risks are often overestimated, however the efforts to educate the public about scientific risks have failed and will fail in the future. Consequently, responding to public attitude, legislature itself will increase risks and “[…] combined with an institutional inability to set detailed, scientific standards, will cause inconsistent, random, and often irrational ... lawmaking”. As a result, the public will feel unprotected by law and decision-makers, which will lead to more political pressure to take action. In this model, public influence and participation in risk management are considered with great suspicion. Accordingly, risk management includes the following elements:

1. Delegation of risk management to a competent and independent administrative body (in order to avoid the influence of media, pressure groups and politics).

2. Clear distinction between risk assessment and risk management.

3. Risk cannot be measured in an abstract way but rather analyzed, comparing various action scenarios, respecting the general principle of coherence.

4. Trust is not aligned with openness, but rather with reputation and perceptions of competence. Consequently, this model relies on a technocratic vision. The bias against industrial lobbying, the polarization of public opinion and groups of interests, and the reinforcement of independent scientific expertise, represent the elements of the traditional top-down approach.

***Co-construction Model*:**

This model distinguishes itself by questioning the way in which technological development projects use experts. Representations of technology come from numerous collected case studies. The works of the new sociology of sciences have progressively come to blame the traditional conception of science as a revelation of universal, independent truths of the social system that they produce. This important work results from taking the methodological path proposed by different kinds of social constructivism (i.e. David Bloor, Bruno Latour or research stemming from the Society for the Social Study of Sciences). It is therefore a criticism of sound science, which ‘melts’ the analysis of risks in the preceding models, and which invites us to place it into a pragmatic perspective.

In this model, both facts and values are taken into account, as underlined by Stirling (1999), and this is not only a democratic issue; it is a matter of analytical rigor because it is the only way of addressing seriously the issues raised by science, technology and innovation. If not, how can we criticize and validate the framing? Why hide and withdraw from the debate concerning what may eventually be changed? The Co-construction Model usually fits in the *participatory* paradigm, since it requires a participatory approach, such as in Latour’s proceduralism. Once again, this model has some legitimacy, but no real efficiency.

## 3.2. RRI theories between consultation and co-construction

The typologies of Michel Callon and Pierre Benoît Joly seek to grasp governance models of science and technology, focusing on their manner of constructing epistemic knowledge. When considering RRI theories, however, we will have to shift from the drawing of scientific expertise to the building of ethical expertise. In other words, we want to scrutinize how ethical issues are dealt with in RRI conceptions, and what are the links between members of society’s value systems and ethical expertise.

This shift is important to notice because it implies a different kind of normativity at stake. In Joly’s Standard Model for instance, scientific (and positive) expertise is opposed to the (irrational) knowledge of lay people. Transposed to ethical expertise, the relevant opposition will not be the fact/value dichotomy any more – values being present in all ethical assessments – but the distinction between rational/ irrational approaches of ethics. In other words, the standard model of ethical issues’ governance in science, technology and innovation would oppose the rational thinking of ethical experts to the irrational and subjective approach of ethics conducted by “laypeople”. Of course, the type of criticism addressed to this model in the context of scientific expertise also applies to ethical expertise: not only social actors can have rational approaches to ethics, but they are needed in the co-construction of ethical issues and answers. Perhaps even more than in the co-construction of scientific knowledge, laypeople, because they hold different and often conflicting moral positions, can fruitfully contribute to the ethical assessment and design of science and technology[[8]](#footnote-8).

With this remark in mind, let us apply Joly’s typology to our RRI conceptions. First, it appears that no RRI approach relied on the mere exclusion of social actors with a pure delegation of social and ethical issues to experts, meaning that there is no illustration of the Standard Model in current RRI approaches. Recalling the INDECT case[[9]](#footnote-9) is a useful way of highlighting the contrast between the Standard Model and the other models. In del 2.2., we showed that the INDECT project delegated all ethical issues to an expert committee and did not include any of the potentially relevant stakeholders (directly or indirectly affected by the security technology at stake in this project). This method of relating norms with their context has to be compared to approaches where participation and/or deliberation are claimed to be central pillars of responsibility. All current RRI approaches insist on having a (even basic) form of participation and public dialogue. From this starting point, our analysis consists in defining thoroughly the nature and role of participation and deliberation promoted in the different approaches.

Similarly, we found that no approach felt under the Revised Standard Model. No RRI perspective defended the idea of a technocratic delegation of risk and responsibility management, and all approaches insisted upon the need of a direct implication of social actors.

Then, all RRI approaches fall under the consultation and the co-construction model. Most of them ground the governance of innovation and research in consultation. The legitimacy of technological development comes from the possibility for social actors to express their values and value systems. To avoid market failures and the backlash of innovation, consulting end-users and other stakeholders is an essential step. However, if consultation introduces an important improvement compared to the Standard Model, it only focuses on the political process by which societal acceptability may be obtained. In other words, risk management is shared with the public but the very process by which knowledge is constructed, and the definition and implementation of responsibility dynamics are still in the hands of experts. In the EC report, and in Sutcliffe’s approach, for instance, although participation plays an important role, it seems mostly focused on economic efficiency: the co-construction of technology is consumer-oriented to prevent firms from missing their targets. According to von Schomberg’s perspective, participation is more broadly conceived. Yet, it remains unclear how including the public will effectively lead to a re-shaping of technology. In this sense, central to this approach (but also to Grunwald’s and Guston’s approaches) is that the political power of decision-making is shared among the citizens through participation. But the improvement of democratic legitimacy of innovation and research does not necessarily requires social actors to share the power of modifying the shape of technology.

Finally, it appears that two of our RRI perspectives (Owen *et al*. and Grinbaum and Groves) come close to the co-construction model. They both strongly insist on the need for participation and deliberation to result in a co-shaping of technology. Responsibility is conceived in a dynamic way though the idea – in the case of Grinbaum and Groves’ approach – of a “duty for care” towards the vulnerability of future users of technical artifacts whose malleability is emphasized. In Owen *et al.*, responsiveness, and the possibility of a step-by step or incremental adaptation of innovation is aimed at a co-production of technology where scientists are responsive towards social actors’ value whereas the latter understand and take into account the imperatives and constraints of innovators and researchers. Both approaches go beyond the mere idea of political legitimacy and social acceptability and promote concrete ways by which research and innovation can adapt to social wishes and values. However, as we will see in chapter 4, they do not defend second order reflexivity and neglect the issue of challenging the framing by which ethical issues and the question of responsibility is constructed and answered.

Table 3.1. summarizes these results.

**Table 3.1. RRI conceptions according to their governance models**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Governance Typologies | Standard | Revised Standard | Consultation | Co-construction |
| von Schomberg (2011, a, b; 2013) |  |  | × |  |
| Grunwald (2011, 2012) |  |  | × |  |
| Guston (2013) |  |  | × |  |
| Owen *et al*. (2012, 2013a) |  |  |  | × |
| Grinbaum and Groves (2013) |  |  |  | × |
| EC Report (2013) |  |  | × |  |
| Sutcliffe (2011) |  |  | × |  |
| INDECT | × |  |  |  |

**Conclusion**: Keeping in mind that the aim of our analysis is different from that of Callon-Joly’s typology (social processes of technology management *vs* RRI theories, construction of scientific knowledge *vs* drawing of normative knowledge), it seems that the center of gravity of theoretical conception of participation and dialogue has shifted. In other words, there is no one now (at least at the level of RRI *theories*) to defend the Standard Model or the Revised Standard Model. In RRI conceptions, participation, dialogue and the inclusion of stakeholder’s desires, values, or value systems is a prerequisite of any “responsible” development of technology. Yet, what has to be depicted more carefully, is the degrees, levels and modalities of participation and deliberation, the kind of actors included and the criteria according to which the efficiency and legitimacy of inclusion is assessed. This will require us to build a new typology.

## 3.3. A typology of governance models of RRI conceptions

Having applied the parameters of the grid to our main RRI conceptions, two main models of RRI governance emerge. The typology we propose echoes the Callon-Joly typology, departing however from it in several ways. First, we try here to build ideal types of theories and not ideal schemes of real processes of technology management. Among other things, this means that the relevance of the typology is not quantitative but rather qualitative (as the number of RRI approaches is not very high). Second, this typology is a work in progress. The Analytical Grid was the first step, while the next steps will include modifying and refining the typology according to the results of case studies that are currently being conducted (WP3 and WP4). Finally, as the models are ideal types (cf. Weber), they never grasp all the complexity of the reality that they are schematizing. Therefore, RRI conceptions might not fit perfectly in those models (cf. Grinbaum and Groves’ approach, below, for instance), not invalidating the categories but nuancing them. Besides, the differences between the models are sometimes a matter of degree by which a component is promoted. For example, where every RRI conception relies on participation or responsiveness, in most cases, they differ as for the role both pillars play, or as for the degree to which each element participates in the overall process.

To be more precise, five characteristics guided us in the construction of this typology.

1) The process of norm construction.

How norms of responsibility are elaborated in RRI theories? We found that two non-exclusive dimensions can coexist. The first one corresponds to a “substantive” way of producing norms: the RRI theory at stake relies on an already existing set of norms (for instance, the Universal Declaration of Human Rights). The second one rests on a very broad form of proceduralism embedded in responsiveness and participation. Responsiveness understood as a capacity of social actors (often researchers, innovators, but also decision-makers) to include social values is supposed to help producing responsible norms. Participation and deliberation should also lead to responsibility in offering a set of normative constraints and moral judgments broader than scientific, political and economic drivers. Invoking participation, deliberation and responsiveness introduces a procedural dimension in a theory, because the content is not assessed as such but determined by the quality of the procedure: the rightness/goodness of norms depends on the properties of the mechanism of norm’s production. However, and this point will be further analyzed in chapter 4, the various forms of participation and deliberation, the (sometimes conflicting) relations between both processes, and even the specific dimensions of responsiveness are only poorly conceived in RRI theories. They remain at a very general level that do not explore the variety of practical and theoretical perspectives that have emerged in the past 50 years to set a) the conditions of a good and legitimate deliberation (the veil of ignorance and the moral qualities of society’s members in Rawls’ theory of justice or the pragmatic presuppositions in Habermas discourse ethics for instance[[10]](#footnote-10)), or b) the requirements for a legitimate participation (issues of representativity, pluralism, etc.). Therefore, the “procedural” dimension of RRI theories has, here, to be very broadly understood. Finally, RRI theories can rely on both a substantive and a procedural way of norms’ production.

2) The approach of ethics (several degrees of consequentialism, identification of ethical issues as in ELSI perspectives or allowance of a broader reflection on the transformations of society, social actor and the paradigms induced by technology as in the case of explorative philosophy).

3) The role of participation as mere consultation (to ensure social acceptability and/or more legitimacy) or as an attempt towards a co-construction of technology.

4) The relation to knowledge (rationalistic approach of future and uncertainty *vs* explorative philosophy as enabling a normative deliberation).

5) The degree of reflexivity (identification of ethical issues *vs* attempt to open the framing)

### 3.3.1. Model 1 – responsibility grounded in social acceptability

1. **The process of norm construction**

In Model 1, RRI is characterized with a method of norm production that is both substantive and procedural. Substantive because it relies on already existing norms (for example, EU norms of sustainability, quality of life, “highly competitive social market economy, aiming at full employment and social progress”[[11]](#footnote-11) etc. With the EU, the set of norms proposed through the Treaty of Lisbon, among others, has been democratically settled (approved by the European Parliament), but in a discrete fashion, as a social agreement established once for all. This set defines the normative horizon of RRI approaches. Then, below this general normative horizon, more practical norms – those shaping technology, research and innovation – allow for a form of proceduralism that is incorporated into the dynamics of “responsiveness” which in turn rests on inclusion. Under the banner of responsiveness, it is claimed that innovators and other social actors become mutually responsive in an attempt to ensure process flexibility and strengthen the ability to shape research and innovation according to societal values.

As René von Schomberg claims it (2013, P. 21), “technical innovators become responsive to societal needs and societal actors become co-responsible for the innovation process”. This inclusive innovation process rests on an alleged capacity of actors to mutually understand the purposes and ends of other social actors and to adapt their values or value systems to those of the others. In the end, this should contribute to elaborate a common normative horizon, since the overall process is aimed at ensuring a co-construction of “societal desirable products”. In other words, the inclusion of social actors and the alleged capacity of researchers and innovators to respond to social need and to affect technologies’ pathway of development according to social values are elements of the procedure that lead *by* *themselves* to responsibility. Yet, we saw in chapter 2 that this presupposition suffers from being both intentionalist and schematizing, offering an already settled way of relating norms and context.

1. **The approach of ethics**

Another salient trait of this model is that the overall ethical framework is often purely consequentialist, grounded in traditional ELSI approaches and/or focused on the ethical outcomes of given technologies. This would leave, just like in the Standard Model, much of the assessment process in the hand of experts. Indeed, our three examples (EC report, Sutcliffe’s and von Schomberg’s approach) rely on technology assessment and technology foresight to apprehend and rationalize the unpredictable future, implicitly carrying a positivist perspective of risks (that can be objectively determined like in the Standard and Revised Standard Models). Outcomes assessment is an important part of the overall assessment process of technology and the ethical aspects of technology should be identified and listed to eventually be dealt with. Cost/benefit analysis is often highlighted as a useful tool of rationalization. To sum up, the first model of RRI does not offer much more than traditional approaches of technology assessment to evaluate the ethical validity and the social desirability of technology. Moreover, it suppose a delegation to these important issues to sets of experts and does not leave much room for a co-construction of the problem.

1. **The role of participation**

Unlike the standard model, however, Model 1 highlights the need for participation and sometimes deliberation. Yet, participation often means mere consultation, focusing mostly on the meeting of societal needs and the avoidance of market failures. Individuals are asked about their values and value systems in order to ensure that “their views have been listened to and taken into account.” Sutcliffe (2011, p. 11). Processes of stakeholder’s involvement seeks to achieve three targets: a) Social acceptability and the legitimacy for research and innovation that consultation contributes to create; b) social desirability. In addition of a general normative horizon embedded in a given and substantive set of norms, the other source of norm production draws on actors’ values. In this model, ethical relevance comes from social desirability, emphasizing the need for actors’ consultation. Finally, consultation aims at c) avoiding costly market failures in revealing consumers’ preferences. Co-construction and the need for shaping technology according to actors’ values are often mentioned. However, it appears as a mere abstract purpose, a wishful thinking that leaves aside three main issues: 1) the practical conditions through which a mutually responsive process of technology development can be implemented; 2) the normative framework though which value conflict (or value systems conflicts) can be resolved (lexicographic order, costs/ benefits, *etc*.), 3) the possibility of a public dialogue as a process of co-construction of norms[[12]](#footnote-12).

1. **The relation to knowledge**

As for knowledge, the model of responsibility grounded on social acceptability holds a rationalistic framework, in which the future is figured out through technological assessment approaches and technological foresight, although the unpredictability of many outcomes of emerging technology is acknowledged. Building scenarios is one of these tools, but it has not the prominent role (as a normative process) it gains in Model 2. Consequently, the articulation between a representation of the future and a procedure of normative assessment relies on risk assessment more than on the precautionary principle, however, not excluding the recourse of the latter.

1. **The degree of reflexivity**

Finally, the type of reflexivity mobilized by this model comes close from first order reflexivity: responsible actors are supposed to identify key ethical issues as a list of problems to be watched and/or answered. This approach presupposes a deterministic relationship between ethical issues and technology that neglects the constructivist power of social actors to define what is considered as an ethical issue. Since the way in which problems are built closely determine the shape of the answers, we can say that norms are, here, mostly established out of the context.

**Table 3.2. The substantive model of responsibility**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | von Schomberg (2011, a, b; 2013) | EC Report (2013) | Sutcliffe (2011) |
| Participation/  deliberation | Consultation for more:   * Legitimacy (social acceptability) * Social desirability | Early stakeholder’s involvement helps to define social desirable products.  Leads to collective responsibility. | Meeting societal needs (beyond sole economic constraints).  Reducing costs of market failures. | Meeting societal needs (beyond sole economic constraints).  Reducing costs of market failures  Social acceptability. |
| Shaping institutions |  |  |  |
| Shaping technology  (Co-construction ) | Aims at shaping innovation beyond economic constraints and cost/ benefit analysis).  Ethics by design. |  | Aim at co-shaping innovation. |
| Production of norms  (ethical relevance) | Substantive | Focused on ethical acceptability.  **Substantive normative horizon** anchored in EU Norms.  **Institutional tools**: codes of conduct, standards.  **Ethical tools**: ethical assessment. | Focused on ethical acceptability.  **Substantive normative horizon** anchored in EU Norms.  **Institutional tools**: raising awareness for RRI through EU funding.  **Ethical tools**: traditional ELSI approaches. | **Substantive normative horizon** anchored in EU Norms.  **Institutional tools**: codes of conduct, standards, scientific committees, *etc*.  **Ethical tools**: traditional ELSI approaches. |
| Procedural | Responsiveness:  “Technical innovators become responsive to societal needs and societal actors become co-responsible for the innovation process.” | Responsiveness (from Owen *et al*.:  “Responsiveness refers to the flexibility and capacity to change research and innovation processes according to public values.” |  |
| Ethical framework |  | Consequentialist. | Consequentialist. | Consequentialist. |
| Relation to knowledge (and their related epistemic tools) | Rationalistic | Foresight in context of unpredictable outcomes ( → Epistemic tools =TA, F + PP). | Traditional RA , TA. | Traditional TA or RTA, impact assessment.  Foresight. |
| Anticipation as a normative tool |  |  | Scenarios, “Anticipatory Governance.” |
| reflexivity | Content reflexivity (1st order) | Dealing with ethical issues raised by technology and innovation. | RRI implies to  “evaluate both outcomes and options in terms of ethical values (including, but not limited to well-being, justice, equality, privacy, autonomy, safety, security, sustainability, accountability, democracy and efficiency). | Traditional ethical assessment (ELSI approaches). |
| Framing reflexivity ( 2nd order) |  |  |  |
| Responsibility |  | Collective responsibility. | Collective responsibility evoked. |  |

**Legend**

F: Foresight

PP: Precaution Principle

RA: Risk Assessment

RTA: Real Time Assessment

TA: Technology Assessment

### 3.3.2. Model 2 – responsibility through responsiveness and deliberation

1. **The process of norm production**

Model 2 only supports a **procedural determination of norms**, based on r**esponsiveness** and **inclusion[[13]](#footnote-13)**. Facing uncertainty and adapting to rapidly changing innovation requires an incremental, and step by step mutual awareness of innovators and researchers for social values and of social actors for innovation. Moreover inclusion is realized by means of deliberation that sometimes is favored through explorative philosophy (Grunwald, Guston) and sometimes just mentioned as such (Owen *et al.*).

In all cases, however, the quality of the results does not lie so much in the content of the set of norms regulating technology (such as in model 1) but in the quality of the procedure. Sometimes, care is added as a way of dealing with the specific issues of innovation and research (unpredictable outcomes in context of limited knowledge). If care could shift the model towards a substantive determination of norms (when virtue ethics is explicitly invoked such as in Grinbaum and Groves (2013)), it can also qualify the rightness of the procedure. Consequently, participation and deliberation are central to model 2 not only as a mode of consultation, but as a commitment and involvement of social actors who reenact the potential of a democratic handling of science and technology. Thus, responsibility is collectively conceived and shared. Responsibility is no longer purely associated with one’s own actions, but rather attached to the dense net of causal chains underpinning events.

1. **The role of participation**

Here, participation is not only meant to ensure the legitimacy of innovation. It aims explicitly at **achieving a co-construction of technology**, involving scientists, social scientists, “stakeholders”, and, more generally, social actors who are no longer mere spectators, but influential and decisive participants. However, acknowledging the practical difficulties of co-constructing research and innovation, a less ambitious but allegedly more achievable end can be the co-shaping, if not of technology, at least of the institutions that mark out the interface between science and society.

1. **The approach of ethics**

This model relies on **anticipatory governance and explorative philosophy as normative tools**. Anticipating possible futures and building scenarios does not only help to face the challenge of uncertainty, but also reveals and builds common normative horizons (and more practical norms). The construction of scenarios is supposed to help to co-construct the problem but also to allow social actors to express their values and value systems, to discuss and argue for them, and perhaps to achieve some sort of agreement. In other words, explorative philosophy is used an enabler of deliberation.

1. **Relation to knowledge**

Against a purely consequentialist and rationalistic relationship with knowledge, the power of imagination and narratives helps different normative systems to express and unfold their complexities. Model 2 aims at going beyond traditional tools of TA, (even real time TA), and ELSI approaches to avoid the paradoxes of knowledge and the limits imposed by radical uncertainty. Promoting responsiveness and deliberation for RRI is one way of overcoming the limits of rationalistic frameworks.

1. **The degree of reflexivity**

Finally, model 2 tends to 2nd order reflexivity that is defined as following by Oran Perez in the context of regulation[[14]](#footnote-14):

« Second-order reflexivity involves an attempt to take a step back from the substantive debates among theoreticians and practitioners of responsive or reflexive regulation in order to examine the presuppositions and commitments that are shared by those who engage in the debate.”(Perez, 2011, p. 760).

In addition to the need of addressing the ethical issues of innovation and research, the “opening up of visions, perspectives and dilemmas” or the possibility of “changing the framing and venue” of issues is an endeavor of reflecting on the two dimensions of the context. Compared to model 1, there is a step further here, as it becomes possible to open the reflection of social actors to the construction of the problems and not only to the finding of solutions[[15]](#footnote-15).

**Table 3.3. The procedural model of responsibility**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Owen *et al*. (2012, 2013a) | Guston (2013) |
| Participation/  deliberation | Consultation   * Legitimacy (social acceptability) * Social desirability |  |  |
| Shaping institutions | Towards a collective responsibility enabled by participative democracy.  Actors are influential and decisive. |  |
| Shaping technology  (Co-construction ) | Aims to *co-produce* innovation through participation. | “Co-production of a new technoscience.” |
| Production of norms  (ethical relevance) | Substantive |  |  |
| Procedural | Care and responsiveness underpin RRI  “Importantly, it not only embeds the concept of responding to a changing information environment, i.e. being *adaptive, but also to responding to the views, perspectives and framings of others* – publics, stakeholders- i.e. being *deliberative*”.  Incremental response against rule-based application of principles (sensitivity to context, flexible in the face of uncertainty). | Anticipatory governance leads to RRI  It includes “a non­linear, co-constructive approach, flexibility in the face of scientific serendipity, an avoidance of top-down policy prescriptions, foresight and flexibility under uncertainty), encouragement of public engagement, participatory technology assessment, and responsible innovation” (p. 11). |
| Ethical framework |  | Limited consequentialism: not only outcomes of actions matter, but also the quality of the process of RRI (care & responsiveness beyond reciprocity and consequentialism): RRI as a “collective commitment of care for the future through responsive stewardship of science and innovation in the present.” | Consequentialist. |
| Relation to knowledge (and their related epistemic tools) | Rationalistic |  |  |
| Anticipation as a normative tool | Foresight & TA are only “entry points”.  Scenarios help to draw desirable and undesirable worlds, as well as our normative horizons. | Anticipation contributes to the production of new norms (normative horizons but also practical norms).  It enhances the reflexivity of scientists. |
| reflexivity | Content reflexivity (1st order) | Dealing with ethical issues raised by technology and innovation. | Dealing with ethical issues raised by technology and innovation. |
| Framing reflexivity ( 2nd order) | Towards a 2nd order reflexivity: a deliberative process that opens up “visions, purposes, questions and dilemmas” ensures “the introduction of a broad range of perspectives to reframe issues and the identification of areas of potential contestation.” | “such proceduralism […] answer(s) the ethical and political questions of whether or not a given course of action is good or bad, right or wrong, just or unjust […] because it provides a different context – changes in framing and in venue – in which the researchers, as decision-makers, have the oppor­tunity to focus on the normative dimensions of their enterprise” (p. 17). |
| Responsibility |  | Responsibility is enabled by care and responsiveness.  Collective responsibility/ Co-responsibility between different actors, institutions, etc. |  |

|  |  |  |
| --- | --- | --- |
|  |  | Grunwald (2011, 2012) |
| Participation | Consultation   * Legitimacy (social acceptability) * Social desirability | Participation enhances trust from society for science and technology. |
| Shaping institutions | *Interdisciplinary ethical reflection* plays a prominent role. Participation is required to incorporate social values into innovation.  Interface between nanos and society has been shaped by ethical reflection. |
| Shaping technology  (Co-construction ) | Aims at shaping innovation instead of shaping technology (motto of TA). |
| Production of norms  (ethical relevance) | Substantive |  |
| Procedural | Incrementalism: (step-by-step, case-by-case laws and norms elaboration in the case of nanotechnologies, for instance).  Proceduralization= “small steps toward extending our knowledge of the possible health or environmental consequences as rapidly as possible and in an ongoing evaluation and assessment of the knowledge situation. The resulting procedure would be an *incrementalism* that is reflective and obligated to constantly question itself.” |
| Ethical framework |  | Consequentialist. |
| Relation to knowledge (and their related epistemic tools) | Rationalistic |  |
| Anticipation as a normative tool | Futurology, i.e., anticipating future scenarios of evolution enhances responsibility. However it is needed to assess the inter-subjective and trans-subjective argumentative quality of such processes. |
| reflexivity | Content reflexivity (1st order) | Dealing with ethical issues raised by technology and innovation. |
| Framing reflexivity ( 2nd order) | Towards a 2nd order reflexivity through the process of constant adaptation and of calling norms into question. |
| Responsibility |  | Traditional view of responsibility (attribution to an actor’s actions according to available knowledge). |

**Specific case of Grinbaum and Groves**

|  |  |  |
| --- | --- | --- |
|  |  | Grinbaum and Groves (2013) |
| Participation | Consultation   * Legitimacy (social acceptability) * Social desirability | No mention of participation.  The need for political deliberation rapidly evoked. |
| Shaping institutions |  |
| Shaping technology  (Co-construction ) |  |
| Production of norms  (ethical relevance) | Substantive | Virtue ethics – virtues of scientists. |
| Procedural | Care and responsiveness (constant adaptation with concern for future generation and vulnerable beings) underpin RRI. |
| Ethical framework |  | Limited consequentialism: not only outcomes of actions matters, but also the quality of the process of RRI (care and responsiveness beyond reciprocity and consequentialism): RRI as a “collective commitment of care for the future through responsive stewardship of science and innovation in the present.” |
| Relation to knowledge (and their related epistemic tools) | Rationalistic |  |
| Anticipation as a normative tool |  |
| reflexivity | Content reflexivity (1st order) | Dealing with ethical issues raised by technology and innovation. |
| Framing reflexivity ( 2nd order) |  |
| Responsibility |  | Responsibility is enabled by care & responsiveness.  Collective responsibility/ Co-responsibility between different actors, institutions, etc. |

# Chapter 4: In search of a reflexive governance theory for RRI

How can we interpret these two models of RRI theories in light of the SIM presuppositions?

## 4.1. RRI models of responsibility in light of the SIM presuppositions.

Model 1 promotes a form of participation which is incomplete, focusing on the acceptance of society more than on the ethical relevance and validity of the system. Model 2 is an attempt to overcome these limits. In promoting a procedural framework where participation and dialogue are aimed at ensuring co-construction, it relies on a deeper conception of democratic science and technology. However, this model does not resist the line of attack developed by Maesschalk and Lenoble (cf. del 2.3.). Promoting responsiveness and/or care as well as anticipatory governance as the main motors of norms production inevitably presupposes that the relation to norms is mentalist, schematizing and intentionalist. The very process by which social actors will come to define and comply with innovation and research norms is only very generally tackled. For instance, how to ensure the ethical relevance of anticipatory governance? Exploring possible scenarios may contribute to the building of a common normative horizon. But how to assess the ethical validity of this bedrock? We would not like to let people think RRI conceptions do not provide any substantial improvement. The essence of the relationship between science and society has been strongly reconsidered within RRI theories and the “center of gravity” of these conceptions shifted from inclusion as a mere legitimation process towards an appraisal of efficient co-construction. But, somehow, this naive belief in a well-intentioned deliberative democracy[[16]](#footnote-16) is still at work.

Participation and deliberation whether as mere consultation or closer from the model of co-construction, may be the basis of the dynamic process of responsibility. But the practical process by which conflicting values are incorporated in the making of technology is never made explicit.

## 4.2. Maesschalck and Lenoble’s conception of reflexivity

Let us try to make this point clearer by recalling Maesschalck and Lenoble’s perspective on reflexive governance.

We already recalled their case against a purely rationalistic apprehension of the moment of the effectuation of norms. We also emphasized how they deny the possibility of rationally anticipating all the possible transformations of an existing living form that are necessarily entailed by the adoption and implementation of a norm. In their words, for a sake of epistemological validity, it is necessary to assume that reason, to exert some possible effects on the real world, relies on something different from it, more precisely on a belief in its power to transform reality.

“This means that every judgment (and therefore every norm since a norm is obviously the result of a judgment) can only be applied and produce a meaning (and therefore effects in reality) at the end of a complex operation which is an exchange between the effects expected by the abstract norm on the one hand and on the other effects rendered possible by the coherences specific to the existing way of life.”(*op.cit*. p 113)

The “other effects” to be involved in the application of norm refer to the effective possible states of the world that may result from the social application of the norm. These effective possibilities are conditioned by the structure of the context, that have to be constructed though a common “moment” of reflexivity that never can be fully anticipated by the formal rules of a dialogical process.

**Reflexivity**

Here we come to one of the most important points of the criticism made by Maesschalck and Lenoble to most of modern theories of norms production (to which we include RRI): their understanding of reflexivity.

The point of departure is that every proposition is a selective « representation » of what makes up our world (for instance, a particular combination between human and non-human beings in Latour’s conception). These various representations are *a priori* in infinite number, and the construction of a common world necessarily involves a selection of these propositions and a negotiation between them, conducted by social actors. The fundamental claim of Maesschalk and Lenoble is that the operation of selection and of negotiation cannot be anticipated *a priori* by some rationalistic calculus.

More precisely, defining a rule implies the identification of an objective to achieve, and hence, a problem to solve. *The construction of the problem*, i.e. the identification of its meaning and the choice of the appropriate behaviors that may help solve it, result from an operation of selection that cannot only come from the formal constraint of a rational justification (*op. cit*. p. 348). It is thus necessary to reflect on the ways to negotiate solutions to, but also to construct, problems. How is the necessity to take action identified among different actors, how is it designated, how are solutions chosen, and how is it affirmed that the solutions will meet the objective implicitly defined by the rule? All these choices cannot only be left to an *a priori* exercise of the reason. The selection of possibilities has to be reflexively constructed out of pure rationalistic anticipations and determinations, as it is supposed by either rational choice theory or theories of deliberation[[17]](#footnote-17).

But what does “reflexivity” mean? Reflexivity is defined as an operation proper to any act of judgment by which the latter is conceived as a “power of possibilization”. It consists in seizing the act of judgment itself in its activity, in its semantic productivity with respect to a given context whereby social actors identify the various possible states of the world that will determine the choice for a rule or a norm. Reflexivity relates to the capacity, of any judgment, to rely on elements that are external to reason in order to product effects on reality – in Maesschalck and Lenoble’s words, “effects of meaning” – in a precise context (*op.cit*, p. 7).

Reflexivity, then, implies a capacity of reconstructing in a reversible way the representations that define the collective equilibriums of a particular living form that has been accepted among a community. This implies that 1) the agenda of the problem at stake (for which the norm or the rule is elaborated) has to be constructed collectively by the concerned actors (and not imposed *a priori* by a rational construct) and 2) that actors themselves “have to be reflexively constructed in relation with their context” (*op.cit*. p. 313), meaning, among other things, that they cannot be reduced to their rational predictable preferences (as neoclassical economics does).

These two conditions enable a reflexive stance through which the perception of the possible states of the world that are accepted by social actors are questioned. A better conception of the effectuation of norms avoiding the mentalist and schematizing error, gives more consideration to the anticipations that are made by social actors regarding the identification of the problems at stake – that cannot be fully given by the formal operation of reason, i.e., for instance, settled by a set of experts, be the most enlightened – and the solutions that can be proposed. It is required from social actors, for whom the rule is designed reflexively, to draw the various possible worlds that might be transformed by the rule – possible worlds that depend on the ways social actors identify problems and solutions, i.e. the transformation of reality implied by the prescriptive rule.

To sum up, ensuring a reflexive governance of technology then aims to “force the decision makers who adopt such a technological development to create the conditions for a negotiation with a view to allowing the inclusion of the diverse dimensions which condition the effective use of this technology and which therefore condition its consequences and the transformations required if these consequences are to be as positive as those which the simple abstract examination of their relevance spontaneously attributes” (*op. cit.* p. 324). Only a common negotiation of the question to be asked and solved and of the ways in which norms will be implemented can favour an effective transformation of a given state of affairs.

## 4.3. Reflexivity in RRI theories

Now, how far do current RRI theories comply with the conception of reflexivity developed by the authors of the *Action of Norms*?

Tables 2 (chapter 2) give a synthetic account of the different interpretations of reflexivity held by the leading authors of RRI conceptions. Mostly, they focus on the outcomes of innovation whether known or not (in the case of radical uncertainty), though: reflexivity includes the traditional ELSI (or ELSA, Ethical, Legal and Social Aspects) ways of dealing with the “issues” of technology (EC Report, Sutcliffe, von Schomberg), but also a questioning of the purposes and motivations underlying technology (Owen *et al*.). The latter opens a possibility of questioning the *framing* of the actors (through a discussion about the ends of technology). However, the “Copernican revolution” by which the co-construction of the context of norms would be the main purpose (actually the *raison d’être*) of inclusion is not at work, yet.

The three remaining perspectives go a step further in attributing (explicitly or implicitly) space for reconsidering the processes by which norms are constructed. David Guston, for instance, argues that anticipatory activities enhance the reflexive capacities of researchers and other social actors in “exploring and assembling current values, knowledge, and plausible scenarios” (Guston, 2013, p. 15). Explorative activities do not only increase our understanding of possible scenarios but they also contribute to the elaboration of a common normative horizon: social actors are given room for expressing their values and value systems, discussing complex issues, exchanging perspectives and adjusting their visions to that of others.

Grinbaum and Groves (2013) also propose to rely on narratives that trace out and unfold” how the consequences of action ramify through time in unpredictable ways” (*op. cit.*, p. 139). In line with Maesschalck and Lenoble, they claim for a limited power of foresight, and – as we mentioned it already in del 2.2 – draw on the philosopher Hannah Arendt to emphasize the finitude of human beings and the limits of their rational capacities of foresight. These premises would call for “rational prudence” and for a conception of RRI based on virtue ethics rather than on a consequentialist framework. But even developing the virtue of responsibility, together with integrity, impartiality, honesty, and all the other virtues that could be required for innovators and researchers, the possibility of moral luck[[18]](#footnote-18) can always turn good intentions into bad outcomes. This is where narratives are one way of building “reflective skills”, which help to challenge the “cultural meaning” of innovation.

Explorative activities, fostering a discussion that is not *a priori* reduced to the outcomes of technology, but that also includes the different meanings it takes on, allow for reflexivity. Clearly, the two above views seek to step out from *given* ethical issues (that simply pop up or have to be discovered), and reject the idea that the construction of the problem is ontologically bound to the properties of technology regardless of its social fabric. On the contrary, a common reflection about future possible scenarios allows actors to argue about their value and value systems – but also to mobilize their narrative, interpretative and reconstructive capacities as argued by J.M. Ferry – while assessing technology, its purpose, outcomes and desirability. Explorative philosophy can be a powerful tool towards a co-construction of the context.

Defending another process that would challenge our framings, Armin Grunwald highlights the virtues of the ethical debate, which in the specific case of nanotechnology did not result in a co-shaping of technology, but at least allowed for a reconfiguration of the dialogue between science, politics and the public. The purpose of inclusion here goes beyond mere consultation and aims to influence (even if only partly) the way in which the problem is conceived: i.e. “shaping the interface between nanotechnology and society”, by means, for instance “of the early demand that research be conducted into the possible toxic properties of nanoparticles, by opening the debate on the relationship between humans and technology in light of the possibilities for technical “enhancement”, and by insisting on the opportunities for participation and on dialog” (Grunwald, 2013, p. 331). The nano-debate emphasized the need to move forward the expert battleground to allow social actors to express and define for themselves what the core “problems” to be investigated were. As such, if the debate did not directly influence the design of nanotechnology, it led to changes in how the issues raised by nanos are dealt with.

These three conceptions of RRI include an intuition regarding the necessity to reflexively construct the context in which responsibility will unfold. In different ways, though, the mentioned authors seek to seize the process by which it is possible not even to shape the design of technology but more fundamentally to co-construct the way in which the issues connected with technology are conceived by different social actors (including researchers and innovators). However, they do not address in a systematic way the question of the implementation of norms, which, in the reflection of Maesschalck and Lenoble, leads to the necessity to reflexively construct the actors and the problems the RRI norms would solve. In other words, the kind of reflexivity these theories put forward, does not give much tools for a common elaboration (interpretation, reconstruction) of norms. Current RRI theories do not explore the complex issues raised by participation and deliberation related with the conflicts of values, motivations and purposes that emerge when gathering differently-driven social actors. In a way, participation and deliberation are deemed to ensure reflexivity, but the “black box” of the means by which reflexivity is achieved is not open. To this respect, it is quite significant that most of the conceptions of RRI suppose that the reflexivity of actors is given: i.e. the conditions by which actors can think about themselves in relative ways are not thought of. Finally, what is missing includes some insight on how the “capacitation” of actors is ensured, how institutional design can build their capacity for reflexivity, and a theory of institutional learning by which the design of participation and deliberation processes can be corrected and modified according to their capacity to ensure a quality dialogue.

## 4.4. The missing reflexive theory of governance

To sum up, the blind spot of current RRI conceptions are linked with the role and the development they assign to participation and deliberation. Often, these two facets of inclusion – a central “pillar” of RRI as claimed in del 2.2. – are supposed to provide all the normative content that is required when speaking about responsibility in research and innovation. But participation and deliberation raise more issues than they answer questions as for the concrete conditions by which a responsible and ethically sound decision can be made out of participation and/or deliberation. For example, moving beyond mere consultation implies that “participants” can be decisive in the final choice. Yet, on which grounds is the “mitigation” between conflicting values to be made? In procedural deliberative theories (Rawls or Habermas, for instance) the intrinsic dynamic of (rational) argumentation is meant to help in building a common ground of norms. But, following Maesschalck and Lenoble, what happens if we have to abandon reason as the unique bedrock of norms’ application? How to practically design participative or deliberative processes that would both avoid the evils of cacophony supporting the survival of the fittest and of a full orchestrated score constraining the participants to play a part already written for them?

Another aspect revealing the still early stage of RRI theories’ conceptual development lies in their common under-theoretical understanding of participation and deliberation that, in addition, are not precisely distinguished. Deliberation is quasi-absent from the proposition made by the EC report (2013) and by Sutcliffe’s approach while it is more or less explicitly mentioned in all other theories. However, most of the conditions and requirements of a legitimate deliberation investigated by modern political theorists and political philosophers such as James Fishkin and Robert Luskin (2005), Dennis Thomson (2008), Joshua Cohen (1989, 1997), Jane Mansbridge (1983), Charles Girard (2010a), John Rawls (1971, 1993) or Jürgen Habermas (1981, 2008) are not even mentioned. Also absent of the reflection are the possible contradictions between participation and deliberation. For instance, including more participants can threaten the possibility of agreeing on something or reduce the quality of the argumentation (Girard, 2010b), for opening one’s own deliberation to that of others creates similar conflicts as the divergences that can oppose moral and political deliberation (Reber, 2011).

Also, if participation is generally conceived as a way of including social actors’ values, deliberation is only very narrowly associated to “public dialogue.” Yet, public dialogue as such does not ensure a qualitative argumentation. And the requisites needed to ensure that public engagement or public dialogue results in quality argumentation (an important dimension of deliberation) are missing in the current reflection on RRI. Finally, if RRI theorists suggest by their vocabulary that the ideal of deliberative democracy infuses their conceptions, the complex mechanisms of such political processes are never defined. Yet, the theory of deliberative democracy could be of great help for RRI conceptions, even if it carries some weaknesses – namely an under-determination of the argumentation (Reber, 2012a) and difficulties raised by a co-argumentation in an interdisciplinary context such as the one RRI conveys – highlighted both at a normative level (by philosophy) and at empirical level (by sociology and political sciences). In spite of that, promising work is done by the collective research led by Jürg Steiner (Steiner *et al*., 2004; Steiner, 2013) that analyses the different characteristics of deliberative democracy: participation, stories (narrative), common good justification, respect, public openness, force of better argument, truthfulness. Recalling the controversies these characteristics initiated but also the empirical results they are related with, this research exposes what are the conditions of deliberation and their outcomes in real life, supported by a robust *Discourse Quality Index* (2 versions). This type of enquiry, associated with a deeper concern for theoretical conceptions and empirical experience of deliberation could be of much helps for RRI theories.

Finally, we can also notice that, paradoxically, most of RRI theories only very rapidly define the concept of responsibility. If we remember the different meanings or understandings of responsibility that we mentioned in del 2.2., current RRI theories broadly adopt responsiveness and accountability (although not always explicitly) as the main layers of responsibility. But they never dig in depth to the implications of these conceptions. How can someone be held responsible? In complex causal chains, how is it possible to assign responsibilities to different social actors? In innovation and research projects are individual, collective or both responsibilities more relevant? If responsiveness designates a constant adaptation and the ability to respond, how can it capture the kind of moral constraint that is involved in responsibility? When responsibility is understood as responsiveness, it sometimes can be interpreted as a cautious way of moving forward (as suggested by the idea of “incrementalism”). However, being cautious does not always protect from irreversible choices. Moreover, as emphasized by Oran Perez (2011, p. 771) in the context of regulation, requiring flexibility (or responsiveness) of regulators (or of innovators and other social actors) does not give them much about the ways to increase their flexibility and their capacities for learning. Actual RRI conception leave in the dark the overall question of a systemic learning favoring continuous improvement, which is though necessarily implied by the idea of responsiveness.

Isn’t it, then, that procedural conceptions of responsibility abandoning the notion of an agent holding his part of responsibility end up as empty boxes? Finally, cannot delegating the construction of normativity to stakeholder’s inclusion be interpreted as a way of avoiding the normative tasks of defining what is responsible or irresponsible?

# Conclusion

In the end, we are left with more questions than answers because a theory of reflexive governance of RRI in which norms are constructed within the context is still missing. Whether relying on a substantive or a procedural process of norms’ production, RRI theories remain relatively blind to the mechanism by which their normativity is created. When the substantive dimension is favored, theories of responsibility delegate the issue to a previously accepted set of norms without challenging them or without defining, for instance, a hierarchy between conflicting values such as sustainable development and competitiveness in a global market economy. While including a form of proceduralism, theories of responsibility rely on responsiveness and anticipative activities to address conflicting values. They remain at a very general level and then avoid a fine and precise discussion about the conditions, the different requirements (institutional settings and capabilities of actors, among others) that should be implemented to ensure a responsible construction of norms.

This negative conclusion should not be understood as a rejection of current RRI theories. Our aim has been here to emphasize some missing aspects of early research on RRI, with an attempt to shift the focus from “ingredients”, “recipes” of RRI towards this central issue of reflexivity understood as the conditions for a common construction of the issue of responsibility. With this work, we hope to impulse a move from a reflection in terms of a *definition* of RRI towards the question of the *governance* of responsibility. In other words, as highlighted by Owen et al. (2012), responsibility designates the practices of scientists, innovators, industrials or engineers involved in project of RI and the processes by which projects address the social and ethical issues raised by the technology they develop. But it also can define the process by which norms are constructed. And a responsible way of constructing norms will suppose a reflexive construction of actors and the context, as well as a room for deliberation that still has to be characterized and institutionalized (Reber, 2012b).

# References

Callon, M. (1998). « Des différentes formes de démocratie technique », *Annales des Mines*, January, pp. 63-73.

Callon M., Lascoumes P., Barthes Y. (2001) *Agir dans un monde incertain. Essai sur la démocratie technique*. Paris: Le Seuil.

Cohen J. (1996) [1997]. *“*Procedure and Substance in Deliberative Democracy.” *In* J. Bohman and W. Rehg (eds.), *Deliberative Democracy: Essays on Reason and Politics*, Cambridge, MA: MIT Press, pp. 407-437.

Cohen, J. (1989). “Deliberation and democratic legitimacy.” *In* A. Hamlin and P. Pettit (eds.), The Good Polity, Oxford: Blackwell, pp. 17-34.

Ferry, J-M. (1991) *Les puissances de l'expérience,* 2 vol. Paris: Éd. du Cerf.

Ferry, J.-M. (1991) *Les ordres de la reconnaissance*. Paris: Éd. du Cerf.

Ferry, J.-M. (2002) *Valeurs et normes: La question de l'éthique*. Bruxelles: Éditions de l'Université de Bruxelles.

Fishkin, J. S., R. C. Luskin (2005). “Experimenting with a democratic ideal: Deliberative polling and public opinion”, *Acta Politica*, 40(3), pp. 284–298.

Girard C., Le Goff A. (ed.) (2010a). *La démocratie délibérative. Anthologie de textes fondamentaux*, Paris, Hermann.

Girard, C., (2010b). *L’idéal délibératif à l’épreuve des démocraties représentatives de masse. Autonomie, bien commun et légitimité dans les théories contemporaines de la démocratie,* PhD., Université de Paris 1.

Grinbaum, A. and Groves, C. (2013). “What is the ‘responsible’ in responsable innovation”. In, R. Owen *et al*. (eds) (2013b), pp. 119-142.

Grunwald, A. (2011) “Responsible Innovation: Bringing together Technology Assessment, Applied Ethics, and STS research”, *Enterprise and Work Innovation Studies*, 7, pp. 9-31.

Grunwald, A. (2012). *Responsible Nanobiotechnology*, Pan Stanford Publishing.

Guston, D. (2013). “Understanding ‘anticipatory Governance’”, *Social Studies of Science*, p. 1-25.

Habermas J. (1984) [1981]. *Theory of Communicative Action Volume One: Reason and the Rationalization of Society,* Boston, Mass.: Beacon Press.

Habermas J. (1987) [1981]. *Theory of Communicative Action Volume Two: Liveworld and System: A Critique of Functionalist Reason*. Boston, Mass.: Beacon Press.

Joly, P.B. (2001). « Les OGM entre la science et le public ? Quatre modèles pour la gouvernance de l’innovation et des risques », *Economie rurale*, n°266, p. 11-29.

Latour, B. (1987). *Science in action: How to follow scientists and engineers through society.* Cambridge, Mass: Harvard University Press.

Lenoble, J., Maesschalck, M. (2003). *Toward a Theory of Governance: The Action of Norms.* Kluwer Law International.

Maesschalck, M.( 2001). *Normes et Contextes*, Georg Olms Verlag.

Mansbridge, J. J. (1983). *Beyond adversary democracy*. Chicago: University of Chicago Press.

Owen R., Baxter D., Maynard T., Depledge M.H. (2009a). “Beyond regulation: Risk pricing and responsible innovation”, *Environmental Science and Technology*, 43(14), pp. 5171– 5175.

Owen R., Crane M., Deanne K., Handy R.D., Linkov I., Depledge M.H. (2009b). “Strategic Approaches for the Management of Environmental Risk Uncertainties Posed by Nanomaterials”. *In* I. Linkov, J. Steevens. (eds.) *Nanotechnologies: Risks and Benefits*, Dordrecht: Springer.

Owen, R., P. Macnaghten, J. Stilgoe (2012). “Responsible research and innovation: From science in society to science for society, with society”, *Science and Public Policy*, 39, pp. 751–760

Owen, R., P. Macnaghten, J. Stilgoe, M. Gorman, E. Fisher, D. Guston (2013a). “A Framework for Responsible Innovation”. *In* R. Owen *et al.* (eds.) (2013b), pp. 27-50.

Owen, R., J. Bessant, M. Heintz (eds.) (2013b). *Responsible Innovation. Managing the Responsible Emergence of Science and Innovation in Society*. London: John Wiley.

Ralws J., (1971). *A Theory of Justice.* Cambridge, Massachusetts: Belknap Press of Harvard University Press.

Ralws J. (1993). *Political Liberalism, The John Dewey Essays in Philosophy, 4.* New York: Columbia University Press.

Reber, B., (2011). « Les risques de l’exposition à la délibération des autres », *Dialogue et pouvoir*, *Archives de philosophie du droit*, Tome 54, p. 261-281.

Reber, B. (ed.), (2012a). *Vertus et limites de la démocratie délibérative*, Paris, *Archives de philosophie*, 74/1.

Reber, B. (2012b). « Argumenter et délibérer entre éthique et politique », in Reber B. (dir.), *Vertus et limites de la démocratie délibérative, Archives de Philosophie*, Tome 74, avril-juin, p. 289-303.

Report of the Expert Group on the State of Art in Europe on Responsible Research and Innovation, 2013, “Options for Strengthening Responsible Research and Innovation”, Chair: Jeroen van den Hoven; Members: Linda Nielsen, Françoise Roure, Laima Rudze, Jack Stilgoe; Rapporteur: Klaus Jacob.

Steiner, J., Bächtiger, A., Spörndli M., Steenbergen M. R. (2004). *Deliberative Politics in Action. Analysing Parliamentary Discourse,* Cambridge (UK), Cambridge University Press.

Steiner, J. (2012). *The Foundations of Deliberative Democracy*. *Empirical Research and Normative Implications*, Cambridge (UK), Cambridge University Press.

Stirling, A. (1999). *On science and precaution in the management of technological risk.* SPRU, Report for ESTO, Brighton.

Thompson, D. F (2008). “Deliberative Democratic Theory and Empirical Political Science,” *Annual Review of Political Science*, 11, pp. 497-520.

von Schomberg R. (2011a). “The quest for the "right" impacts of science and technology. An outlook towards a framework for responsible research and innovation”. In M.Dusseldorp, R. Beecroft (eds.) "Technikfolgen abschätzen lehren. Bildungspotenziale transdisziplinärer Methoden*".* Springer Verlag, pp .394.

von Schomberg, R.( 2011b). “Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields”, European Commission, Brussels.

von Schomberg, R. (2013). "A vision of responsible innovation". *In* R. Owen, *et al*. (eds.) (2013b), pp. 51-74.

1. Del 2.2. and del 2.3. [↑](#footnote-ref-1)
2. The distinction is important. Claiming that participation and deliberation are now at the center of RRI approaches that not mean than practices have fundamentally evolved. Cf. forthcoming deliverables 3.2, 3.4, and 4.2. [↑](#footnote-ref-2)
3. See del 2.3. for more information on the parameters. [↑](#footnote-ref-3)
4. C.f. Del 2.2. for a presentation of the INDECT project. [↑](#footnote-ref-4)
5. To recall the title of Maesschalck and Lenoble’s 2003 book. [↑](#footnote-ref-5)
6. The use of the fact/value dichotomy here can seem hazardous as it traditionally opposes science to ethics and not two different types of scientific knowledge (rational knowledge of scientists and irrational beliefs of laypeople). Yet, the position of P.B. Joly is quite original because it considers the epistemic values created and held by scientists. In the standard model, these epistemic values would be opposed to the irrational set of knowledge of lay people in science. [↑](#footnote-ref-6)
7. There are actually two reasons explaining the width of public’s question: 1) the fact that some of them face real risks (opposed to the abstract theoretical thinking of experts, 2) the much broader conception of science held by the public, compared to that of scientists. [↑](#footnote-ref-7)
8. The co-construction of a common normative horizon appears then as a more realistic objective than the co-construction of scientific knowledge. [↑](#footnote-ref-8)
9. See del 2.2. for a presentation and an assessment of this project from RRI perspective. [↑](#footnote-ref-9)
10. That corresponds to an “ideal situation of speech” that Habermas then abandoned in favor of the pragmatic presuppositions. In Habermas (2005, p. 89), he identifies four of them as being the most important: 1) no one capable of making a relevant contribution has been excluded, 2) participants have equal voice, 3) they are internally free to speak their honest opinion without deception or self-deception, and 4) there are no sources of coercion built into the process and procedures of discourse. [↑](#footnote-ref-10)
11. To quote Article 2 of the Treaty of Lisbon (2007). [↑](#footnote-ref-11)
12. As we mentioned it, those gaps imply that participation is reduced to mere consultation and tends to play the same role than in the Standard Model where opposite voices are listen to in a first time but then are conceived as irrational or as resulting from ignorance, which has to be fought by adequate education and communication. [↑](#footnote-ref-12)
13. In chapter 4, we analyze the limitations of these two models. In the case of model 2, the main issue lies in the absence of a theoretical approach of the implementation of norms. Responsiveness, participation and deliberation can be promoted as enablers of responsible innovation but without specifying the concrete conditions by which they will be put into practice, they remain as wishful ideals without flesh. [↑](#footnote-ref-13)
14. See del 2.2. and del 2.3. for further analysis of second order reflexivity. [↑](#footnote-ref-14)
15. We will see in chapter 4 what the limits of this attempt are. [↑](#footnote-ref-15)
16. which is never defined and whose requisites are never emphasized. [↑](#footnote-ref-16)
17. Perhaps sometimes too narrowly conceived. [↑](#footnote-ref-17)
18. The term, first introduced by Bernard Williams and also discussed by Thomas Nagel, refers to circumstances whereby a moral agent is assigned moral blame or praise for an action or its consequences even though it is clear that this agent did not have full control over either the action or its consequences. [↑](#footnote-ref-18)