

Responsible Research and Innovation and the Governance of Human Enhancement

PREPRINT

Guido Gorgoni

Department of Political Science, Law, and International Studies,

University of Padova, via 8 febbraio, 2, 35122 Padova, Italy

e-mail: guido.gorgoni@unipd.it

Cite as: Gorgoni, G. Nanoethics (2018). <https://doi.org/10.1007/s11569-018-0326-9>

Read-Only version freely available at: <https://link.springer.com/epdf/10.1007/s11569-018-0326-9>

Abstract

This article aims to explore the debate on human enhancement (HE) from the perspective of the evolutions of responsibility paradigms, and in particular from the perspective of the so-called Responsible Research and Innovation (RRI) approach. The aim is not to explore the arguments pro or contra the ethical legitimacy and/or technical feasibility of human enhancement, but rather exploring if, and how, RRI perspective can shape the debate on human enhancement (and vice-versa).

In particular, the human enhancement debate will be read through the lenses of four main responsibility paradigms we sketch by examining both the historical and conceptual evolution of the responsibility idea, as well as the dynamics of its ascription. In order to provide a useful scheme for interpreting human enhancement, RRI will be characterised as a distinctive responsibility model that can subsequently be used to frame the debate on HE with a particular emphasis on its normative implications, as well as on its social and political significance.

1. Introduction

The quest for the responsible development of science and innovation has a long history, particularly within the technological field. In order to deal with the management of both the intended and unintended outcomes of innovation, many approaches have been

elaborated throughout the years. Methods and approaches such as the multiple variants of technology assessment and stakeholder engagement, the consideration of ethical, legal and social implications of research (ELSA), and 'midstream' modulation of science [1] are all attempts to translate the "responsibility" into the policy discourse on science and technology. The more recent notion of Responsible Research and Innovation (RRI) can be seen as one of these attempts. However, we believe that RRI offers a more ambitious framework, outlining a paradigmatic shift for the understanding of responsibility in science and technology governance.

Human Enhancement can be briefly characterised as "the intentional improvement of individuals' capacities with the help of technical or biomedical interventions in or on the human body" [2] (see also paragraph 3). Shelley-Egan and others methodologically distinguish between "enhancements" with a "small e" (those already existing) and a "capital E" (speculation about the future, mainly those envisioned by transhumanist discourse), claiming that "the debate over human Enhancement is less relevant for governance of current research and innovation pathways, as the technologies and applications discussed are of a speculative nature" [3]

Whilst this approach can be accepted as a methodology for narrowing the scope of the analysis, this separation cannot be fully maintained here, as human enhancement is intrinsically both normatively framed and ideologically charged. That is to say that the discourse on HE (as the mentioned authors do recognise) is not a purely scientific one, but is mainly a societal and political one.

We argue that RRI is uniquely placed to address this social and political salience of HE, insofar as it is an equally normatively and ideologically committed concept, which has the ambition of steering the development of research and innovation in a manner that would benefit society.

We should therefore (a) compare RRI with alternative responsibility paradigms (section 2); (b) examine how the governance of human enhancement is affected by those paradigms in order to underline what the potentialities of RRI might be (section 3); and finally, (c) explore how RRI is and can be influenced by the specificities of HE and its discourse (section 4).

We should therefore (1) locate RRI with respect to different responsibility paradigms, then (2) see how the governance of human enhancement is affected by those paradigms in order to stress what the potentialities of RRI might be; finally, (3) it is also worth considering how the debate on HE can shape the debate on RRI and the understanding of this notion.

2. Paradigms of responsibility

The historical evolution of the notion of responsibility, particularly in the legal field, has produced different paradigms. The different meanings of responsibility can be understood by organizing them around two distinct modes and two temporal dimensions.

Regarding the modes, we can distinguish between a passive and an active sense of responsibility. The passive pole is related to the idea of being held responsible by somebody else (morally or, typically, legally) and it corresponds to the idea of an imputation of responsibility, where responsibility is understood as the obligation to bear the consequences of an action (liability) or as the obligation to take into account one's duties and give an account of it (accountability). The active pole of the idea of responsibility is linked to the idea of a voluntary assumption of responsibility without relying on pre-established duties (unlike liability) nor waiting for *ex post* accounts (unlike accountability); this anticipatory and active meaning of responsibility is captured by the idea of *responsiveness*, which shapes responsibility more in terms of a quality or an attitude than of a process or its outcome [4].

A second crucial distinction for understanding the different meanings of responsibility includes its axis of temporal directions, which can be retrospective or prospective [5].

Retrospective responsibility is backward-looking, i.e. past-oriented, and is essentially linked to the idea of a *reaction*, which shapes the idea of responsibility in terms of sanction, compensation or justification, and is essentially linked to the ideas of liability and accountability.

Prospective responsibility is forward-looking, i.e. future-oriented, and is essentially linked to the idea of *assuming* and *actively exercising* responsibility, both in the sense of complying with some duties, but also by (pro)actively assuming responsibilities when the contents of our duties and tasks are not or cannot be established in advance. Taken in this sense, responsibility is called "prospective"[6] in that responsibility is not an *ex post* judgement over a certain state of affairs, but it concerns a possible future state of affairs.

Considering the two semantic poles we described above and the two temporal directions underlying the different understandings of responsibility, different general paradigms of responsibility can be distinguished according to their shifting logic and combination of these elements.

In particular, three main paradigms can be identified [7]:

(1) The paradigm of *fault*, corresponding to the traditional moral and legal idea of responsibility as linked to a faulty causation by the agent. This paradigm of responsibility is essentially *retrospective*, as it is based on an *ex post* judgment on a past action (and possibly on its sanction), and *passive* as somebody is made responsible by somebody else.

(2) The paradigm of *risk*, in which the legal focus of responsibility is geared toward

guaranteeing victims against damages (without reference to individual fault) rather than on sanctioning the “responsible” person(s), whose involvement in producing or not producing the damage becomes irrelevant under the “objective” logic of compensation. This model of responsibility is indeed *prospective* in that it aims to anticipate the occurrence of damages by means of risk management techniques, and *active*, in the sense that it favors risk anticipation through organisational measures and insurance.

(3) The paradigm of *precaution*, as a reaction to a situation of epistemic uncertainty that cannot be domesticated by means of risk prevention. As the direct or indirect outcomes of research and innovation practices cannot be fully anticipated (e.g. the effects of the use of chemical products in agriculture and their effects on the ecosystem, the effects of genetically modified organisms on the biosphere, etc.), it becomes crucial to anticipate also the undesirable outcomes of techno-scientific activities. This paradigm is both *prospective*, as it is oriented towards undetermined potentially negative effects which do not fall within the two former paradigms of responsibility, and *active* as it requires balancing both facts (data) and normative assumptions (values) case-by-case, without certain and established criteria for fully objectively assessing the (marginal but relevant) risks it addresses. This paradigm is linked to the idea of *precaution*, which emerged first within ethical reflection [8] and which has subsequently been enacted as a legal principle, initially in the form of soft law [9], and then as a general principle of European Union law [10].

Within those paradigms, different meanings of the word “responsibility” do coexist. Indeed responsibility has been defined as “a syndrome of concepts” [11] to the degree that nine different meanings of responsibility can be identified [12]. How is RRI located amongst these paradigms?

RRI admittedly aims to cope not only with *uncertainty*, but also with the *indeterminacy* of the outcomes of innovation. It therefore aims to include the precautionary paradigm, but also aims to improve it by relying mostly on the proactive and future-oriented senses of responsibility. Despite its deliberate soft forms of institutionalisation, RRI presents some distinctive features. Indeed, a common understanding of RRI emerges across the different definitions, which use different terminology and have distinguished orientations and focus, but share some common characteristics. These common characteristics can therefore be identified as defining the core features of the concept of RRI [13] [14]:

- a) Responsibility is oriented towards the future rather than towards the past: the specific approach of RRI aims to steer the innovation processes from within, according to societal values and needs, which are held in common.
- b) Responsibility is seen as a collective and participative process shared across different actors with different roles and powers along the innovation process: inclusion is therefore one of the most prominent dimensions of RRI.
- c) Responsibility is characterised as being proactive more than reactive, as it is

intended mainly as a driving factor of innovation processes rather than as a remedy for their failures: “Responsible research and innovation involve proactively seeking information about legal conduct as well as doing the right thing, whether there is a compliance mechanism or not” [15].

- d) The most challenging and problematic characteristic is the need for concrete engagement by the relevant actors in RRI with societal stakeholders.

These characteristics, derived from its formulation, seem to make RRI stand apart from the other responsibility paradigms, as it combines some of their elements in a new fashion. Indeed, RRI can perhaps be considered a new governance paradigm [13] that goes beyond the traditional emphasis on fault and punishment, risk and compensation, uncertainty and precaution, as it aims to *steer the innovation process from the inside* towards societal goals rather than coping with its (actual or anticipated) unwanted and unintended externalities (see Table 1). RRI aims to promote the integration of wide societal concerns in research and innovation processes, thus widening their traditional scopes and inscribing them into wider contexts. Moreover, RRI stipulates cooperation between societal actors, hence it is based on the responsabilisation of both innovators and stakeholders, promoting a governance model centered on the adoption and the practical implementation of (self-)regulatory instruments such as codes of conduct, guidelines, technical standards, reporting, and audits. In this way, RRI comprehensively combines and integrates various earlier approaches and methods, such as Technology Assessment or the Precautionary Principle. The position and the relationship between the different paradigms is illustrated in Table 1.

TABLE 1 HERE

These are the promises of RRI more than its reality. More than promises (or maybe we should say *like* promises), these are indeed the normative features of RRI – that is, the features required in order to realise RRI in practice. The distance between the rhetoric of all the responsibility discourses and their realisation can nevertheless be great. This distance makes its normative implications even stronger. This is to say that despite the many resources and tools for realising it already provided by many EC funded programs, which could lead to a successful codification and practical implementation of the paradigm, RRI is still in its infancy and the force of its declared ambitions has to be widened and reaffirmed if we really want to take seriously its promises and not accommodate the unavoidable rhetoric accompanying its practical implementations.

I suggest that a demanding and challenging subject such as HE has the twofold effect of stimulating an explicit assessment of RRI implications and, at the same time, of clarifying its claims and promises. In other words, both ideas assert many promises for the future: RRI could be, in principle, the governance model of HE, and indeed it also looks to be the one with greater potentialities, both for the perspectives it opens and for managing the state-of-the-art of the already ongoing HE interventions and technologies, as there are

already some practical implementations of HE which RRI could provide a governance framework for.

But, as we shall see, this relation could also be inverted: in being a projection over the future with definite and ambitious aims and means, HE somehow forces RRI to better define its aims and simultaneously to raise its expectations and demands.

The fourth element characterising RRI highlighted by Wickson and Forsberg [1], namely the willingness of the actors involved to both *act* and *adapt* according to the central features characterising RRI, shows the frailty of this idea and its subsequent vulnerability. Indeed, being that RRI is essentially dependent more on voluntary commitments and soft regulatory measures rather than on formal regulations makes it more vague and unsure as a normative standard. At the same time, this constitutes its appeal and its potential strength, and this could be a key element for its success, as it allows for some leeway in its implementation, which could be advantageous in allowing for its adoption by a larger audience. This uncertain regulatory status is linked to the problematic normative status of a soft form of regulation considered not only in itself but also in its relationships and interactions with formal (“hard”) regulation [16].

3. Responsibility paradigms and the governance of Human Enhancement

In this section, we discuss the impact that different responsibility paradigms can have on the governance of human enhancement techniques or, put another way, we explore the merits of each paradigm in the governance of HE. It will be argued that RRI could be a good candidate for the governance of HE, but only if some non-obvious conditions are met. Firstly it is necessary to clarify what HE refers to:

"Human enhancement, also called human augmentation, is an emerging field within medicine and bioengineering that aims to develop technologies and techniques for overcoming current limitations of human cognitive and physical abilities [...] human enhancement technologies [...] [are] techniques that improve human functions beyond a normal range" [17].

HE intervenes on fundamental human capacities, be they at the physical or cognitive level, which poses problems in terms of responsibility, necessitating an ethical and a legal framing. One of the main issues is the evaluation of the intrinsic merits and the limits of HE, which are far from being uncontroversial.

In turn, given that it affects basic human capacities, HE may alter agency and therefore is likely to induce some changes in both the distribution and the attribution of responsibility. By directly affecting the expectations and roles of the different societal

actors, this in turn redefines the terms of the implicit “social contract” between science and society [18].

Here we shall consider HE as an object of responsibility, while in the next we will consider responsibility as an object of HE.

In this sense, it seems clear that the paradigm of fault and that of risk are obvious candidates for regulating HE-related activities: the use of techniques with effects that are not fully known, or exposing “patients” to possible long-term damages do advocate for the use of liability and that of risk management, depending on the situation. Therefore, in principle, those regulatory techniques can be applied to HE just as they are to the medical profession.

Nevertheless, both techniques presuppose a well-established balance of values, rules and knowledge which is so far lacking in the HE domain, since:

- a) HE merits are contested according to different value options (strongly promoted by transhumanists but contested by others on the basis of a different scale of values);
- b) consequences and side-effects of HE techniques, especially in the long term, are not yet well known (for instance, if we think of the long-term effects of off-label prescription drugs);
- c) the boundaries of HE cannot be drawn in a clear manner between what can be labelled as enhancement and what cannot be labelled as such, so that the distinction becomes slippery.

These features, along with the implicit or explicit promises of a better future accompanying HE, do require a prospective governance approach focusing on societal values, rather than a reactive one. This perspective corresponds with forward-looking responsibility approaches, namely the Precautionary Principle and RRI (only if “taken seriously,” as we will argue).

We shall briefly discuss their merits and their limits as governance models for HE.

Despite being a future-oriented responsibility model, the Precautionary Principle is still conceived of and used as a remedy when the consequences of innovation may be in conflict with the wealth or even the survival of society. In my opinion, what distinguishes the Precautionary Principle from RRI is not its inner logic, which is that of the anticipation of responsibility and its extension also to the broad non-measurable consequences of innovation, nor its underpinned epistemology, which ultimately in both cases is linked to a context of epistemic uncertainty, but rather – and crucially – to its context of application.

Indeed, the Precautionary Principle has been conceived as an anticipatory remedy against the undesirable outcomes of innovation activities, either by inverting, diverting or blocking their path, thus remaining somehow in a context in which positions obey an

adversarial logic.

Instead, RRI aims to bring the anticipatory logic of precaution into the context of cooperation between innovators and society. What specifically distinguishes RRI from the Precautionary Principle is the integration of an anticipatory precautionary approach within the innovation process.

What distinguishes the two approaches, therefore, is not primarily their "functioning principle" (by which I mean their basic epistemology and their inner logic), but the quality of the context: the real novelty of RRI as a governance approach is not to be found in its epistemological stance, but rather in the attempt to shape a different context of actors involved, values considered, timing and place of decision, etc.

The alleged political nature of the Precautionary Principle determines its force, making it quite controversial as a legal principle. That feature is precisely what RRI takes and brings to another context, which can be characterised by the cooperation between science and society (to put it simply) and not by their confrontation.

This shift does not require the shaping of new principles. It requires the shaping of new power relations, i.e. changing the equilibrium between the forces of science and society we have so far [19]. This is more of a cultural and political change than a purely ethical/legal issue, and therefore is harder to achieve. This also explains the relevance of the fourth element stressed by Wickson and Forsberg [1] in summarizing the essential features of RRI :

1. A specific focus on addressing significant societal needs and challenges
2. A research and development process that actively engages and responds to a range of stakeholders
3. A concerted effort to anticipate potential problems, identify alternatives, and reflect on underlying values, and
4. A willingness from relevant actors to act and adapt according to 1–3.

It is precisely this last element that shows how RRI is linked essentially to voluntary measures and how it is shaped by the creation of a cultural context that includes law (and which is formed and conditioned also by law) but that cannot be identified in a set of rules and principles. The common defining features of such approaches are: (a) their non-mandatory nature (i.e. the fact that they do not generate legal obligations), but nevertheless (b) the fact that they contain commitments that go beyond existing legal requirements or regulatory standards.

The coupling between those two characteristics puts RRI in an ambitious and yet problematic position, which is that of aiming to pursue genuinely *political* goals (in the broad sense of the definition of the conditions for living together in society) through

voluntary measures. In other words, RRI implies: (a) the assumption of obligations without coercion and (b) outside the institutional spheres of political representation.

Whilst the non-coerced assumption of obligations has been widely discussed and the various voluntary instruments used for implementing some regulatory standards explored, the relationship between RRI and political representation deserves closer attention, as it requires examining its normative implications, be they explicit or implicit.

With regard to the former, several practical instruments have been in place for decades, for instance, under the label of Corporate Social Responsibility, with its methods of practical implementation that seem to be appropriate to implementing RRI in practice. Voluntary approaches include a broad spectrum of possible arrangements spanning from industry self-regulation to negotiated agreements between government and industry (co-regulation).

Besides those self-regulatory instruments, a wide typology of tools specifically aimed at implementing RRI in practice is available; many of these tools have been developed in the FP7 European Union-funded programs specifically dealing with RRI. Concerning in particular the issues of identifying and managing responsibility in RRI, we could mention the “Responsibility Navigator” [20] (to cite one belonging to the very first stream of projects dealing with the definition of RRI, its scope and its ambitions, and which was elaborated via a ‘co-construction method’), which aims to supporting participating actors in identifying, developing and implementing RRI measures in specific contexts. In order to do so, the Responsibility Navigator provides some entry points to the fundamental questions of responsibility, which by definition are contested, as they involve multiple actors with different (when not conflicting) points of view [21]. Following that first stream, other projects that have been funded aim to develop instruments and tools [22] to translate RRI into concrete practices in different fields such as industry [23] or higher education [24]. the merits or demerits of which still remain to be assessed.

With regard to RRI and political representation, it is necessary to examine the conditions for a non-unilateral setting of reference values, standards and practices, which points at another issue *à la page* nowadays: that of stakeholders' engagement in research and innovation processes.

Stakeholders' engagement is undoubtedly required for a successful allocation of responsibilities, but nevertheless it is not sufficient in itself to grant an acceptable balance between societal goals and economic impacts: RRI cannot be based solely on voluntary instruments and there is a need to blend regulatory flexibility with democratic safeguards [25].

The need for a “democratization” of research and innovation is a consequence of the changed relationship between science and society: as long as science is producing rather than reducing uncertainty, and as long as that affects some essential aspects of our living in societies, the scientific and the political debate become intertwined; the augmented

risks cannot be addressed only through expertise or technological applications, so that managing uncertainty is no longer a prerogative of public decision-makers, and the decision-making processes requires more transparency and public participation [26]. Indeed there is a serious power imbalance between the general public (stakeholders) on one side and innovators and policy-makers on the other side, which has to be taken into account when advocating for public participation and deliberation in policy-making and in setting the research agenda.

The rationale behind the inclusion of societal actors in science [27] is that of filling a double gap between scientific innovation and society, namely that concerning the production of knowledge, and that concerning the legitimacy of the choices that are made. Increasing the democratic nature of innovation process governance is not only a matter of facilitating representativeness within the participatory processes, but also of overcoming a reductive framing of societal issues in participation initiatives [28]. Indeed participatory approaches cannot grant a qualitatively better decision-making by themselves:

“Participation as such does not insure an automatic positive outcome of the process. The genuineness and efficacy of this idea depends basically on who participates (and how) and on the link between participation and the decision-making process” [29].

The issue, in sum, is not whether the public should have a say in technical decisions, but how to promote more meaningful interaction among policy-makers, scientific experts, corporate producers, and the public [27].

Therefore, if we want to give credit to RRI as a new governance approach, and in particular to its endeavour of attempting to add morally and politically relevant functionality to the innovation process [30], I think that RRI has to acquire a stronger *constitutional* identity [31].

We therefore have to maintain that RRI aims at creating an *agora* for the science-society debate, which constitutes its genuinely political mark: “the *agora* is a domain of primary knowledge production – through which people enter the research process” [32]. Along with these premises, it is possible to group the various definitions of RRI and to distinguish two fundamental ways of approaching and constructing this notion, each of them with definite features and normative implications. The literature therefore has to differentiate between a “socio-empirical” and a “normative” approach to RRI [33].

In the socio-empirical definitions of RRI, values and goals are the result of a contextual consensus, which does not provide a stable normative anchoring for RRI governance; within this framing, there is no remedy internal to RRI that can correct possible contrasts with fundamental rights, which appear to have a weak status.

This is the reason why more normatively oriented versions of RRI seem to be more convincing, especially in light of the declared aims and ambitions of the RRI idea. A

normative approach maintains the necessity of guaranteeing fundamental rights, which are not only part of the Rule of Law, but also belong widely to the legacy of the contemporary European philosophical, political and legal culture [34]. Therefore the high political ambitions of the RRI idea – if taken seriously, according to its definitions and to the benefits proclaimed by the European Commission (EC) – require reference to values belonging to the constitutional traditions of European countries, in particular those embedded in the Charter of Fundamental Rights of the European Union and in the European Convention on Human Rights.

The reference to fundamental rights protection should not be seen as a ‘threat of adjudication’ once removed from the ‘threat of regulation’, but rather as a reference to a common level of protection within the European Union.

It follows that voluntary (self-)regulatory initiatives should grant a level of protection of fundamental rights compatible with the current standards defined at the European level (both by the European Union and by the Council of Europe, as specified and integrated by the Court of Justice of the European Union and by the European Court of Human Rights). The reference made to human rights does not imply taking societal goals and values as fixed once and for all in advance and outside of society; indeed the content of fundamental rights is subject to contextual specification and actualization over time.

The inclusive dimension of RRI is called on precisely to structure the disagreements and conflicts between different (if not diverging) interpretations of fundamental rights and disputes about finding the correct balance between them.

Therefore, the integration of Human Rights within RRI is precisely what can characterise RRI as a distinctive governance approach to the management of the responsibilities of research and innovation – at least if we want to take it seriously and give credit to its ambitions. Otherwise the risk would be that “the undirected disposability of purposes can create the risk that RRI could be realised in ways that de facto contradict its premises, thus becoming worthless rhetoric or an instrument for covering purposes other than its authentic promises” [6].

Therefore, a major implication of adopting an RRI governance approach for HE is that of inscribing the questions of Enhancement in a wider societal perspective: "bringing an RRI lens to the context of use of prescription drugs for cognitive enhancement can contribute to raising the level of discussion from individual level to collective or societal level to consider both societal needs and expectations and implications for society" [3].

The adoption of an RRI perspective therefore demands us to evaluate the specific merits of HE in the societal context, in particular assessing if, and how, it can respond to societal needs and challenges. RRI can add a multifaceted perspective to its analysis, challenging HE underlying values and asserted benefits, and shifting the discourse from concrete enhancing techniques to the factors and values that both drive HE and are promoted by such discourse (e.g. competitiveness).

If we examine the features of RRI, it is worth noting how many of them (if not all) are intended to shape the context of the debate: a commitment to openness and transparency, ensuring pluralism through stakeholder engagement; promoting anticipation and reflection on actions and values; encouraging responsiveness, conceived as an attention to the various, and often conflicting, normative assumptions of societal actors. Significantly, the analysis of the merits of RRI for the governance of Human Cognitive Enhancement (HCE) [3] has pointed to the potential contribution of HCE in solving societal challenges, which precisely reflects the intrinsic political nature of RRI.

Whilst I do share these conclusions, I think it is necessary to examine the implicit assumptions behind them, namely the reasons why we should give RRI a chance and consider RRI capable of tackling such a demanding task.

In this sense, taking the HE debate seriously implies taking RRI seriously in turn – that is to say, specifying the normative requirements not only implied by the RRI model, but also those that it would be reasonable to expect from RRI, if it has to be given such a credit and a role. In other words, this implies (re)constructing some more precise normative implications of RRI that could be derived from its definitions and the roles it is called to fulfil, according to the rhetoric accompanying it.

As for our present discussion, the interaction of HE with the different responsibility paradigms will be the object of the subsequent section.

4. Human Enhancement influence on RRI

In this section, I will try to illustrate how human enhancement can affect responsibility, and in particular, the pivotal elements of each responsibility model. Indeed, it seems clear that HE can affect all the responsibility paradigms as it has the potential to significantly influence their inner logic and mechanisms, altering the way they work. Human enhancement has a direct impact on the established criteria of responsibility that we find, for instance, in law as – by definition – enhancement techniques do intervene on the basic elements defining human capacity, therefore affecting the process of both ascribing responsibilities *ex ante* and evaluating or judging responsibilities *ex post*. As normal bodily or mental functions are augmented, the question of the threshold used for ascribing and evaluating responsibility has to be questioned across the paradigms of fault, risk, and precaution.

fault

The paradigm of fault follows the idea that responsibility depends on capacity and causality [35]. In particular, responsibility is ascribed to a capable person if there is a causal link between their action and the occurrence of an event (there is never ascription of responsibility without capacity, whilst the ascription of liability without causality is possible as an exception). Within this framing, it is clear that HE interferes with the very basic conditions of responsibility, and in particular as the augmentation of the bodily or mental function claims to also be an enhancement of capacity, the question arises whether this should justify increasing or decreasing the responsibility threshold. Therefore, both bodily and cognitive enhancement seem to directly affect responsibility, namely regarding the ascription of duties and the imputation of liabilities. Indeed, as “responsibility follows capacity,” a possible “natural” consequence would be that of expecting more from people with higher capacities, therefore ascribing to them higher responsibilities than to “ordinary” people. This is not the right place to discuss the issue extensively, but it is worth highlighting some problems it presents: whether discriminating “enhanced” persons would be justified and fair; who should determine the higher threshold, and how; and what kind of responsibilities we might expect from enhanced individuals.

The problem is therefore whether it would be justified to impose new (and more demanding) duties on “enhanced” people, in particular when cognitive enhancement interventions are deliberately taken in the context of highly specialized professional practices (i.e. surgery) for the purpose of minimizing the risk of bad outcomes or increasing the success rate of their intervention [36].

Indeed it seems that the creation of new specific duties is currently thought of as being problematic more for reasons linked to the reliability of the enhancement techniques (e.g their long-term side effects) than to the legitimacy of the expectations. At the opposite end, it appears that some expectations about the duty to enhance for some professionals in some circumstances are likely to arise in a near future [37], as the standard responsibility threshold might be affected by the evolution of capacities linked to HE techniques. Without engaging in an extensive debate over the issue here, it is worthy observing that RRI could have effects precisely on the definition of the legitimacy of those expectations: while acknowledging that capacities are defined only by the measure of a bodily or mental function (which is an issue to be discussed, as it is not self-evident), it remains that normative expectations are, or should be, also societally-defined. Therefore, the contribution of an RRI approach in this sense would be that of questioning, in the sense of opening the debate on a public level regarding the implicit or explicit normative expectations about HE and enhanced individuals.

risk

Enhancement techniques could have effects on risk-management techniques, both in the sense of decreasing some risks by virtue of the enhanced capacities, and of increasing other risks for the same reason. More specifically, for example, cognitive enhancement techniques on some professionals (e.g. in the medical profession) – that can augment some capacities like concentration or resistance to stress – could help lower the risk of mistakes, thus decreasing the risk of compensation. On the other hand, as the long-term effects of some enhancement techniques are not well-known yet, the risk to health of the enhanced individuals could be increased.

As risk can only be distributed but never eliminated, the reasons behind the re-distributions of damages between the individuals (a damage not compensated) and the society (compensation of damages suffered by individuals on the basis of an insurance system) must to be subject to public scrutiny, as they shape an “equilibrium between social forces” in society, which in turn depends on some fundamental societal values (as we will discuss later, this is one of the reasons that could count in favour of an RRI-based approach to HE). A risk-based approach to responsibility could lead to justifying both the imposition of special enhancement duties on certain professionals in order to increase the benefits for society.

precaution

Unlike the risk-based approach, the precautionary approach does not require the certainty of a causal link between the action or the event and the possible harm, as it intervenes on non-measurable risks and on uncertain causal links: “according to the precautionary principle, lack of scientific certainty about the potential harm should no longer constitute an obstacle for risk-preventive actions” [38].

It seems to me that HE can interfere with the precautionary approach, as it could alter the evaluation of the unforeseeable consequences that the principle deals with, as well as the horizon of solutions available for damage recovery. The proactive, interventionist approach of HE can influence risk assessment and alter the evaluation of uncertainty. Unpredictability can be seen as an opportunity, unforeseen events as occasions to exploit by using HE techniques [39]. In this way, the precautionary approach could be undermined by de-potentiating both its logical foundations and its intended effects.

RRI

The RRI discourse is imbued with a specific rhetoric of responsibility that has many aspects: the commitment towards societal goals, openness and participation, and reflexivity and responsiveness, to cite the more relevant ones. The relationship between HE and RRI seems to be complex and ambiguous. It is true that the RRI paradigm is

based, albeit non-exclusively, on voluntary measures going beyond duties that are mandated by law, as we have noted before. Nevertheless, it is necessary to consider that RRI is not an individual form of responsibility, but a collective one, and the core of RRI is the creation of a context of participation and discussion about societal values [25]. In this sense, the capacities to be enhanced are considered from the viewpoint of their ethical, legal and political significance, rather than from the perspective of the specific bodily and cognitive alterations they entail.

It appears therefore that if HE has an impact on RRI, it is not in the sense of *altering* the functioning of its constitutive elements, but rather that of pushing RRI to be more explicit about its intrinsic normative commitments at the ethical, political and legal level.

Without revoking the voluntary nature of the rules and standards through which RRI is mainly realised (as it is realised essentially through soft law and self-regulatory tools), nevertheless we can maintain that RRI is a fundamentally *normative* concept, which poses the need for a *normative steering* of research and innovation oriented towards the achievement of fundamental societal goals, revealing its underlying *political* nature.

In the EU context, stressing the political nature of RRI means giving it not only scope but also content in line with European political values, which can be summarised by the Rule of Law and the protection and promotion of Human Rights.

This political significance of RRI does require reference to some common values defining some fundamental features of society and polity. These values can be found in the principle of the Rule of Law, which is part of the legal-political heritage of many European countries, and in the protection of Human rights, in particular those embedded in the Charter of Fundamental Rights of the European Union and in the European Convention on Human Rights.

This is a crucial element, as it provides a solid normative framing for RRI without silencing the controversies around values, which are at the core of RRI as a governance approach: as there is not a stable hierarchy between them, and it is necessarily a contextual balancing in each case, the conditions for their compression cannot be left purely to a contingent consensus between innovators and stakeholders.

Therefore the reference to fundamental rights is necessary in order to frame the context of the values implied by RRI, and the subsequent reference to the Charter of Fundamental Rights of the European Union is justified both as it has been produced in the context of the European Union and because the content of the rights have been updated according to changes in society, as well as social progress and scientific and technological developments (particularly recognising data protection as an autonomous fundamental right beyond privacy), which seems to fit with the RRI idea.

The reference to fundamental rights protection therefore should be seen as reference to a common level of protection of fundamental values within the European Union. The

content those values assume in different contexts should certainly be specified through participatory processes, but also through the case-law of the European Court of Justice and that of the European Court of Human Rights, which could define the standards of their protection and the limits and conditions of their limitation.

The role played by horizontal dynamics between fundamental rights in this sense appears crucial. In this sense, the inclusive deliberation processes advocated by RRI are called upon to structure the conflict of interpretations between fundamental rights, the content of which is subject not only to contextual specifications, but also has to be actualized over time: the integration of fundamental rights into research and innovation could also help reduce the risk of adverse technology reception by society, effectively playing the steering function expected from a governance model.

5. Final thoughts

This paper argues that RRI seems better suited than other governance models to address the issues raised by HE. To do so, RRI should, however, be taken “seriously”. We believe this means that, in the EU context, RRI should be explicitly anchored in human rights and their legal instruments, according to what I called a “constitutional identity” of RRI. Indeed, a normatively-framed version of RRI seems to grant flexibility on one side, but within a framework of rights and duties. Actually, the integration of Human Rights within governance mechanisms and instruments can be seen precisely as the element characterising RRI as a distinctive governance approach compared to other forms of governance, if we want to give credit to its ambitions [6].

From this viewpoint, we can certainly maintain that inclusion, and in particular stakeholder participation, is undoubtedly an essential feature of RRI, but inclusion cannot and should not overcome the reference to fundamental rights, in particular those more closely associated with a collective dimension.

This approach is relevant to HE, as it seems that it is not possible to separate enhancement from its framing in terms of values, be they explicit or implicit. Like RRI, HE is inherently political in nature. Taking for granted its premises and promises is a political stance in itself. Therefore, given the nature and the width of its actual or potential impacts on society, its merits and limits must be subject both to the possibility of public scrutiny and to subsequent approval or refusal.

RRI has the potential, if taken seriously as an explicitly normative governance mechanism, to steer the debate on the fundamental, normative assumptions of HE. In turn, the intrinsically political and ideological nature of HE makes the apparently idealistic ambitions and features of RRI as essential elements of the responsible governance model which HE requires.

References

1. Wickson F, Forsberg E-M (2015) Standardising Responsibility? The Significance of Interstitial Spaces. *Science and Engineering Ethics* 21:1159–1180. doi: 10.1007/s11948-014-9602-4
2. humanenhancement | Enhancement & RRI. In: humanenhancement. <https://www.responsibleenhancement.eu/enhancement-rri>. Accessed 3 Jul 2018
3. Shelley-Egan C, Hanssen AB, Landeweerd L, Hofmann B (2018) Responsible Research and Innovation in the context of human cognitive enhancement: some essential features. *Journal of Responsible Innovation* 5:65–85. doi: 10.1080/23299460.2017.1319034
4. Pellizzoni L (2004) Responsibility and Environmental Governance. *Environmental Politics* 13:541–565. doi: 10.1080/0964401042000229034
5. Cane P (2002) *Responsibility in Law and Morality*. Hart Publishing, Portland, OR
6. Arnaldi S, Gorgoni G (2016) Turning the tide or surfing the wave? Responsible Research and Innovation, fundamental rights and neoliberal virtues. *Life Sciences, Society and Policy* 12:1–19 . doi: 10.1186/s40504-016-0038-2
7. Ewald F (1993) Responsabilité. In: Arnaud A- J (ed) *Dictionnaire encyclopédique de théorie et de sociologie du droit*. Lgdj, Paris
8. Jonas H (1984) *The Imperative of Responsibility*. University of Chicago Press, Chicago, IL
9. United Nations Conference on Environment and Development (1993) *Rio Declaration on Environment and Development*. United Nations, New York
10. (2016) *Treaty on European Union*
11. Vincent NA (2011) A Structured Taxonomy of Responsibility Concepts. In: Vincent NA, Poel I van de, Hoven J van den (eds) *Moral Responsibility*. Springer Netherlands, pp 15–35
12. van de Poel I (2011) The Relation Between Forward-Looking and Backward-Looking Responsibility. In: *Moral Responsibility. Beyond free will and determinism*. Springer, pp 37–52
13. Arnaldi S, Gorgoni G, Pariotti E (2016) Responsible Research and Innovation as a governance paradigm: What is new? In: *Navigating Towards Shared Responsibility in Research and Innovation. Approach, Process and Results of the Res-AGorA Project*. Karlsruhe, pp 23–29
14. Burget M, Bardone E, Pedaste M (2017) Definitions and Conceptual Dimensions of Responsible Research and Innovation: A Literature Review. *Science and Engineering*

Ethics 23:1–19 . doi: 10.1007/s11948-016-9782-1

15. Iatridis K, Schroeder D (2016) *Responsible Research and Innovation in Industry*. Springer International Publishing, Cham
16. Shaffer GC, Pollack MA (2009) Hard vs. soft law: Alternatives, complements, and antagonists in international governance. *Minn L Rev* 94:706
17. Brey P (2009) Human enhancement and personal identity. In: *New waves in philosophy of technology*. Springer, pp 169–185
18. Arnaldi S, Bianchi L (2016) *Responsibility in Science and Technology. Elements of a Social Theory*. Springer Netherlands
19. Wright D, Gutwirth S, Gellert R, Friedewald M (2011) Precaution and privacy impact assessment as modes towards risk governance. In: Schomberg R von (ed) *Towards responsible research and innovation in the information and communication technologies and security technologies fields*. Publ. Office of the European Union, Luxembourg, pp 83–97
20. Res-AGorA Project (2016) *Responsibility Navigator*. <http://responsibility-navigator.eu/>. Accessed 25 Oct 2017
21. European Commission *Responsibility Navigator* [press release]. http://cordis.europa.eu/news/rcn/124643_en.html. Accessed 26 Oct 2017
22. RRI-Tools. <https://www.rri-tools.eu>. Accessed 26 Oct 2017
23. *Responsible-Industry*. <http://www.responsible-industry.eu/home>. Accessed 26 Oct 2017
24. *Heirri - Integrating RRI into Higher Education Institutions*. In: *Heirri - Higher Education Institutions & Responsible Research and Innovation*. <http://heirri.eu/>. Accessed 26 Oct 2017
25. Gorgoni G (2018) Voluntary measures, participation and fundamental rights in the governance of research and innovation. *ORBIT Journal - An Online Journal for Responsible Research and Innovation in ICT* 1:1–21. doi: 10.29297/orbit.v1i4.72
26. Boisson de Chazournes L (2009) New Technologies, the Precautionary Principle, and Public Participation. In: Murphy T (ed) *New Technologies and Human Rights*. Oxford University Press, Oxford, New York, pp 161–194
27. Jasanoff S (2003) Technologies of humility: Citizen participation in governing science. *Minerva: A Review of Science, Learning & Policy* 41:223–244
28. Wynne B (2007) Public Participation in Science and Technology: Performing and Obscuring a Political–conceptual Category Mistake. *East Asian Science, Technology and Society* 1:99–110. doi: 10.1215/s12280-007-9004-7
29. Gianni R (2016) *Responsibility and Freedom: The Ethical Realm of RRI*. John Wiley & Sons

30. Van den Hoven J (2013) Value sensitive design and responsible innovation. *Responsible innovation: Managing the responsible emergence of science and innovation in society* 75–83
31. von Schomberg R (2013) A Vision of Responsible Research and Innovation. In: Owen R, Bessant J, Heintz, Maggy (eds) *Responsible Innovation*. John Wiley & Sons, Ltd, pp 51–74
32. Nowotny H, Scott P, Gibbons M (2003) *Introduction: Mode 2'Revisited: The New Production of Knowledge*. Springer
33. Ruggiu D (2015) Anchoring European Governance: Two Versions of Responsible Research and Innovation and EU Fundamental Rights as 'Normative Anchor Points.' *NanoEthics* 9:217–235 . doi: 10.1007/s11569-015-0240-3
34. Yeung K (2004) *Securing compliance: a principled approach*. Hart Pub, Oxford; Portland, Or
35. Vincent NA (2013) *Neuroscience and Legal Responsibility*. Oxford University Press
36. Maslen H, Santoni de Sio F, Faber N (2015) With Cognitive Enhancement Comes Great Responsibility? In: Koops B-J, Oosterlaken I, Romijn H, et al (eds) *Responsible Innovation 2*. Springer International Publishing, Cham, pp 121–138
37. Santoni de Sio F, Faulmüller N, Vincent NA (2014) How cognitive enhancement can change our duties. *Front Syst Neurosci* 8:131. doi: 10.3389/fnsys.2014.00131
38. Osimani B (2013) An epistemic analysis of the precautionary principle. *Dilemata: International Journal of Applied Ethics* 149–167
39. Pellizzoni L (2012) Strong Will in a Messy World. *Ethics and the Government of Technoscience*. *Nanoethics* 6:257–272. doi: 10.1007/s11569-012-0159-x

Tab. 1 - RRI and other responsibility paradigms

Paradigm	Time orientation	Mode	Criterion	Mean	Target	Regulatory mechanism
fault	retrospective	passive (subjection)	liability	sanction	negative outcomes	hard law
risk	prospective	active (anticipative)	risk-assessment	compensation	negative outcomes	hard law
precaution	prospective	active (preventive)	deliberation	expertise	negative outcomes	hard law, soft law
RRI	prospective	active (proactive)	responsiveness	participation	negative and positive outcomes	self-regulation, hard law, soft law