Retooling Techno-Moral Scenarios. A Revisited Technique for Exploring Alternative Regimes of Responsibility for Human Enhancement

Simone Arnaldi

Abstract The techno-moral scenarios (TMS) approach has been developed to explore the interplay between technology, society and morality. Focused on new and emerging sciences and technologies, techno-moral scenarios can be used to inform and enhance public deliberation on the desirability of socio-technical trajectories. The article presents an attempt to hybridise this scenario tool, complementing the focus on ethics with an explicit acknowledgement of the multiple meanings of responsibility and of the plurality of its regimes, i.e. the institutional arrangements presiding over the assumption and assignment of responsibilities. We call this integrated technique 'rTMS' to stress the continuity with the original technique and, at the same time, to highlight the additional element we aim to develop: responsibility. The article describes this approach and illustrates a loosely standardised procedure that can be used to organise and conduct public engagement workshops based on rTMS.

Keywords Human enhancement · Techno-moral scenarios · Responsibility · Ethics · Public engagement

Introduction

'Human enhancement' (HE) has been labelled as the intentional effort to improve individuals' performance with the help of technical or biomedical interventions [1]. Far-fetched visions of HE may appear to be close to science fiction, with stories about mind uploading, physical immortality, and mergers with artificial superintelligences. However, exoskeletons, implants, and genetic engineering are certainly contemporary technological possibilities, which have already (or may have in the near future) enhancing applications in sport, the labour market, education and the military.

Beyond the practical outcomes of the current or hypothetical applications, the logic of enhancement infiltrates the representation we have of ourselves, our values and aspirations. As long as enhancement becomes a 'synonym' of fulfilment, the exploration of human enhancement technologies cannot neglect that their trajectories are inextricably linked to morals and social practices. It is not by chance, therefore, that HE has generated a heated debate, which has often focused on the social desirability and the moral legitimacy of the scientific and technological means to achieve enhancements, or of enhancement itself. In this regard, critics have seen HE as a violation of human nature ([2, 3]), as a threat to equity [4] or safety [5], ultimately characterising enhancement as a force unleashed to disrupt society and polity [6]. On the opposite side, the advocates of HE contend that enhancing individuals can give them greater freedom to shape the life they want to live [7], increase their welfare [8], and, ultimately, benefit society as a whole [9].

Despite its vivacity, this conflict over the characterisation of HE has led to a stalemate, with little dialogue between opposing camps. We could say that enhancement enthusiasts sometimes fail to recognise that these

S. Arnaldi (✉)
Department of Political and Social Sciences, University of Trieste, Piazzale Europa 1, I-34127 Trieste, Italy
e-mail: sarnaldi@units.it

technologies will have to be socially accepted and democratically legitimated if they are to be brought to market or even made available through public health systems, for instance, in Europe. The critics of enhancement neglect the fact that enhancement technologies are already in widespread use; depending on how enhancement is defined, such technologies have been commonplace for centuries [10].

This article revisits the techno-moral scenarios (TMS) technique as a potential contribution to overcome this stalemate. The TMS approach was developed to explore the interplay between technology, society and morality. Focusing on new and emerging sciences and technologies (NEST), TMS use narratives about hypothetical future situations to explore and reflect upon the possible moral consequences of technological change, thus informing public deliberation on the desirability of technological options.[1] The approach that this technique has to morality is descriptive, and its principal characteristic is the recognition that morality changes over time and that technologies are an engine of this change. While it is certainly true that technological trajectories are also oriented by morality and society, it is also true that technologies regularly lead to moral and social change. "We constantly see NEST uprooting established moral routines. These disturbances manifest themselves as controversies about how to re-establish a 'fit' between NEST, our moral world and us. […] In this way, NEST can lead to moral change, although it never determines whether such a change will occur or the direction of that change" ([12], 120).

The closure of controversies and the (re-)establishment of this 'fit' marks the emergence of new "moral regimes" [11]. Moral regimes are constituted by institutionalised practices, rules, and procedures in which abstract moral principles are embedded. These regimes define the positions and tasks, duties and obligations, rewards and sanctions, of and for social actors vis-a-vis their participation in the opening, unfolding and closing of ethical controversies about new technologies. In our view, "moral regimes" essentially describe 'responsibility regimes', insofar as they refer to the institutional arrangements presiding over the assumption and assignment of responsibility for the moral consequences of technological transformations. This view of moral regimes seems confirmed by the TMS literature, which considers moral regimes as a particular form of an "accountability regime", that is, the institutionalised practices, rules, and procedures for which the actors involved in science and innovation are held accountable by their various (and changing) stakeholders in society [13]. The notions of moral and accountability regimes suggest that the alternative resolutions of ethical controversies assume the form of specific institutional responsibility arrangements.[2]

The coupling between regimes and the closure of ethical controversies depends on the fact that responsibility arrangements are not morally neutral. For instance, some of these arrangements can define responsibility from a consequentialist perspective (what makes someone responsible is the 'quality' of the consequences of their actions), while others can take a different stance based on virtues (what makes someone responsible is their moral character). A consequentialist approach is likely to stress the dimension of accountability, whereas a virtue-based approach may emphasise individual responsibilisation. In HE, these alternative options can be translated in governance and regulatory approaches that, for instance, are based, respectively, on controls and sanctions, or on education and the self-regulation of the social actors involved. This moral relevance makes responsibility regimes participate in the ethical controversies about new technologies. For instance, to follow up on the previous example, an approach based on controls and sanctions can be considered a threat to personal autonomy, whereas an approach that is based on self-regulation can be perceived as ineffective in enforcing rule compliance or in

[1] Boenink, Swierstra and Stemerding define morality as "the set of values and norms that a specific community considers very important, because they refer to legitimate interests, mutual obligations and/or views of the good life. Although the precise boundaries of this set may be contested, morality largely exists in the form of implicit beliefs, routines and practices". Morality is different from ethics, which is "the reflection and debate on the relevance and status of (parts of) morality; ethics, that is, is reflexive morality. […] [A]nyone questioning or debating moral values and /or norms engages in ethical activity" ([11], 3).

[2] There is a second, narrower connection between responsibility and the TMS literature. Stemerding [14] and Douglas and Stemerding [15] use TMS in the framework of responsible research and innovation (RRI). RRI is an approach to science and innovation governance developed mainly in the European Union (EU) policy environment, which pursues the alignment of research and innovation activities with societal goals and needs by way of participatory approaches fostering the mutual responsibilisation of science and innovation actors ([16, 17]). In this context, TMS are used to foster the moral imagination of researchers as a preliminary step to begin a dialogue with societal stakeholders to explore the social and needs priorities that can be responded to by research.

preventing undesired consequences. In other words, the different definitions and regimes of responsibility may contradict the moral principles and practices of one or more social groups, or they can be considered insufficient to effectively translate specific value orientations and moral beliefs into policies and actions.

A focus on responsibility, responsibility regimes, and their participation in techno-moral change, is the original feature of this article and its contribution to the development of TMS. The goal is not to correct TMS, but to complement this technique by addressing this specific aspect, which was present and yet, in our view, undertheorised in the original technique. For the sake of clarity, we will refer to the expanded procedure we propose as 'rTMS' to stress the continuity with the original technique and, at the same time, to underscore the additional element we aim to develop: responsibility. Like their 'parent technique', the characteristics of rTMS make them suitable to answer to questions that address the nature and extent of the moral consequences of new and emerging technologies, and the recurring patterns of the public debate on these consequences and their causes. In addition, they pay closer attention to the interdependencies, alignments, contradictions, and conflicts between particular resolutions of ethical controversies about new technologies and specific responsibility regimes. Like the original techno-moral scenarios, rTMS address these issues from a descriptive point of view. The point is not to define a set of specific ethical standards to assess the ethical acceptability of HE technologies or of substantive criteria and procedural conditions according to which the HE interventions can be considered 'responsible'. Like morality, responsibility evolves and responsibility regimes are contingent on the situation observed, including for aspects such as the criteria that qualify somebody as 'responsible', the actors who assume or are assigned responsibility for something, and the means used to enact responsibility in social processes and configurations. Following this line of thinking, rTMS recognise the presence of multiple moral standards that are unevenly distributed in society and that can change, and do change, over time. They appreciate that moral change involves controversy when established standards and routines are challenged and that the resolution of these conflicts entails the association of alternative closures with specific responsibility arrangements.

With their descriptive approach, rTMS provide an alternative point of view on HE, which has the potential to overcome the stalemate that often characterises the public and academic debate. It does so by inviting a focus not on the outright normative assessment of HE technologies and by instead anticipating the possible combinations and associations of morals, technologies and responsibility. rTMS do not appraise different moral standards, nor they assert predefined definitions of responsibility. They are instead a powerful tool to examine the compatibility, consistency, and conditionality of alternative technological options, moral standards and principles, and responsibility arrangements. In doing so, they can offer suggestions on the implications of different technologies for moral and responsibility regimes, the legitimacy of different responsibility arrangements, and the limits and prospects for the mutual adaptations of technologies and morals. In our view, the improved understanding of the relations between morality, responsibility, and technologies that we can gain from applying rTMS, can foster our capacity to outline the possibilities and conditions for building robust moral and responsibility regimes for HE, anticipating our future challenges and fostering the imagination of possible solutions.

To illustrate the features and functioning of rTMS, the main part of the article begins with a presentation of the specific characteristics of techno-moral scenarios. Then, the concept of responsibility is discussed and alternative responsibility paradigms are described. Finally, these elements are combined, and the loosely standardised procedure to run an rTMS exercise is described.

Techno-Moral Scenarios

The scenario technique is aimed to capture future dynamics by elaborating alternative and hypothetical accounts of future situations as well as the decisions and events that allow one to move forward from the present to such futures. They are not forecasts of the most likely future; they do not identify the most likely pathway to the future and estimate uncertainties. Rather, they describe different, internally consistent, "worlds, not just different outcomes in the same world" ([18], 147), balancing plausibility, consistency and multiplicity [19]. Unlike single-line projections and forecasts, they aim at conceptually "reframing" [18] our views of the future, challenging entrenched assumptions, re-perceiving the environment and discovering new strategic opportunities and risks.

This technique is not new to the field of science and technology, and countless scenario studies have been produced. This article focuses on the specific application TMS. This technique has been designed to explicitly consider the normative aspects of the interactions of science, technology and society. Normativity is not new to scenario building. Normative scenarios have long been developed to outline preferable futures and assess the path that leads to such desirable future situations. They have "explicitly normative starting points, and the focus of interest is on certain future situations or objectives and how these could be realised" ([20], 728). Normativity in scenarios goes hand-in-hand with an inclusive approach to involving social and organisational actors in the scenario-building process. In normative scenarios, the word 'expert' has a broad meaning. Participants in scenario-building can be consulted either for their scientific or technical knowledge on the matter that is examined or because they have stakes in the decision-making process [21]. The role of visions in the identification of a desirable future, the importance of values in the analysis and building of future reality, and the social responsibility of the futurist are emphasised by the authors and studies that adopt such a normative stance (for example, Barbieri Masini and Medina Vasquez: [22], 51). Traditionally, the scenario literature views the normative approach as the opposite of an exploratory one. The former sets the objectives to achieve in a future situation and defines the conditions, actions and motivations leading to such achievement. The latter projects expected futures based on what we know of the past and of our present [23].

TMS blur this sharp distinction: they *explore* normativity and make morality the subject of the scenario exploration, examining the interplay between science, technology and morals. As Swierstra, Stemerding and Boenink explain, TMS address the "moral consequences", not the "*morally relevant*" consequences, of technologies. "The latter do receive attention insofar many scenario exercises are im- or explicitly *normative*, aiming at some common good. When [TMS] focus on the moral consequences, by contradistinction, [they] are not primarily interested in applying moral standards to [new technologies], but in the opposite, descriptive, question: how might the [new technologies] affect current moral standards and practices? Of course, these moral consequences will be of interest from a normative, ethical point of view too, but it is important to distinguish describing possible moral change from evaluating it" ([12], 121). An important element in the analytical approach of TMS is therefore the focus on technology and on how it leads to moral change. Through the exploration of "patterns or chains of [ethical] arguments: reactions and counter-reactions" ([11], 12), TMS examine controversies and conflicts and their unfolding over time. This vantage point of observation provides the opportunity to scan long-term moral changes.

This dynamic perspective differentiates TMS from other approaches aimed at assessing the moral desirability of technologies such as 'ethical technology assessment' (eTA) [24]. eTA which appears to mostly focus on the assessment of a new technology according to a set of categories and dimensions that are given analytical and moral relevance either by the researchers or by the social actors participating in the assessment. Therefore, "eTA seems to work with a rather limited timeframe, taking small steps at a time. Since morality usually evolves on the long term, this method may have difficulty anticipating such long-term changes" ([11], 5). A second, major difference between TMS and eTA involves the acknowledgement that technology itself affects and changes moral values. While the eTA approach also acknowledges this influence, technology's impact on morality appears as a residual aspect, or at least as only one element among many others.

TMS differ from other scenario approaches because of this more sophisticated stance on morality and ethics. Individual scenarios in normatively oriented scenario exercises tend to be morally uncontroversial: they incorporate a definite moral stance, or definite moral stances, and actions and behaviours are shaped accordingly. For instance, two different individual scenarios on sustainability can feature, respectively, two distinct situations in which households can behave either in an environmentally conscious or in a careless, eco-destructive way. Similarly, a scenario on cognitive enhancement may feature two different social actors, whose behaviour is oriented by their opposite views of, say, pharmaceutical enhancement. In individual scenarios, these positions rarely change and we can learn about possible moral controversies more from the combination and comparison of different scenarios, rather than from individual scenarios themselves ([11], 6–7).

The importance of focusing on the moral consequences of technologies, and the rationale behind TMS, is the increasing importance and scope of the "soft impacts" of new technologies, i.e., their influence

on culture, norms, beliefs, morals, and social relations. These peculiar effects of technologies challenge the traditional "regime of accountability" for technologies and their management. In this traditional regime, quantifiability, clear and noncontroversial harm and direct causation are the specifications required for impacts to be considered "sufficiently 'hard' (in the sense of objective, rational, public, and concrete) by technology and policy actors to accept and allocate accountability" ([13], 7). However, the more pervasive technologies are and the more "intimate" relations we have with them [25], the less effective these criteria will be in determining their effects. There are plenty of examples of these technologies, from ubiquitous computing, sensors, implants and prostheses, to mundane and omnipresent smartphones. These "intimate technologies" display effects that are irremediably soft: qualitative, ambiguous, and/or indeterminate, whose very identification is a matter of subjective definition and co-produced by users themselves. The premise of this difference is that "in the case of hard impacts the core normative question, that is: whether the technological impact itself is indeed (un)desirable, is considered to be answered already. As a result of this basic normative consensus, hard concerns come with the expectation that they can be solved on the basis of empirical facts" ([13], 9). Instead, the definition of soft impacts is enmeshed with subjectivity and based on disagreements about values rather than uncertainty and ignorance about facts. They pose normative challenges rather than factual ones ([13], 10) and the logic of their management is based on trade-offs rather than on containment, remediation strategies, or the deployment of safety technologies [26]. The moral ambiguity of these intimate technologies "goes deeper as it is caused by the destabilisation of the normative and moral routines that we rely on to assess the (un)desirability of the impacts of those technologies" ([13], 11). On this distinction between hard and soft impacts and the subjective nature of the latter, HE is telling. While there is an unproblematic consensus that HE technologies should do no physical harm to those who choose to enhance themselves, the purpose of enhancements, i.e., why they should be used, is subject to scrutiny and controversy.[3]

One of the critical aspects of reflecting upon future technological developments and their impacts is the risk of engaging in "speculative ethics" [27], i.e., mistakenly addressing hypothetical developments and far-fetched visions of often implausible futures as if they were actual situations [28]. To reduce this risk, techno-moral scenarios invite constant reference to past ethical controversies and to established tropes of ethical argumentation to allow the anticipation of future ethical debates in a sensible and controlled way. "To keep the imagination of the storywriter in check, [...] it is advisable to start from existing historical and sociological knowledge of past interactions between technology on the one hand and society and/or morality on the other. This can be done on several levels. First, we can start from *patterns in the unintended soft impacts* of emerging technologies. Second, we can use the *recurring tropes and types of arguments in ethical debates on* new technologies. In addition, finally it is possible to distinguish *different rates of change on different levels of society or morality*" ([29], 58).

Incorporated in brief fictional accounts of future situations, techno-moral scenarios therefore describe the dynamic processes leading to specific moral and technological futures and can be used to elicit the responses of stakeholders with regard to their desirability, thus offering "cues for reflecting on the possibilities to intervene in or prevent a certain course of events" ([29], 57). By focusing on individual protagonists responding to specific situations (see our discussion of 'personas' in the Section titled "Expanding Techno-Moral Scenarios"), scenario descriptions enable the picturing of these futures in a "non-abstract and non-generalizing form[, helping] imagine what a novel technology's impacts may look like on a very concrete and mundane level. How will it affect the lives of individual people or specific groups in a variety of situations? [...] This has two advantages. First, it enriches the imagination of what the future might look like by adding details. In addition, it invites the audience to identify with one or more of the protagonists. This makes the future more vividly present, and it may even evoke new perceptions of the existing world" ([29], 57).

Responsibility Paradigms[4]

We might refer to TMS as the 'parent technique' of rTMS. As we explained in the introduction, the latter

[3] "It should be stressed that the distinction between hard and soft impacts is not neutral or descriptive. Instead, it is a largely rhetorical distinction brought into play by one group of powerful players (policymakers and technology actors) for practical – or strategic – purposes" ([13], 7).

[4] This section is based on Arnaldi and Gorgoni [30]. See also Gorgoni, in this journal issue, for a very similar, but not identical, description of responsibility paradigms [31].

draw from TMS and try to complement them, introducing an explicit reflection on responsibility. In the introduction, we described the extent to which TMS addressed the issue and we concluded that the reference to responsibility is not absent in this technique. The notion of a "moral" [11] or "accountability" regime [13] unequivocally refers to responsibility and to rules, practices, institutions, in which it is embedded. Moreover, the scenario narratives produced in the context of TMS necessarily discuss and describe the contingent institutional arrangements that preside over the assignment/assumption of responsibilities for ongoing and prospective techno-moral change. However, this topic has been largely left implicit and no specific framework for its analysis and description has been provided in TMS.

To define such a framework, we must acknowledge that responsibility is more than a single concept; it is "a syndrome" of related concepts [25], as several authors have detailed (see, as examples, [32–35]). While a review of these classifications is beyond the scope of this article, we rely on Gorgoni's work to examine how the conceptions of responsibility have changed (see, e.g., Arnaldi and Gorgoni [30]). This author developed the work of François Ewald on the historical evolution of responsibility to distinguish four "paradigms" for this concept that can be useful in summarising and contextualising the multiplicity of approaches to responsibility. We briefly describe these paradigms in the paragraphs that follow.

The *paradigm of fault* corresponds to the traditional moral and legal idea of responsibility arising from faulty causation. This form of responsibility is based on the moral (and legal) obligation of individuals to respond to their faults. The liability of faulty agents can result in their being subjected to adverse treatment (sanctions). This model of responsibility is retrospective insofar as responsibility is established by the assessment of past actions according to a set of given criteria and rules.

The *paradigm of risk* replaces sanctions with compensation. In industrial modernity, the idea of risk and the mechanisms of risk management through insurance effectively disassociated responsibility from fault and (faulty) action, separating compensation from liability. This view of responsibility relies on notions such as event, victim and risk (calculation), which make the agent's contribution to the production of damage irrelevant for the compensation to be granted. In this paradigm, the logic of compensating victims for the damages they suffered prevails on the retribution of the faulty agent. This responsibility paradigm places on abstract calculations (risk assessment) and systemic mechanisms (risk management and prevention) the burden of adverse events and their consequences. As opposed to the paradigm of fault, here, responsibility is disconnected from the moral qualities of the agent, which were instead essential in the former. In the fault paradigm, the actors are responsibilised by being accountable for the negative consequences of their actions and, because of the faulty nature of their behaviour, they can be subject to retribution. The paradigm of risk rests on the idea of social solidarity rather than individual responsibility. Paradoxically, this shift leads to the de-responsibilisation of the agents, as their contribution to the production of the damage is irrelevant for the compensation mechanism to operate. Moreover, this model of responsibility is prospective insofar it anticipates the probability of adverse consequences by means of risk calculation and management. However, at the same time, it remains linked to a retrospective logic in that it anticipates the occurrence of damage but provides compensation for a damage that actually occurred.

The *paradigm of safety* ("paradigm of precaution", in Gorgoni's article for this special section [31]) is centred on the idea of precaution. The paradigm of fault assumes an identifiable author or a discernible causal chain to assign responsibilities. The paradigm of risk demands the identification of a causal chain, which can lead to events and not necessarily to an agent. The uncertainty surrounding science and technology and their outcomes challenges both tenets. The precautionary principle addresses this problem by framing responsibility as prevention rather than ascription and subjection to sanction or compensation. Precaution enters when the undesirable harmful consequences of scientific innovation cannot be ruled out under the criteria of risk governance. The precautionary principle creates neither new forms of liability nor new criteria for risk assessment. Instead, there is an emphasis on actors' prospective responsibility: agency is reasserted, and responsibility is again moralised. However, here, responsibilisation is predominantly negative: the moral agent is required to abstain from action to avoid (prevent) negative consequences.

Responsible research and innovation (RRI), the recent governance approach for science and innovation in the EU, can be viewed as an emerging perspective on responsibility that creatively combines elements from the three paradigms of fault, risk and safety, and

innovates them [16, 17]. Despite some differences, the literature on RRI largely shares a common understanding of responsibility and its dimensions [36]. RRI is prospective such as the risk and precaution paradigms. However, it does not seek to compensate or prevent the negative consequences of innovation. It instead aims to steer innovation and its outcomes and ensure that they are aligned with societal needs and values. RRI aims to be a driver and not a constrainer of research and innovation, which means that the responsibility of social actors goes beyond legal obligations and includes the voluntary commitment to shape ethically acceptable and societally desirable research and innovation trajectories. Delivering on this commitment requires collaboration: unlike liability and risk, responsibility is first and foremost mutual and shared across different actors with different roles and powers throughout the innovation process. Arnaldi and Gorgoni [30] maintain that this original combination of features sets RRI apart from other responsibility paradigms as it goes beyond the traditional emphasis on fault and punishment, risk and compensation, uncertainty and precaution. It aims at steering the innovation process from the inside towards societal goals rather than coping with its (actual or anticipated) unwanted and unintended effects.[5] Table 1 summarises the main characteristics of each paradigm.

As described above and summarised in Table 1, these four paradigms are characterised by their temporal orientation, the distinctive means they rely upon for enacting responsibility, and the level on which imputation of responsibility is adjudicated (individual or collective). Taken together, the paradigms offer an overview of the conceptual diversification of responsibility and capture the different positions of these paradigms with regard to two opposite semantic poles that are always present when responsibility is discussed, i.e., the opposition between active and passive responsibility. Active responsibility implies a deliberate will to take responsibility for something. We associate active responsibility with expressions such as acting or behaving responsibly. Passive responsibility refers to responsibility imputation, and it is associated with concepts such as causality, liability, and accountability. In general, it is fair to say that the different definitions of responsibility oscillate between these two poles. Among the paradigms, Gorgoni considers safety and RRI to principally refer to an active conception of responsibility as actors are expected to assume *ex ante* responsibility for their actions, while the contrary applies to the paradigm of fault, which is based on the *ex post* imputation for the consequences of faulty behaviour. As we have seen, risk is ambivalent, containing both elements of active and passive responsibility, as, on the one hand, it seeks to actively anticipate the probability of adverse events by means of risk assessment, prevention, and management, while, on the other hand, it remains linked to a passive logic in that compensation is provided after responsibility is ascertained.

In the elaboration of the rTMS technique, we use these four paradigms as ideal-types of responsibility, as the references upon which "responsibility regimes" are modelled. We therefore maintain the descriptive perspective adopted by the original TMS technique to discuss moral and accountability regimes; however, we propose a more articulated framework to analyse and describe the hypothetical responsibility arrangements that can emerge in the wake of moral controversies about new technologies, and as a consequence of the disruptions they cause to moral principles and practice. The following section illustrates how the procedure of TMS can be complemented to add this explicit reflection on the dimension of responsibility.

Expanding Techno-Moral Scenarios: The rTMS Technique

As we explained in the introduction, this paper attempts to complement the TMS approach by explicitly considering the issue of responsibility in science and technology. This section retraces the procedure defined by Boenink, Stemerding and Swierstra to create TMS [11] and adapts the steps they highlight to accommodate this additional focus of attention. Again, it is important to note that responsibility is already present in the original design of TMS. Moreover, TMS have been used in the context of RRI-related projects [14, 15]. However, this article elaborates the reference to responsibility in order to include the paradigms described in the Section titled "Responsibility Paradigms".

In rTMS, the four responsibility paradigms serve as ideal-types of responsibility that can be used to sketch

[5] We follow Gorgoni in naming these four responsibility paradigms. However, this author has argued that the features of RRI reflect a broader understanding of responsibility that is rooted in a specific political-economic framework, namely neoliberalism [30]. The term 'responsibilisation' has been used to describe the definition of responsibility typical of this political-economic context and might well be used to label the fourth paradigm identified by Gorgoni [37].

Table 1 The paradigms of responsibility

Paradigm	Guiding principle	Means	Target	Dimension	Temporal orientation
Fault	Liability	Sanction	Negative outcomes	Individual	Retrospective
Risk	Damage	Compensation	Negative outcomes	Systemic	Prospective/Retrospective
Safety	Uncertainty	Precaution	Negative outcomes	Collective	Prospective
RRI	Responsiveness	Participation	Negative and positive outcomes	Collaborative	Prospective

Source: Adapted from Arnaldi and Gorgoni [30]

different regimes, which correspond to them in terms of: (1) active or passive responsible entities, i.e., the entities (actors, artefacts, events) that either take or are assigned responsibility and are subjected to (positive or negative) consequences because of this assignment [38]; (2) temporal orientation, i.e., whether responsibility is prospective or retrospective; (3) the means of enactment, i.e., the arrangements in which the various characterisations of responsibility are embedded.

According to the scenario logic, the paradigms can be used to illustrate hypothetical situations, internally consistent descriptions of the regimes presiding over the assumption and assignment of responsibility for the moral consequences of technological transformations. In the mechanics of scenario-building, the paradigms allow for the delineation and discussion of alternative scenarios for responsibility. In the TMS literature, there is no precise indication of the number of alternative scenarios to be created. On this issue, Boenink says "scenario builders can decide to construct two (or even more) closures of a specific controversy, after which the future may diverge in different directions, and alternative futures may be imagined" ([29], 63). Our proposal is to create four 'raw', alternative 'responsibility scenarios', each of them based on a specific paradigm, and to then revisit them to create alternative combinations leading to the drafting of one or more final 'meta scenarios' based on the consistent combination of elements from all of them (see Fig. 1).

As in the original TMS technique, the centrepiece of rTMS is the use of vignettes. Vignettes are "short stories about hypothetical characters in specified circumstances, to whose situation the interviewee is invited to respond" ([39], 105). Vignettes have been used in sociology to explore normative beliefs, as they allow researchers to elicit situationally specific answers from respondents [39] who are confronted with "precise references to what are thought to be the most important factors in [their] decision-making or judgment-making processes" ([40], 94). Whereas, in quantitative survey designs, vignettes are usually followed by a predefined set of answers, when they are used in a qualitative context, their non-directive nature causes them to resemble projective techniques in psychology, allowing respondents to define the meaning of the situation represented. The qualitative characterisation of vignettes "reorients research towards the meanings respondents ascribe to situations, and accounts of discursive practice and rhetorical positions. […] Questions about validity, where a vignette's 'realness' is understood in terms of veracity (the closeness of its approximation to reality), are replaced by questions about the meaningfulness of a vignette to participants in research" ([41], 83). Texts, videos, and audio have all been used as stimuli to elicit respondents' opinions [41]. All these types of materials have one thing in common: their narrative form.[6] Techno-moral scenarios use vignettes as stimuli to encourage group reflection on the moral issues stemming from specific developments in science and technology. Essentially, vignettes are scenarios themselves as they narrate hypothetical future encounters of technology and morality. Both written and video materials have been used.[7]

In the literature, the creation of techno-moral scenarios follows three successive steps. First, the writers of scenarios/vignettes "sketch the moral landscape" of the future technology that is being examined with regard to the technical feasibility of the developments that are analysed. The active scanning of past controversies is crucial in this phase of the process to establish similarities and analogies for the considered hypothetical cases. Second, hypothetical moral controversies are explored. The promises and the objections to the creation, application and diffusion of technical artefacts are examined and contrasted to anticipate possible controversies.

[6] For a general appraisal of vignettes in social science research, see Hughes and Huby [42].

[7] For an example of a video scenario, see, for example: https://www.youtube.com/watch?v=xGQ6Cp1dC4c (accessed September 15, 2018).

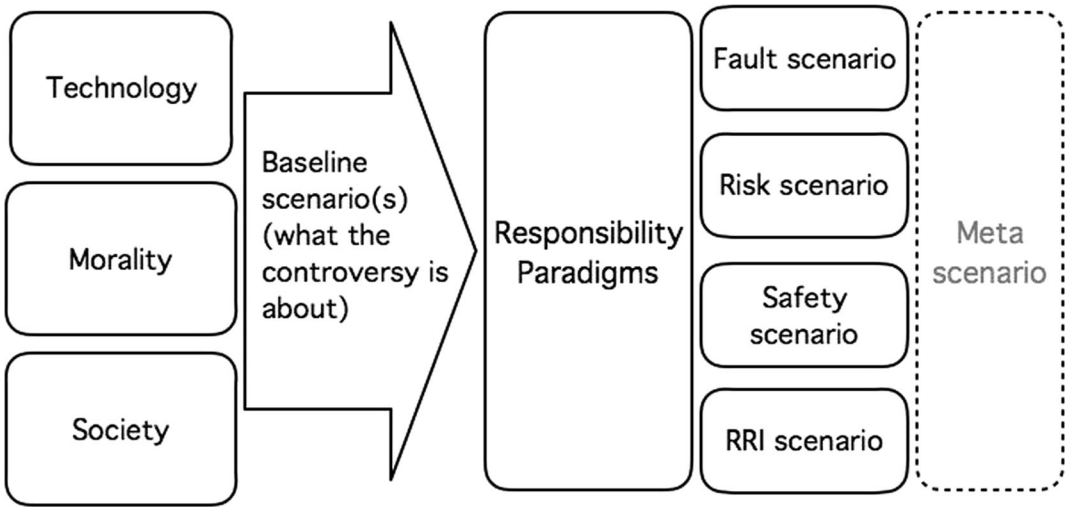

Fig. 1 – Generating responsibility scenarios: an overview of the process

Recurring tropes and types of ethical arguments, as well as analysis of past controversies, are used to generate hypothetical evolutions. Potential closures of controversies are generated. Third, the views and arguments listed in step 2 are weighted, and plausible resolutions are selected. The second and third steps can be rerun to extend the analysis further in the long term (this procedure is described in [13, 29]). To the best of our knowledge, TMS use techno-moral vignettes as stimuli for reflection; however, their development has largely remained the task of the analysts. In the scenario-building parlance, the scenario-building process is handled by the researchers, while scenarios as products are communicated to the public or to stakeholders as stimuli for subsequent reflection.

The technique presented here largely follows the TMS procedure, but departs from the original example in a few significant ways. First, we structure a process that can involve the public and stakeholders in the creation of scenarios and not only in their discussion. While this choice may be productive in terms of the 'engagement potential' of rTMS, it demands that a few passages of their procedure are (loosely) standardised to ensure their replicability in the context of scenario workshops. This loose standardisation includes the introduction of ad hoc questions and aids to focus the attention of workshop participants on aspects that TMS originally entirely leave to the expertise, creativity, and analytical acumen of the researchers, such as the exploration of the technological and moral landscape, as well as of the promises about the impacts of technological options and the objections raised against them. Second, we consider responsibility paradigms and responsibility regimes to be a complement to the original procedures and thus add specific activities that refer to this aspect. The following paragraphs illustrate the scenario-building process as we have revisited it. Each step is described and then summarised in a table, followed by an excerpt from a mock scenario exemplifying the outcome of each part of the process.[8]

Step 1 – Sketching the landscape: In the original TMS technique, the purpose of this step is to delineate the technological development to be discussed and to chart current relevant moral beliefs, practices and norms. Some historical background on both aspects is also provided to present an overview of past related or similar controversies and how they were dealt with [11]. For instance, an assessment of the current status and future prospects of pharmaceutical enhancement in the workplace can refer to past experiences with sport doping.

If one considers the scenario narratives developed in the TMS literature (see, e.g., [12, 13]), one can easily find that an important aspect that is covered in this phase of the process is the examination of the political, regulatory, and social context in which a controversy occurs. This analysis is, of course, unavoidable, and scenario narratives can hardly be exempted from exploring and

[8] To conduct a rTMS exercise, the facilitator can use material aids and tools, such as cards and a 'game' board, that can prompt and support the conversation among the workshop participants. In this article, we do not present these details for the sake of clarity and due to limited space.

describing these aspects. The Authors are well aware of this dimension, and they explicitly mention the fact that examining past social trends may "suggest specific path dependencies in the development of morality" and "may be used to decide which developments are more or less likely to be widely accepted" ([11], 10). However, as for responsibility regimes, it appears to us that this element is barely discussed in the literature on TMS. This different focus is presumably due to the intent to underscore the emphasis on morality, thus differentiating techno-moral scenarios from other, less specific, scenarios on the social impacts of technologies. Moreover, the researchers conducting TMS studies were perfectly aware of the way in which this social dimension could be addressed in their work, and they may not have perceived the need to dedicate a specific activity to this exploration. In rTMS, however, we assume that an exercise can be run with lay people as a participatory tool for creating scenarios. Therefore, for the sake of clarity and to ensure that this aspect is covered in the exercise, we think it is important to explicitly add the 'socio-economic landscape' to this initial, preliminary prospection and provide simple guidance for its examination. With regard to the 'socio-economic landscape', we refer to the macro-trends that set the stage for the relationships between technologies and society, the specific practices in which technologies are embedded as well as the social mechanisms through which values, institutions and normative expectations are distributed, solidified or transformed, such as institutionalisation and socialisation.[9]

As a loose guide to defining such relations, we suggest framing them according to two broad categories: 'competition' and 'cooperation'. Competitive relations describe the effort of two or more parties acting independently to secure benefits or resources (competition can lead, occasionally, to open conflict). Cooperative relations denote actors working together for a common purpose. The logic of using categories so wide is essentially a pragmatic one and it has the primary goal of ensuring both flexibility and minimal guidance for collective discussion and reflection. Moreover, the distinction between cooperative and competitive relations is significant in every context of human action; however, it has a peculiar validity for HE, in which performance, and performance augmentation, is a key issue. For instance, competition in the economy and the professions may exert pressures on individuals and encourage, or even compel, them to enhance themselves. Table 2 summarises this first step of the procedure.

Preliminary to this exploration of the landscape (or as a Step 0, as one may prefer to say), we suggest to develop a "persona", a fictional character to be featured in the scenarios. Personas are a technique used in the field of human-computer interaction (HCI) design, where descriptions of fictitious users are "used as the foundation for outlining a persona-scenario that investigates the use of an IT system from the particular persona's point of view" ([44], 58), thus generating design ideas that are tailored to different user groups in terms of requirements and fitness with the context of use. Personas are used in scenario-building beyond HCI; however, they do not represent the only form that a scenario can take. We decided to use personas in our scenarios as, in our view, this choice reinforces the concrete, vivid and evocative nature of scenario narratives: "[t]he intention behind the inclusion of personas is very much in keeping with the narrative nature of scenarios, to tell the story of a set of individuals and the lifestyle choices they might make, to try and personalise the scenarios and make them more accessible, more tangible for decision and policy makers" ([45], 21). The characteristics of the fictional protagonists will depend on the goal of the exercise, and reflect the specific segments of the population or the social actors considered by the analysis (e.g., women, a specific profession, a distinct industrial sector).[10] In general, the use of personas can contribute to the goal of conveying information in a "non abstract, non generalizing form" and is considered an asset of (techno-moral) scenarios ([29], 57). Table 3 lists the characteristics of the persona pictured in the mock scenario.

The beginning of the scenario is an example of the way in which the first step of the procedure can be translated into a narrative of the future:

[9] The distinction is based on the work of Talcott Parsons [43]. Socialisation refers to the incorporation of values and norms in the psychological structure of individuals; institutionalisation refers to the incorporation of select normative models in the social system of incentives and sanctions.

[10] Personas are built upon research data on users and, more generally, social groups and strata. Due to the mock nature of the scenario presented in this article, which was used for testing the procedure with students, the features of the persona who is the scenario protagonist were arbitrarily decided by the author.

Table 2 A description of Step 1: Sketching the landscape

Dimension	Aspects	Outcome	Questions
Techno-scientific landscape	Expertise, knowledge, technical feasibility and trajectories, areas of uncertainty	Overview of current technological options/artefacts	What are the features of the technological options/artefacts?
Moral landscape	Values (abstract principles) > institutions (general models of behaviour) > normative expectations (reciprocal orientation of actions)	Overview of the moral entities that are relevant to the technological options/artefacts that are considered	What are the past/current moral controversies about the options/artefacts? Are there similar artefacts?
Socio-economic landscape	Macro-trends, stakeholders, institutionalisation mechanisms (incentives and sanctions), socialisation	Overview of the socio-cultural trends, constellations of actors, and institutional mechanisms associating technical choices and moral entities	How are technological options and moral entities socially organised?

My name is Frans, good morning. Is it good, though? Certainly, this morning was special. I was summoned by my boss into her office. And when she calls you, then you are in trouble. I must say that after 15 years in the Financial Management Company, I did not expect to be faced with such a stark choice: being replaced by a computer or becoming able to work like a computer.

Neural implants have been used for decades in medicine. My father has one: he has Parkinson's, and, after he received the implant, he has been in much better shape. However, this is entirely different. I knew that the financial industry had to cut costs and that competition is fierce. Also, I knew a few colleagues who were fired because of the automation program, but I never thought I would be told to have an implant placed in my brain.

We need to cut costs, we need to be more productive, but do we really need something in our head to download the training? Do we really need something in our head that stimulates our attention all day long?

Table 3 – Characteristics of the "persona" used in the mock scenario

Name	Frans
Sex	Male
Age	45 y.o.
Profession	Financial analyst
Nationality	Dutch
Civil status	Married
Children	Yes

The brief text introduces the technology that is the matter of discussion (a neural implant) in a mundane situation, in which our protagonist is summoned by his manager in a labour environment that is marked by competitive pressures. Established medical uses of implants are contrasted with enhancement ones, which are discussed in the scenario.[11] As it will be clear in the next passages, these implants for brain enhancement are available on the market.

Step 2 – Generating controversies: the second step of the process explores the promises associated with technologies and the objections they raise. Like the previous one, this phase closely follows the original TMS technique, and it is aimed at generating and discussing potential ethical controversies about new technologies. Promises about anticipated technological developments are identified and described, and the moral objections they may raise are listed and discussed. Generating and discussing these topics can be difficult in a workshop attended by persons with little or almost no knowledge of the topic. Similar to what we did with the socio-economic landscape in Step 1, we therefore tried to identify and suggest some categories that can be

[11] This scenario is based on the work of the students of the Galilean School of Advanced Studies of Padova University (Italy), in which rTMS were first tested in may 2017. The classroom discussion and the scenario moved from the mission of Elon Musk's start-up 'Neuralink'. The company plans to build a 'neural lace', a brain-computer interface technology enabling bi-directional communication between the human brain and computers. The goal of Musk's venture is to allow humans to run external digital devices and fast download and upload information and data from/to computers to improve their cognitive performance in areas such as memory and information processing. There are many news articles describing Neuralink's goals (e.g., [46]) and critiquing the feasibility of Musk's ideas (e.g., [47]).

potentially useful to guide and stimulate the discussion in the context of a scenario workshop.

Accordingly, we propose distinguishing promise in two broad groups, describing two opposite, encompassing purposes of technology development. The first category is called 'protection', emphasising the role of technology in prevention and protection from threats, either physical or otherwise. The second category is called 'promotion', considering technologies as conditions to provide benefits, economic or otherwise. For instance, the development of safety technologies is justified based on the promise that they can keep us safe from environmental or even technology-induced threats (for instance, radiation, pollution). Conversely, enhancement technologies can be designed and applied because they promise to improve our performance according to disparate metrics such as economic output, saved time from sleep, or cognitive achievement. Of course, opposite types of promises can be used to characterise the same technology. For instance, in human enhancement, this happens when enhancing neuroimplants are supported to improve our work-related performance and/or as necessary modifications to save human beings from irrelevance or worse in a world populated by robots and artificial intelligence.[12]

With regard to objections, Boenink, Swierstra and Stemerding [11] suggest that recurring argumentation patterns and insights from the research on ethical, legal and social implications of technology can be used to formulate objections to these promises and thereby generate hypothetical controversies centred on issues such as the desirability of consequences, rights and obligations, distributive justice, and conceptions of the good life. For instance, arguments based on justice are central in critiques claiming that enhancement will increase inequalities. Moreover, the same authors emphasise that controversies also feature a second, higher-order, level that refers to "another cluster of […] arguments that does not deal directly with the new technology, but deal in more general terms with the relation between technological and moral development" ([12], 126). For instance, a debate initially centred on the consequences of a technology, for example, neurochip brain stimulation, can turn into a debate about the deontology and virtues of those who develop, and perhaps those who sell such devices. Finally, Swierstra and Rip [48] also refer to meta-ethical arguments, which capture different attitudes to the ongoing adaptations between society and technologies. These arguments can either consider technology-induced transformations of society as inevitable or view this transformation as a form of corruption. This position presumes a definite, stable and 'good' state of individuals and society, which are gradually or abruptly vitiated by technology.

Table 4 summarises this second step of the procedure.

The second excerpt of our scenario illustrates how this second step of the procedure describes the dialectic of promises and objections.

> They say my productivity will benefit from the implant. I am told my salary will follow suit.
> But I am not sure I will do it. They say the implant is safe, but having it in my brain means I will never be able to switch it off. I do not want to risk overheating my brain. I need it!
> Also, I have to pay for it, and they are damn expensive. I have a mortgage, and I cannot spend a whole lot of money on it, so who knows if the implant I can afford will be reliable. And what happens if there are malfunctions and I cannot buy a new one or repair it? I would probably be fired…
> The big bosses, their chips are paid for by the company. The costs we are cutting are probably paying for new, better chips, too. They will be able to do all the work without employees if they continue this way. What about us, those who do not have such high salaries?
> And after all, what would be the pleasure in learning and the satisfaction of doing a good job with some effort?
> I am not alone in having this opinion. The European Association of Financial Analysts sent a petition to the Health and Industry directorates of the European Commission and to the European Parliament. I hope they will soon take measures and protect and support us.

The promise of the neural implant (boosting work-related performance and, as a consequence, the user's salary) is criticised on different grounds. First, its 'hard impacts' are challenged: is it dangerous to the human brain? Then, the objections quickly move to moral

[12] This distinction has the same pragmatic logic as the differentiation of cooperative/competitive social relations in Step 1.

Table 4 A description of Step 2: Generating controversies

Dimension	Aspects	Outcome	Questions
Exploration of promises	Reasons for the creation, application and diffusion of the technical artefacts (protection, promotion)	Generating technical trajectories, according to a principle of plausibility	What are the plausible developments of the artefact that is considered?
Exploration of objections	Objections and critiques to the creation, application and diffusion of the technical artefacts	Exploring potential controversies according to typified moral argumentation (e.g., distributive justice, privacy, good life)	What are the possible moral objections and controversies?

grounds, mentioning issues of distributive justice (does the cost of the neural implant create social inequality and an unfair disadvantage for those who cannot afford it?) and of human identity (is it morally preferable to perform work-related tasks with effort? Is this something that has to do with our human identity?). This latter aspect introduces a different type of moral reasoning: while issues of safety and justice are treated from a consequentialist point of view, this latter objection establishes a link between individual effort, the moral character of the agent, and the moral worth of his actions. As explained above, focusing on single individuals does not preclude exploring how such controversies can be significant for society as a whole. For instance, in this case, we speculate that a professional association (the name we report is a fictive one) takes the concerns of its members to regulators (it is of course an outcome of the workshop discussion whether this hypothesis is a plausible and relevant development of the controversy that is discussed).

Step 3 – Closure and responsibility regimes: The third step of the process explores the possible compositions of the controversy and the regimes of responsibility accompanying these outcomes. Steps 1 and 2 define the context and the controversy, completing what we might consider to be a 'baseline scenario', sketching the hypothetical technology examined, the social, moral and ethical contexts and dynamics, and the possible controversies surrounding the technological developments that are assessed. Step 3 complements the analysis of techno-moral change described in Steps 1–2 with the possible "responsibility regimes" that preside over the assignment/assumption of responsibilities for the 'soft', wider impacts of technologies. In this step, four scenarios can be generated, one for each responsibility paradigm (fault, risk, safety, RRI), representing the 'pure' ideal-types of hypothetical responsibility regimes and systematically appraising the particular institutional arrangements that enable and constrain specific solutions to moral controversies. As we explained in the beginning of this Section, three elements are explored as they are constitutive of alternative regimes of responsibility: (1) active or passive responsible entities, (2) temporal orientation, (3) means of enactment. Drawing from the TMS literature [19], we distinguish the means in three categories: (1) technical (artefacts are changed to incorporate moral objections), (2) moral (institutions and normative expectations change to accommodate technical artefacts), (3) socio-political (sanctions and incentives are created to orient behaviour and technical solutions).

The development of the four scenarios allows the analysts or the participants in a rTMS workshop to create an inventory of responsibility arrangements that correspond to each of the responsibility paradigms. These four 'raw' scenarios can be useful in anticipating and systematically exploring the effects of different responsibility regimes on techno-moral controversies and their closing. As we stated in the introduction, responsibility arrangements are not morally neutral. Because of their moral significance, responsibility arrangements are not interchangeable in terms of their coherence and compatibility with moral standards and principles. This systematic exploration allows us to chart the plausible conditionalities, alignments, conflicts, and inconsistencies between moral and responsibility regimes. "Responsibility regimes" therefore fully participate in techno-moral change, and they are not their mechanical outcome. The intent of creating alternative responsibility scenarios is not to decide beforehand the types of responsibility regimes that are relevant

Table 5 Description of Step 3: Closure and responsibility regimes

Dimension	Aspects	Outcome	Questions
Exploring regimes of responsibility (one for each paradigm)	Centres of imputation (who is responsible?)	Identification of the entities that take part to the distribution of responsibility in a specific regime	Who takes the responsibility to act for closing the controversy? (assumption of responsibility) - Who/what is ascribed responsibility for the emergence of controversy? (ascription of responsibility) - Who is held accountable for the emergence of the controversy? (subjection to responsibility)
	Temporal orientation	Identification of the time-orientation of the responsibilities that are distributed in a specific regime	Is responsibility forward-looking or backward--looking?
	Means of enactment	Identification of the means that are used to support specific responsibility assignments and configurations: (1) technical (artefacts are changed to incorporate moral objections), (2) moral (institutions and normative expectations change to accommodate technical artefacts), (3) socio-political (sanctions and incentives are created to orient behaviour and technical solutions).	What is the way in which responsibility is realised enacted?

or legitimate but to determine the specific regimes of responsibility that can be associated with different moral controversies and their closure alternative closures (see also Table 5).

The following excerpts of the mock scenario describe alternative hypotheses relating to the different responsibility regimes, each of which suggests a possible closing of the controversy around the use of neural implants for improving work-related performance. The following paragraphs are brief examples of how alternative scenarios for the fault and risk paradigm could be developed.

Fault paradigm

I am particularly worried about the functioning of the chip. It is not a medical one. Am I sure it works in the same way? Now producers pay big fines if their products are not safe, but the brain is yours, not theirs. I know there are many who are suing the producers in cases of product failure, and their employers for making them work much longer and with much fewer breaks than the product guidelines require. I also know there is a consumer association that gives them legal assistance.

It should at least be mandatory for the employer to include this benefit in the employment contract, so that one can afford safe, reliable implants. I guess it is like dental care, and probably more important. But the boss wants results, not happy employees. It is not by chance that I know at least three colleagues who took their chips and remained silent, only to get better results in their evaluations… As I said, I have a mortgage and I need a job. So, never mind: I'll forget my concerns and I'll take the chip, too.

Risk paradigm

I am particularly worried about the functioning of the chip. It is not a medical one. Am I sure it works in the same way? I know that producers are insured for malfunctions and they pay, but, after all, you still may have your brain injured. They solve the problem sending you the product instructions, but that's theory, not the real world: nobody tells them that the boss makes you work for much longer and with much fewer breaks than the product instructions require.

At least, the employer should include regular neurological exams as part of the health benefits in the employment contract, so that those who cannot afford expensive implants can prevent brain injuries. I guess it is like dental care, and probably more important.

> But the boss wants results, not happy employees. It is not by chance that I know at least three colleagues who took their chips and remained silent, only to get better results in their evaluations… As I said, I have a mortgage and I need a job. So, never mind: I'll forget my concerns and I'll take the chip, too.

This brief passage shows that a hypothetical fault scenario may emphasise sanctions for the implants' producers if they fail to comply with safety standards or legal action to establish the liability of producers and employers. As an alternative, a risk paradigm may focus on possible compensations for users in the case of product failure or in introducing an insurance scheme or a product guarantee for implants that fail to function properly. Mandatory schemes for the employers to financially support employees who are requested to buy the implants or who are willing to monitor their health consequences may be instead seen as an answer to the potential of the neural implants to increase inequalities.

While the four responsibility scenarios describe 'pure ideal-types' of responsibility regimes, it is unlikely that different regimes (or their components) inspired by alternative responsibility paradigms do not co-exist in the real world. However, the various elements of different regimes are not necessarily compatible among them. For instance, one might argue that addressing sport doping from the point of view of the RRI and Fault paradigms can lead to incompatible solutions (e.g., voluntary compliance versus sanctions). Conversely, arrangements based on risk and precaution (risk and safety paradigms) or risk and liability (risk and fault paradigms) can be compatible, and they often are, for instance in the different treatments of (legal) responsibility deriving from wrongdoing or negligence.

Assessing these (in-)compatibilities is the last step of the process, and it is aimed at creating one or more "meta-scenarios" that incorporate the variations of responsibility arrangements identified in the four 'raw' scenarios in a plausible and coherent combination. The logic of this step is close to that of morphological analysis in scenario planning. Morphological analysis systematically explores the relationships between alternative elements and dimensions in a system to eliminate the incompatible combinations of factors and create plausible and internally consistent configurations [19]. In rTMS, this assessment has the purpose of designing more realistic responsibility regimes, which can have a better adherence to the trajectories that are more likely to develop "without succumbing to predictive claims" ([11], 10).

To exemplify this part of the process, our mock "meta-scenario" combines aspects from the risk (compensation in case of product failure) and fault (sanctions in case of company's wrongdoing) paradigms.

Meta-scenario

> I am particularly worried about the functioning of the chip. It is not a medical one. Am I sure it works in the same way? I know that producers are insured for malfunctions and they pay, but, after all, you still may have your brain injured.
> I know there are many who are suing the producers in the case of product failure, and they are suing their employers for making them work much longer and with many fewer breaks than the product guidelines require. I also know there is a consumer association that gives them legal assistance. That's good, but regulators should step in more aggressively to sanction any wrongdoing, from producers and employers alike.
> It should at least be mandatory for the employer to include this benefit in the employment contract, so that one can afford safe, reliable implants. I guess it is like dental care, and probably more important.
> But the boss wants results, not happy employees. It is not by chance that I know at least three colleagues who took their chips and remained silent, only to get better results in their evaluations… As I said, I have a mortgage and I need a job. So, never mind: I'll forget my concerns and I'll take the chip, too.

Concluding Remarks

This paper focused on the triad of morality, responsibility and technologies. The point of convergence of these three dimensions is the elaboration of a technique for creating scenarios that explore and anticipate their co-evolution. This technique draws from techno-

moral scenarios (TMS). Our goal was not to correct the TMS technique but to complement it with an explicit reflection on responsibility paradigms and on responsibility regimes, i.e., the institutional arrangements presiding over the assumption and assignment of responsibility for ongoing and prospected techno-moral change. We labelled this expanded approach 'rTMS' and maintained the descriptive and exploratory perspective of the original method.

In the context of our reflection on responsibility and human enhancement, we referred to TMS because they were created to anticipate and appraise the 'soft' social and moral impacts of new and emerging sciences and technologies (NEST). The chief characteristic of these soft impacts is their subjective definition, which causes them to elude the traditional regimes of technology assessment and management. HE technologies are certainly characterised by this type of impact. As they are intended to reshape our very identity, and their promises mostly concern the fulfilment of desires and aspirations, the subjective nature of their effects, and of the corresponding assessment, is indisputable. As we have shown, this peculiar aspect of HE has led to a stalled public and academic debate in which critics and advocates argue either about the social and moral legitimacy of the scientific and technological means used for enhancement goals or about the legitimacy of the goals themselves.

The descriptive approach of TMS to morality seemed to us a way out of this impasse, and the logic of rTMS follows it. This choice required a shift in focus: no longer were we interested in elaborating a set of moral standards or responsibility criteria according to which HE technologies could be assessed but rather we sought to outline a technique for exploring the alignments, contradictions and interdependencies of or between HE technologies, morality, and responsibility.

The creators of TMS framed moral change on three levels [11]: the macro-level of abstract moral principles; the meso-level of moral regimes, i.e., the institutions, rules, procedures that translate principles in more concrete requirements; and the micro-level, which concerns the discussions and negotiations in which moral issues and practices are negotiated in contingent situations. This paper's framing of responsibility is also based upon this tripartite model, distinguishing abstract paradigms, institutional arrangements (to which we have referred as "responsibility regimes"), and the modes in which responsibility is enacted in particular social contexts and interactions. By using personas in scenario narratives, rTMS aims to create a micro-level entry point for analysing the three levels from the lens of the hypothetical experiences of individuals, representing 'ideal types' of social groups and actors.

However, it is clear from the article that our main focus of attention is on the 'regimes' or arrangements on the meso-level, which we identified as the site where responsibility and morality meet. This choice is probably unsurprising, as this article is written by a sociologist tinkering with morality and responsibility. Furthermore, this emphasis is coherent with the practical goal of this work: developing a participatory technique that is essentially interested in pragmatically anticipating and systematically exploring the institutional arrangements shaping this encounter, in order to delineate hypothetical moral and responsibility regimes which are coupled with the ethical controversies around new technologies and their alternative resolutions [12]. This work still has a practical-normative significance, insofar as it fosters our capacity to build robust moral and responsibility regimes for HE technologies by systematically appraising the transformations of morality and responsibility.[13]

As we have said, our goal was to expand the reach of TMS by developing in a more detailed and explicit way the reference made by this technique to responsibility. Combining responsibility and morality facilitates the exploration of the relationships between these two dimensions. It is up to the reader to judge whether we have succeeded in combining these two dimensions in the rTMS procedure. Certainly, if we did, this result was attainable within the narrow confines of this technique and its application, while it was much beyond the scope of our work to engage in a thorough conceptual and theoretical understanding of the relationships between morality and responsibility. This narrow and pragmatic focus was a condition of feasibility of this effort; however, it nonetheless represents the main limitation of our work, the further development of which would certainly benefit from a more comprehensive approach.

[13] While this article applies rTMS to HE, the application perimeter of this technique is not limited to this subject and covers the 'soft impacts' of new and emerging technologies in general, rather than of specific domains.

Public engagement techniques are often open-ended, and they are frequently modified to improve their functioning. We hope rTMS will not be an exception, and we invite contributions and critiques that can strengthen not only their practical implementation but the very theoretical foundations on which they rest.

Acknowledgements The author would like to thank two anonymous reviewers for their helpful comments. I wish to thank the participants in the mutual learning workshop on "Responsibility and human enhancement" held in Padova, Italy, on May 22, 2017, for their precious comments and suggestions on the first draft of this article. In particular, we would like to express our gratitude to Tsjalling Swierstra who acted as the discussant of this paper during the workshop. Finally, I gratefully acknowledge the funding from the Independent Social Research Foundation (ISRF), Flexible Grants for Small Research Groups program, which made it possible to conduct this research as part of the project "Responsibility and Human Enhancement. Concepts, implications and assessments".

References

1. Sauter A, Gerlinger K (2014) The pharmacologically improved human: performance-enhancing substances as a social Challenge BoD - Books on Demand, Norderstedt
2. Kass LR (2003) Ageless bodies, happy souls: biotechnology and the pursuit of perfection. The New Atlantis 1:9–28
3. Sandel MJ (2009) The case against perfection: ethics in the age of genetic engineering. The Belknap Press of Harvard University Press, Cambridge, MA
4. Garcia T, Sandler R (2008) Enhancing justice? NanoEthics 2:277–287. https://doi.org/10.1007/s11569-008-0048-5
5. McVeigh J, Evans-Brown M, Bellis MA (2012) Human enhancement drugs and the pursuit of perfection. Adicciones 24:185–190
6. Fukuyama F (2003) Our posthuman future: consequences of the biotechnology revolution. Profile Books, London
7. Harris J (2011) Enhancing evolution: the ethical case for making better people. Princeton University Press, Princeton
8. Savulescu J, Sandberg A, Kahane G (2014) Well-being and enhancement. In: Savulescu J, Meulen RHJ ter, Kahane G (eds) Enhancing human capacities. Wiley-Blackwell, Chichester, p 1–18
9. Bostrom N, Roache R (2011) Smart policy: cognitive enhancement in the public interest. In: Savulescu J, Meulen RHJ ter, Kahane G (eds) Enhancing human capacities. Wiley-Blackwell, Chichester, p 138–149
10. Meacham D (2015) The subject of enhancement: augmented capacities, extended cognition, and delicate ecologies of the mind. The New Bioethics 21:5–19. https://doi.org/10.1179/2050287715Z.00000000063
11. Boenink M, Swierstra T, Stemerding D (2010) Anticipating the interaction between technology and morality: a scenario study of experimenting with humans in bionanotechnology. Studies in Ethics, Law, and Technology 4. https://doi.org/10.2202/1941-6008.1098
12. Swierstra T, Stemerding D, Boenink M (2009) Exploring techno-moral change: the case of the ObesityPill. In: Sollie P, Düwell M (eds) Evaluating new technologies. Springer Netherlands, Dordrecht, pp 119–138
13. Swierstra T (2015) Identifying the normative challenges posed by technology's 'soft' impacts. Etikk i praksis - Nordic Journal of Applied Ethics 9. https://doi.org/10.5324/eip.v9i1.1838
14. Stemerding D (2015) iGEM as laboratory in responsible research and innovation. Journal of Responsible Innovation 2:140–142. https://doi.org/10.1080/23299460.2014.1002171
15. Douglas CM, Stemerding D (2014) Challenges for the European governance of synthetic biology for human health. Life Sciences, Society and Policy 10:6. https://doi.org/10.1186/s40504-014-0006-7
16. von Schomberg R (2013) A vision of responsible research and innovation. In: Owen R, Bessant J, Heintz M (eds) Responsible Innovation. John Wiley & Sons, Ltd, Chichester, pp 51–74
17. Stilgoe J, Owen R, Macnaghten P (2013) Developing a framework for responsible innovation. Res Policy 42: 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008
18. Wack P (1985) Scenarios: shooting the rapids. Harv Bus Rev:139–150
19. Amer M, Daim TU, Jetter A (2013) A review of scenario planning. Futures 46:23–40. https://doi.org/10.1016/j.futures.2012.10.003
20. Börjeson L, Höjer M, Dreborg K-H, Ekvall T, Finnveden G (2006) Scenario types and techniques: towards a user's guide. Futures 38:723–739. https://doi.org/10.1016/j.futures.2005.12.002
21. Roubelat F (2000) Scenario planning as a networking process. Technol Forecast Soc Chang 65:99–112. https://doi.org/10.1016/S0040-1625(99)00125-0
22. Masini EB, Vasquez JM (2000) Scenarios as seen from a human and social perspective. Technol Forecast Soc Chang 65:49–66. https://doi.org/10.1016/S0040-1625(99)00127-4
23. Barbieri Masini E (1993) Why futures studies? Grey Seal, London
24. Palm E, Hansson SO (2006) The case for ethical technology assessment (eTA). Technol Forecast Soc Chang 73:543–558. https://doi.org/10.1016/j.techfore.2005.06.002
25. Est R van, Rerimassie V, Keulen I van, et al (2014) Intimate technology: the battle for our body and behaviour. Rathenau Institute, The Hague
26. Luhmann N (2005) Risk: a sociological theory. Aldine Transaction, New Brunswick
27. Nordmann A (2007) If and then: a critique of speculative nanoethics. NanoEthics 1:31–46. https://doi.org/10.1007/s11569-007-0007-6
28. Nordmann A, Rip A (2009) Mind the gap revisited. Nat Nanotechnol 4:273–274. https://doi.org/10.1038/nnano.2009.26
29. Boenink M (2010) Imagining the future: how vignettes and scenarios might improve ethical reflection on synthetic biology for health purposes. In: Szebik I. (ed) Ethics and clinical applications of synthetic biology: an interdisciplinary dialogue. SYBHEL project, Bilbao, p 55–64
30. 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 12:6. https://doi.org/10.1186/s40504-016-0038-2
31. Gorgoni G (2018) Responsible research and innovation and the governance of human enhancement. NanoEthics 12(3). https://doi.org/10.1007/s11569-018-0326-9
32. Vincent NA (2011) A structured taxonomy of responsibility concepts. In: Vincent NA, van de Poel I, van den Hoven J (eds) Moral responsibility. Springer, Dordrecht, p 15–35
33. Hart HLA (2008) Punishment and responsibility: essays in the philosophy of law. Oxford University Press, Oxford
34. Davis M (2012) "Ain't no one here but us social forces": constructing the professional responsibility of engineers. Sci Eng Ethics 18:13–34. https://doi.org/10.1007/s11948-010-9225-3
35. van de Poel I, Nihlén Fahlquist J, Doorn N, Zwart S, Royakkers L (2012) The problem of many hands: climate change as an example. Sci Eng Ethics 18:49–67. https://doi.org/10.1007/s11948-011-9276-0
36. Forsberg E-M, Quaglio G, O'Kane H et al (2015) Assessment of science and technologies: advising for and with responsibility. Technol Soc 42:21–27. https://doi.org/10.1016/j.techsoc.2014.12.004
37. Shamir R (2008) The age of responsibilization: on market-embedded morality. Econ Soc 37:1–19. https://doi.org/10.1080/03085140701760833
38. Arnaldi S, Bianchi L (2016) Responsibility in science and technology. Elements of a social theory. Springer VS, Wiesbaden
39. Finch J (1987) The vignette technique in survey research. Sociology 21:105–114. https://doi.org/10.1177/0038038587021001008
40. Alexander CS, Becker HJ (1978) The use of vignettes in survey research. Public Opinion Quarterly 42:93. https://doi.org/10.1086/268432
41. Wilks T (2004) The use of vignettes in qualitative research into social work values. Qualitative Social Work: Research and Practice 3:78–87. https://doi.org/10.1177/1473325004041133
42. Hughes R, Huby M (2004) The construction and interpretation of vignettes in social research. Soc Work Soc Sci Rev 11:36–51. https://doi.org/10.1921/17466105.11.1.36
43. Parsons T (1991) The social system. Free Press, New York
44. Madsen S, Nielsen L (2010) Exploring persona-scenarios - using storytelling to create design ideas. In: Katre D, Orngreen R, Yammiyavar P, Clemmensen T (eds) Human work interaction design: usability in social, cultural and organizational contexts. Springer, Berlin, pp 57–66
45. Milne S (2009) Scenarios and personas: towards a methodology for portraying the carbon intensity of UK lifestyles to 2030, Scott Milne, RESOLVE working paper 06–09. University of Surrey
46. Fourtané S (2018) Neuralink: how the human brain will download directly from a computer. https://interestingengineering.com/neuralink-how-the-human-brain-will-download-directly-from-a-computer. Accessed 29 Sep 2018
47. Regalado A (2017) With Neuralink, Elon Musk promises human-to-human telepathy. Don't believe it. https://www.technologyreview.com/s/604254/with-neuralink-elon-musk-promises-human-to-human-telepathy-dont-believe-it/. Accessed 29 Sep 2018
48. Swierstra T, Rip A (2007) Nano-ethics as NEST-ethics: patterns of moral argumentation about new and emerging science and technology. NanoEthics 1:3–20. https://doi.org/10.1007/s11569-007-0005-8