31 March 2023, Preprint v1 · CC-BY 4.0

Research Article

Social responsibility, disciplinary moral identity, and not-so-value-free biomedical research(ers)

Vincenzo Politi[1]

1. Universitat Autónoma de Barcelona, Spain

For the expanded concept of social responsibility promoted by several science policies, scientists ought to steer their research towards socially desirable ends and to consider not only the quantifiable harmful consequences of their work, but also the transformative impacts that innovation may have on society. Some science policy frameworks seek to implement a sense of social responsibility in science and technology through a series of interventions at the 'midstream stages' of research that aim at encouraging the cultivation of critical and moral reflection among researchers. Such interventions include interdisciplinary collaborations between STEM and Medical, on the one hand, and scholars from the Humanities and the Social Sciences, on the other; as well as the requirement of engaging with the public and other stakeholders, in order to acknowledge and integrate the stances of different societal actors. The philosophical backbone of these science policies is the rejection of the so-called Value-Free Ideal, which states that non-epistemic values must play no role in the internal stages of research and in the justification of scientific claims. While there are fields in which researchers may regard themselves as detached from social or ethical concerns, in fields that are highly value-laden and whose main objectives are already socially beneficial, such as in biomedical research, these policies may lead to problematic consequences. In particular, the proposed interventions may reinforce biomedical researchers' disciplinary moral identity and self-perception of moral excellence. Such a reinforcement could hinder interdisciplinarity and public engagement, that is the very measures designed to make research socially responsible. Further investigation into the disciplinary moral identity of scientists in general, and biomedical researchers in particular, must be encouraged for the improvement of responsible research in biomedicine.

doi.org/10.32388/OSWB8G

1. Introduction

For a long time, the so-called *Value-Free Ideal* was taken as the foundation of the relations between science and society. With its demise, a new conception of the role of science (and scientists) within society have emerged. For the newly emerging *co-production ideal*, scientists ought to steer their research towards socially desirable ends and, to do so, they must integrate the views and stances of other stakeholders and of the general public into their work. The new ideal has inspired the development of several science policy frameworks that aim at institutionalising a sense of social responsibility in science through a series of interventions that are supposed to encourage the cultivation of reflexivity and social openness among researchers. Such interventions include interdisciplinary collaborations between STEM and Medical, on the one hand, and scholars from the Humanities and the Social Sciences, on the other; as well as the requirement of engaging with the public and other stakeholders, in order to acknowledge and integrate the stances of different societal actors. What these policies seem to presuppose is that researchers do not already engage in reflexivity and do not already consider the stances of other societal actors. They presuppose, that is, the *descriptive* validity of the very Value Free Ideal that they deem as untenable.

It is possible that some researchers working in specific fields do indeed regard themselves as detached from social or ethical concerns and reason in a ‘value free’ fashion. This may depend on how they have been trained and by the (often tacit) methodological assumptions dictating how research ought to be conducted in a particular field. However, those who work in highly value-laden research fields, the main objectives of which are already socially beneficial and ‘people-oriented’, may have a different *disciplinary moral identity*. The problem, in such cases, is not whether scientists make use of value judgments, but whether they make use of the *right* value judgments.

In this article I will argue that, in highly value-laden fields such as biomedicine and the health sciences, disciplinary moral identity could pose some problems. In particular, some of the interventions promoted by the science policies may actually have unwanted consequences. In particular, such interventions may reinforce biomedical researchers’ self-perception of righteousness and moral excellence. Such a reinforcement could hinder interdisciplinarity and public engagement, that is the very measures designed to make research socially responsible.

In the next section, I briefly explain the philosophical backbone of the current science policies and the kind of social responsibility they aim at institutionalising. In section 3, I clarify why such policies are

doi.org/10.32388/OSWB8G

deemed as necessary for biomedicine and the health sciences by discussing some of the potential societal implications of contemporary biomedical research, especially when coupled with innovative developments in technology. In section 4, however, I suggest that we reconsider the issue of social responsibility for ‘not-too-value-free’ fields such as biomedicine from a different philosophical perspective. Such a perspective not only considers the ‘value-freedom’ and ‘value-ladenness’ of research fields as a matter of degrees but also how, depending on such degrees, members of different scientific communities may share a more or less strong disciplinary moral identity. In sections 4 and 5, I consider the potential risks associated with an excessive reinforcement of biomedical researchers’ self-perception of moral excellence. The moral identity of scientists in general, and of biomedical researchers in particular, need to become an object of empirical studies as well as of philosophical analysis.

2. From one ideal to the other

For a long time, scientists were allowed, and actually expected, to freely pursue the fundamental truths of Nature while leaving to others, such as politicians and policy makers, the responsibility of implementing the results of their research into society. The philosophical foundation of such a model of the science-society relation is the so-called Value-Free Ideal (Lacey 1999; Kincaid, Dupré and Wylie 2007). This ideal states that, although science is guided by epistemic or cognitive values (such as simplicity, consistency, and so on; see Kuhn 1977, McMullin 1983, Laudan 1984), the influence of non-epistemic values (such as moral and political judgments) is admissible only in the external stages of research (for example, at the agenda-setting stage), but must have no place in its internal stages (for example, in theory appraisal or in the justification of scientific claims). One of the main motivations for holding the Value-Free Ideal is a defence of the epistemic authority of science: the closer scientific knowledge production is to society and its politics, one may argue, the more subject to manipulations and corruption it becomes, and the less authoritative as a guide towards the truth the institution of science risks to be. Hence the need for a clear ‘social contract’ between science and society, regulating tasks and dividing duties and responsibilities (Polanyi 1962).

In the past few decades, STS scholars, science policy makers, and philosophers have challenged the Value-Free Ideal. Several philosophers, for example, have maintained that science is intrinsically ‘value-laden’ and, among other things, they have argued that: the epistemic/non-epistemic values distinction is untenable (Rooney 1982); non-epistemic values are indeed necessary in many of the internal stages of research (Douglas 2009); non-epistemic values do not compromise scientific objectivity and its authority (Koskinen 2020); sometimes non-epistemic aims (i.e., a ‘fast solution’) may have a priority on epistemic aims (i.e., the ‘right solution’; Elliott and McKaughan 2014).

The demise of the Value-Free Ideal lead to the development of an alternative ideal, in which science is not regarded as inhabiting a detached and perhaps privileged ‘Ivory Tower’ but, rather, as an institution which is deeply embedded in, and open to, society (see, among many others, Funtowicz and Ravetz 1993; Gibbons et al 1994; Kitcher 2001, 2011; Douglas 2009; Kourany 2010). According to the new ideal of the science-society relation, which is known under the name of *co-production ideal*, scientists ought to align their work to the needs of other stakeholders, steer research towards socially just and desirable ends, and actively collaborate with a plurality of societal actors (Jasanoff 2004; MacNaghten 2021).

The co-production ideal promotes an ‘expanded’ conception of scientists’ responsibility. That science ought not to harm people and society, of course, is unquestionable. Committees and professional integrity codes are already in place to prevent harm, by regulating and putting legal and ethical constraints to research. The potential implications of scientific and technological research, however, are not only clear-cut and easily quantifiable harms to be prevented (as in the case, for example, of technological devices that may clearly compromise the health of end-users). Science and technology, in fact, may also involve so-called ‘soft impacts’ (as in the case of technologies transforming how people interact, socialise, and perceive themselves; see van der Bug 2009). Although necessary, therefore, quantitative risk/benefit analyses for preventing harms may not be sufficient for socially responsible research, which also requires a critical reflection on the transformative and disruptive power of innovation (Jasanoff 2016; Forssén et al 2011). Such reflexivity and openness to societal concerns cannot be regulated only ‘from the outside’, as existing legislative and ethical frameworks do, but must be encouraged ‘on the inside’, that is among working researchers.

The sense of social responsibility promoted by the co-production ideal requires the integration of social and value diversity in research. Philosophers have analysed the epistemic benefits of socially diverse scientific communities (see Harding 2015, Intemann 2010, Jaggar 2004, Rolin 2019, Wylie 2003), and the moral and political arguments in favour of more diversity in values and stances driving scientific research (Longino 1990, 2002). Science policy makers have developed various frameworks such as, for example, Technology

Assessment (TA) (Grunwald 2018), the Ethical Legal and Social Implications framework (ELSI) (Fisher 2015), and Responsible Research and Innovation (RRI)(von Schomberg 2013; Owen et al 2013), targeting the internal or 'midstream' stages of research, and aiming at encouraging the cultivation of a reflexive attitude among scientists and engineers (Fisher, Mahajan and Mitcham 2006). Midstream interventions may include *interdisciplinary collaborations* between lab researchers and scholars from the Humanities, in order to make research communities more diverse and pluralistic, and the requirement of an active *public engagement*, aimed at involving different stakeholders in research.

The convergence of philosophical and science policy literature on topics such as the social responsibility of science and the need to make scientific communities more diverse and socially engaged has been recently discussed by Koskinen (2022), who also analyses why, oftentime, what works in the philosophical theory fails to be properly institutionalised in practice. In this article, however, I will not discuss whether and why the interventions suggested by RRI or similar policies may fail to bring out the expected results. Rather, I will consider the unexpected or even undesirable consequences that these policies may have on some 'not-so-value-free' fields of research, such as biomedicine and the health sciences. Before doing so, however, it is necessary to see whether and up to which point biomedicine is required to be 'socially responsible' in the sense promoted by the current policies (in the sense, that is, of taking into consideration the potential transformative impacts of its results).

3. How 'transformative' is innovative biomedical research?

The new science policies that aim at institutionalising social responsibility in research have been developed out of a concern for the socio-ethical challenges posed by the Human Genome Project as well as by the innovative technologies employed and produced in emerging fields such as synthetic biology, nanotechnology, AI, and robotics. Several funding organisations now require biomedical research projects to comply with some of the new policies.

One might wonder, however, whether current research in biomedicine and in the health sciences is 'innovative' in the sense captured by these policies. To begin with, the idea of 'innovation' underlying RRI and similar frameworks is often linked to plans of national economic growth driven by the financial exploitability of the outcomes of 'Research and Innovation' (R&I) projects. Instead of assuming it as intrinsically beneficial, some have argued that R&I-driven economic growth deserves further reflection and critical assessment, especially in view of its environmental (un) sustainability and societal impacts (de Saille and Medvecky 2016). The far more modest problem discussed in this article, however, concerns the fact that the main objective of much biomedical research does not seem to be the production of immediately marketable goods. Biomedical projects may be ‘innovative’, but in the sense that they may modify the clinical practice and the healthcare system.

For these reasons, some have argued that the challenges posed by the rise of new health technologies are different from those tackled by current science policies. Therefore, innovative biomedical research may require a governance framework of its own (Pacifico Silva et al 2018; Lehoux et al 2018; Lehoux et al 2019). Here I will remain agnostic on whether we need to develop a brand new policy specific to biomedical and health research. Once again, my worry is more general: it has to do with the clarification of the potential societal implications of biomedical research such that the implementation of a policy framework for social responsibility is justified and actually needed. This is a more fundamental issue, since the need for a policy to make biomedical research socially responsible could be unclear or even questioned.

As explained in the previous section, the co-production model was born out of a reaction against the Value-Free Ideal. The problem is that biomedical research may not look so ‘value-free’ to begin with. On the one hand, its main aims (such as finding new effective treatments, improving the existing ones, contributing to patients’ well-being both during and after treatment, and so on) appear to be socially desirable in themselves. On the other hand, biomedical research already avoids potential harm by complying with the national and international standards of research ethics and integrity, while its potential results appear unlikely to have drastic disruptive implications for society. To understand why there is a need for a governance framework to align biomedicine to society it is worth looking at its possible ‘transformative’ potential.

Much of contemporary biomedical research is characterised by its increasing use of advanced bio-banking techniques, AI-based diagnostic tools, data analysis software, and similar technologies. As an example of why this kind of research is regarded as potentially transformative, I will mention the potential implications of those biomedical projects that aim at the design and development of AI-based diagnostic tools such as mechanistic models capable of forecasting the effects and effectiveness of drug therapies on patients, as well as the overall course of their disease. Based on the analysis of patients’ data, these models can help to identify the optimal drug combinations, dosing, and scheduling for each individual patient (Barbolosi et al 2016). In this way, expensive drugs and precious time will be saved and, in theory, patients will avoid ineffective treatments and adverse side effects.

Apart from the obvious potential benefits of this kind of research, there are some complex and non-easily quantifiable implications to consider and that go beyond the widely discussed general problem of the undesirable effects of potentially biassed algorithms in AI.

AI-based machine learning tools in healthcare come with the promise of increased precision. However, some fear the risk of an *overuse* of biobanking techniques and laboratory tests (Brownlee et al. 2017), that can easily lead to an increased number of false-positives and, therefore, of *overmedicalization* and *overdiagnosis* (Redberg 2016; Vogt et al 2019). Moreover, while there is already awareness of the potential fallacies of human experts, on the one hand, and of the potential issues surrounding AI devices, on the other, there is still little evidence that the 'humans-*plus*-AI' complex distributed decision-making system is better than a fully human clinical panel. The introduction of AI-tools in clinical panels, in fact, may not be a summative process, in which the precision of the technological devices is simply added to human expertise. Some empirical evidence indeed suggests that the interactions between AI and clinical experts may actually affect clinicians in a negative way, for instance by reducing their diagnostic skills (see, for example, Hoff 2011; Povyakalo et al 2013). Alongside the risks of overmedicalization and overdiagnosis, therefore, some fear that too much reliance on AI may lead to a sort of cognitive and decisional 'laziness', that would increase the risk of *less careful diagnosis* (Keil 2017).

It should not be forgotten, moreover, that AI-tools are introduced in healthcare in order to *support* clinical reasoning, not to *determine* clinical decisions (van Baalen, Boon and Verhoef 2019). Especially in cases where clinical experts' opinions diverge from the AI-based predictions, the actual 'weight' of the innovative devices in the clinical reasoning process remains unclear. There is the possibility that the disagreements between clinicians and machine learning devices may generate an increase in the demand of 'second opinions', leading to lengthier clinical paths as well as to increased confusion and anxiety among patients (Cabitza 2019).

Innovative healthcare technology may also have unexpected impacts on patients' experience of their own illness and therapy. Empirical studies suggest that patients may mistrust AI-based diagnostic tools, or that they may perceive a loss in their own autonomy as well as in the 'humanity' of the physician-patient relationship (Esmaeilzadeh 2020). Finally, future research may end up developing highly precise 'next-generation' healthcare AI-tools, which patients could use by themselves, outside the clinical setting, and without the mediation of a physician, with unexpected cultural and societal effects.

The organisational and social impacts of innovative AI-based tools in healthcare, briefly mentioned in this section, cannot be assessed only by focussing on the problem of biassed algorithms, nor are they easily quantifiable. Since these issues are not 'technical problems', they do not have a merely 'technical' resolution. These issues, in other words, are not about the design of optimal machine learning devices but, rather, about how such devices will be deployed and which kind of transformation in clinical practice and in patients' personal experience they will bring about. Because of the potential negative effects of an over reliance on AI, the 'best' innovative technological tools may not be 'good' for healthcare after all. This is also the reason why all these issues should not be reflected upon by AI engineers alone, but also by the biomedical researchers involved in projects that aim at the development of innovative healthcare tools and treatments.

In short, although its aims seem to be straightforwardly socially desirable, innovative biomedical research has the power of modifying clinical reasoning, clinical practice, the healthcare system, the physician-patient relation, and even the self-perception of one's own illness and therapy, in ways that are hard to predict. Biomedical researchers' responsibility towards society, therefore, goes beyond the prevention of easily identifiable and quantifiable harms. Avoiding maleficence does not guarantee beneficence: solving the well-known technical issues with AI does not guarantee that a technically good AI-based healthcare tool will have the right impact. In the light of these considerations, the requirement for biomedical research projects to comply with the new policy frameworks appears plausible.

4. Disciplinary moral identities

In the previous sections I explained how the co-production ideal was developed as a reaction against the Value-Free Ideal; and I have also talked about the potential social impacts that motivate the compliance of biomedical research to the policy frameworks inspired by such an ideal. Often the discussion about these issues is conducted at the very general level of the different ideals of the science-society relation. Two crucial points, however, should be taken into consideration.

First, between the ideal of an absolutely value-free science and the ideal of a value-laden science perfectly aligned with the needs of society, there exists a varied spectrum of research traditions. When it comes to actual rather than ideal practices, value-freedom and value-ladenness may well be a matter of degrees. This is not problematic in itself. That research in theoretical mathematics may be value-free, for example, does not seem to pose particular concerns. By contrast, a reflection on societal impacts of science and technology may be endemic and drive research in more innovative and applied fields. There are even cases of fields that, by expanding their range of application, may *become* more or less value-laden. For example, the once purely theoretical quantum mechanics is now getting applied in the development of new technologies, such as the so-called ‘quantum computer’, which may have a transformational impact on society. While the general discussion of the value-free and the co-production ideals is an important entry point to tackle some practical issues, the heterogeneous plurality of actual disciplinary and research traditions must not be forgotten, especially when it comes to the development and application of science policies and governance framework.

Second, not only is it important to distinguish between different degrees of value-freedom of the sciences, but also of the *scientists*. The much contested Value-Free Ideal is a prescriptive thesis: it states that non-epistemic values must not play any role in the justification of scientific claims. It is not a descriptive thesis: that is, it does not state that actual scientists, as a matter of fact, never make value-judgments. Even if it were found out that actual scientists *do* make value-judgments, supporters of the Value-Free Ideal could still argue that they *ought not* to do so. At the same time, however, the policy strategies that aim at implementing a sense of social responsibility in research seem to presuppose that, indeed, researchers are not already engaged in such a reflection. They seem to presuppose, in other words, that not only the sciences but also the scientists are, or at least tend to be, value-free. In the same way in which they consider the general public to have a ‘knowledge deficit’, in other words, several science policies and governance frameworks seem to regard scientists as characterised by a ‘moral deficit’; often, however, both the public and the scientists these policies are built for are strong idealizations, ‘imagined’ rather then real individuals (see Åm et al 2021).

Some empirical evidence suggests that, in some fields, researchers carry on their activity in a rather ‘value-free’ fashion, without reflecting upon the societal implications of their work. A recent ethnographic study conducted on a large engineering lab in the US has concluded that engineers do not think about the societal implications of their work, neither spontaneously nor after they have been explicitly asked to do so (Lee et al 2019). At the same time, it is important to specify that there can be many different reasons for why researchers do not engage in critical reflection in an explicit or systematic way. Many of these reasons do not have to do with researchers’ inability to be reflective or with an indifference towards society. As pointed out by Wolpe (2006), the reasons for scientists’ lack of explicit engagement with socio-ethical issues may range from feeling inadequate to tackle them, to the belief that it is not their duty to do so, to the misrepresentation of ‘ethics’ as a constraint to research (i.e., a set of rules and prohibitions) rather than as something that may positively drive it. Rather than just assuming that STEM or Medical researchers need an 'external' help to engage in critical reflexivity, it should be wondered whether and how much they do so already, especially when it comes to research areas that have never been regarded as value-free to begin with.

The two points are indeed connected. If value-freeness is a matter of degrees, it may well be the case that the members of some research communities are *already* more engaged with moral and social considerations than others (whether they acknowledge the value-ladenness of their way of reasoning or not). Depending on the field they are working on and on the research tradition they have been trained to, that is, scientists from different disciplines may possess different *disciplinary moral identities*.

Osbeck and Nersessian (2017) have recently examined the idea that, in virtue of their membership to a scientific community, in turn acquired through a lengthy special training, scientists from different disciplines may possess a different disciplinary identity. Osbeck and Nersessian examine how different disciplinary identities may pose an obstacle to interdisciplinary research. Following their argument, "in interdisciplinary science there are often multiple (sometimes conflicting) norms and values – differing ideas about what constitutes good science" (Osbeck and Nersessian 2017:233). Their focus, however, is mainly on epistemic issues. In fact, Osbeck and Nersessian speak about scientists' disciplinary identities in terms of *epistemic* identities. Following on Osbeck and Nersessian's study, and expanding upon their perspective, I will examine how different disciplinary *moral* identities, which depend on the degree of value-ladenness of different research traditions, may pose some problems to the attempt of integrating a sense of social responsibility in science.

Before discussing how disciplinary moral identities could play a role in the implementation of social responsibility in biomedical research, a few more specifications about the concept of identity are necessary. 'Identity' is an irreducibly complex concept and it may refer to different levels, such as the individual and the collective level. For example, 'personal identity' depends on individuals' subjective experiences and it is a concept employed in psychology. 'Group identities', by contrast, are sociological and political concepts. These two levels, however, are often intertwined. For instance, when we say that "Mary is a lawyer", we are talking about Mary's personal identity as a practising expert in the field of jurisprudence, but such an identity is provided by her membership to a professional group. The 'sense of one's own self', in other words, often depends on one's relations with the others and with a social and cultural context, which in parts shapes individuals. At the same time, shifts in individuals' awareness and preferences may lead to social and cultural changes. For these reasons, identity could (and perhaps should) be considered as intersecting the individual/psychological and the collective/social level (Tajfel and Turner 1986; Turner and Onorato 1999).

Because of such a complex interplay of individual and collective levels of identity, it is important to stress that different individuals belonging to the same group, and therefore shaped by the same collective identity, may nevertheless behave and experience themselves in different ways. For example, when in the next section I will say that biomedical researchers may already have a strong disciplinary moral identity, I will not mean that every single biomedical researcher is driven by the same moral commitments towards humans. Many biomedical researchers may well be motivated by rather selfish reasons (such as the desire of recognition), or may be simply indifferent to actual patients. The reasons and motivations of *individual* biomedical researchers notwithstanding, however, biomedical research, as a *collective* enterprise, has its own epistemic objectives (i.e., 'discovering the *right* cure for x') that are also regarded as morally good and socially desirable (i.e., 'finding the cure for x is *good*').

It is also important to stress that, although collective identity plays a role in shaping individual identity, it may do so in ways that may defy one's own self-representation. Individuals belonging to some groups may misunderstand or misrepresent some of the characteristics of their group's identity. In every political dispute, for example, members of different parties will consider themselves to be on 'the right side'. Similarly, when it comes to disciplinary identities, individual scientists may simply fail to notice some of their own characteristic traits. In the case of biomedicine, researchers may even regard themselves as reasoning in a value-free fashion (that is, they regard themselves as being concerned solely with the discovery of 'pure facts'), without noticing the value-judgments they employ in their reasoning. Instances of this issue have been indeed discussed (see, for example, de Melo-Martín, Kristen Intemann 2012). In the same way in which the disciplinary epistemic identity discussed by Osbeck and Nersessian (2017) may be a 'tacit' and non-explicit perspective endorsed by scientists, so their disciplinary moral identity may guide their choices without their full awareness.

5. Interdisciplinarity for responsibility: integration, delegation, or moral reinforcement?

Implementing a sense of social responsibility in science and technology requires structural and cultural changes in research conducting institutions. Some of these changes may simply be ‘mechanical’, targeting some institutional procedures. This is the case, for example, of the formal requirement of improving gender equality and the representation of marginalised social groups in research teams through specific hiring mechanisms (Leuschner 2015). The requirements of publishing Open Access and of finding new avenues for dissemination (such as blogging, or writing popular articles for the general public) also count as rather ‘mechanical’ and structural changes. These are requirements that simply demand researchers to follow new rules, without aiming at modifying how they conduct research, how they value it, and how they perceive their own role in society. Cultural changes, however, are more difficult to make. There are different strategies to obtain such changes ‘from the inside’, that is by targeting the midstream stages of research. One of such strategies consists in encouraging *interdisciplinary collaborations* between lab researchers and scholars from the Humanities.

Interdisciplinarity is believed “to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice” (NAS 2005:2). There exists a vast philosophical and STS literature about the value of interdisciplinary research for the resolution of complex ‘real world’ problems. With the emergence of the co-production model and of several new science policies, some scholars have begun to consider not only the problem-solving power of interdisciplinarity, but also its potential to bring value diversity in research and make it more socially responsible, open, and reflexive as a consequence. The interaction between STEM and Medical researchers, on the one hand, and humanists, on the other, is believed to bring a variety of moral values and social perspectives into innovative projects.

Since ‘health’ and related concepts are inherently complex and require more than a reductively biological explanation, some even argue a form of ‘radical’ interdisciplinarity between biomedical researchers and humanists (Clarke, Ghiara and Russo 2019). In particular, it has been suggested that interdisciplinary biomedical research teams should include *in situ* ethicists who would analyse the moral and social ‘embeddedness’ of some research practices and even provide an active ‘ethical guide’ to the researchers in the lab (Gardner and Williams 2015; Jongsma and Bredenoord 2020). These proposals have been advanced in the wake of the so-called ‘empirical turn’ in bioethics and in clinical research ethics (Borry, Schotsmans and Dierick 2004, 2005), the aim of which is to draw upon empirical data and observations of research practices in order to improve the context sensitivity of normative bio-ethical claims (Musschenga 2005).

As a strategy for the institutionalisation of social responsibility, however, interdisciplinarity is not problem free. By itself, the establishment of interdisciplinary research teams does not guarantee that social awareness and reflexivity will successfully be integrated in the internal stage of research. Within interdisciplinary research groups, in fact, there is the concrete possibility that scientists and engineers will keep regarding socio-ethical issues as something that their humanist colleagues will have to look after. Even when working on projects that comply with policies such as ELSI or RRI, and which require the participation of ‘embedded’ humanists or social scientists, many STEM researchers have still the tendency to regard the reflection on the societal impact of their work as ‘extraordinary’, that is not common, not expected (Hennen, Hahn and Ladikas 2021). STEM researchers, in other words, may simply pay lip-service to the science policy by hiring some sociologist or philosopher “to look after the social responsibility part” of their project. In this way, awareness and reflexivity do not really get implemented in the internal stage of research. As some scholars have already noticed, in other words, interdisciplinarity by itself may easily lead to the *delegation* of social responsibility to SSH scholars, rather than its integration in research (Delgado and Åm 2018; Sigl et al 2020).

Granted that there are many problems with the design of interdisciplinary teams for the institutionalisation of social responsibility and reflexivity in research, I want to focus on the underlying assumptions behind the idea that interdisciplinary interactions may foster social responsibility. The assumption, that is, that researchers do not already engage in the kind of critical and ethical reflection that interdisciplinary collaborations with scholars from the Humanities would supposedly elicit. As already discussed in the previous section, this way of distinguishing between ‘unreflective’ lab researchers and humanists with a ‘moral expertise’, on the other hand, is far too simplistic.

Biomedicine and the health sciences are ‘not-so-value-free’ fields. Within the value-free/value-laden spectrum, that is, biomedicine and the health sciences appear to be already oriented towards socially desirable ends and often engaged with at least part of other societal actors and stakeholders (such as patients’ organisations, pharmaceutical companies, and so on). The production of reliable biomedical knowledge for therapeutic ends may indeed be regarded as *both* an epistemic *and* a moral objective. In other words, "[the] medical and health sciences are [...] underwritten by moral and social values having to do with easing human suffering and preventing premature death. These values influence decisions – in the early stages of scientific inquiry – about which research projects ought to be pursued, and they also play a role in decisions about how biomedical research ought to be conducted" (Rooney 2017:34). One may even go as far as arguing that the inextricability of epistemic and moral aspects in biomedical research influences not only researchers' decision-making process, but also their own self-perception: "[m] any biomedical researchers describe their work as the search for *Truth*. Their scientific quest to understand the mysteries of the human body and processes of disease and healing demands a personal commitment that makes a life of biomedical research a *moral calling*" (Heitman 2000:S41).

To summarise this argument: biomedicine and the health sciences are in themselves value-laden, therefore researchers in these fields make use of value-judgements (even more and more often than researchers in other fields), work towards socially desirable ends, reflect upon the impact of their work on others, and may even regard themselves as morally superior, or at least 'just'. This amounts to saying that biomedical researchers, in virtue of the value-ladenness of their field, possess a pretty strong disciplinary moral identity.

There are not many empirical studies, if at all, directly addressing the issue of the disciplinary moral identity of scientists, in general, and of biomedical researchers, in particular. There is however some empirical evidence that suggests that biomedical researchers do indeed regard themselves as both epistemic and moral agents, sometimes even displaying some sort of 'moral superiority'. Antes et al. (2010) have studied how many biomedical researchers who took special ethical training modules became so overconfident to the point of developing a sort of 'self-enhancement bias'; as a result, it was observed a decrease in their help-seeking behaviour. After a 2-year quantitative study, McCormick et al (2012) also concluded that self-confidence in the strength of their own moral intuitions is actually a barrier to a deeper engagement to reflecting critically about the socio-ethical implications of their work among biomedical researchers.

That 'moral training' may reinforce the sense of moral superiority, or even 'infallibility', of biomedical researchers, as suggested by these studies, of course does not imply that the same problematic effect will be inevitably obtained in interdisciplinary collaborations. However, the implication of the existence of different disciplinary identities ought to be taken seriously. In their interdisciplinary interaction with *in situ* bioethicists, for instance, biomedical researchers may end up reinforcing their commitment to the values of traditional bioethics (that is, autonomy, beneficence, nonmaleficence, and justice; Beauchamp and Childress 1979) that drive their research, shape their objectives and, therefore, is part of their disciplinary moral identity. The problem is that these fundamental bioethical values cover only partially the complex ‘moral terrain’ of biomedical research, the discussion of which is too often modelled after the debates about the clinician-patient relations and it is therefore conceived almost exclusively as ‘therapy-oriented’ (Joffe and Miller 2012). Such an approach could not be helpful in considering the kind of socially transformative impacts that innovative biomedical research may have, such as those mentioned in the third section. Actually, that biomedical research cannot be reduced to finding the best treatments or therapies, without any reflection about their wide-ranging implications, is what motivates some policy and governance frameworks. However, if biomedical researchers have the belief that their only objectives are indeed finding the best treatments or therapies, and that such objectives are also morally good, and if such a belief is reinforced through moral training or interdisciplinary interactions with bioethicists, then their disciplinary moral identity may actually be an obstacle to their development of a wider and more nuanced sense of social responsibility.

Finally, it must be stressed that current innovative biomedical research is *already* a highly interdisciplinary enterprise. Apart from interacting with scholars from the Humanities and the Social Sciences to become more socially responsible, biomedical researchers collaborate with biostatisticians, bioinformatics, computer modellers, bioengineers, and so on. They collaborate, that is, with researchers with different disciplinary identities. As mentioned in the previous section, Osbeck and Nersessian (2017) have already discussed the epistemic issues arising from the conflict of different disciplinary identities in interdisciplinary research. There can be, however, also moral issues. To paraphrase Osbeck and Nersessian: in interdisciplinary science there are often multiple (sometimes conflicting) norms and values – differing ideas about what constitutes *socially responsible* science. The issue is of course exacerbated if the interaction with bioethicists may reinforce the moral perspective of biomedical researchers, putting them in a position of ‘moral authority’ within the interdisciplinary group. Once again, the problem is not that biomedical researchers, in this case, would be ‘wrong’ or ‘bad’, but only that their perspective is not sufficient to tackle the wide ranging societal implications of their research.

The problem, in short, is not so much that biomedical researchers do not make value-judgments, but that they do not make the right value-judgments necessary to the kind of social responsibility required by current policy frameworks and by the co-production ideal.

6. Patient engagement, public engagement, and co-production

Since the implementation of science and technology in society is a process that involves different agents forming a complex network, and since the long-term transformative societal impacts of research are difficult to predict, the co-production model recognizes the importance of integrating a plurality of stances and perspectives. In this model, scientists and other societal actors ought to collaborate in order to produce socially desirable results together. The co-production ideal, in other words, puts a major emphasis on the 'openness', 'inclusiveness', and 'democratisation' of responsible research (Eigi 2017; Jasanoff 2017; Maassen and Weingart 2005). For this reason, the new policy frameworks require researchers to engage with different 'stakeholders' as well as with 'citizens'.

Determining what these terms actually refer to is often problematic (Bensaude-Vincent 2014; Woolley 2016). Nor is it always clear *how* public engagement should be conducted in practice (i.e., at which stages of research, by which members of a research group, targeting which part of the 'public'; see Ivani and Dutilh Novaes 2022). Moreover, it is becoming questionable that public engagement actually serves the function of making research socially responsible. Some fear, for example, that researchers would not become more open to social concerns and to the plurality of needs and perspectives by engaging with the public. Rather, they would use public engagement to gain and reinforce consensus (Wynne 2006; Thorpe and Gregory 2010; Lezaun and Soneryd 2007).

These well known issues with public engagement get a further layer of complexity in the case of biomedicine. *Prima facie*, it may seem that biomedical researchers already 'engage with the public', thus satisfying one of the requirements for socially responsible research. Especially in so-called 'translational medicine', researchers *already* seek the input, involvement, and collaboration of patients and patients' organisations (Woolf 2008; van der Scheer 2017). 'Public engagement', however, means more than 'engagement with the patients'. It includes, in fact, engagement with many different societal actors. What remains unclear, then, is how biomedical researchers should prioritise, balance, and integrate the views of different societal actors (that is, patients and non-patients) into research.

Patients' health and well-being is the primary objective of biomedicine. From a biomedical perspective, would it be (epistemically and morally) right to weigh non-patients' views and inputs as much as the views of people affected by an illness, whose well-being or even lives may depend on the particular directions research will take? If patients' views should always count more than the other stakeholders' views, why should biomedical researchers engage with other stakeholders to begin with?

Dilemmas like this arise from the conceptualization of responsibility in terms of 'being responsive' to the public (Stilgoe, Owen and Macnaghten 2013). In the co-production model of science, in fact, researchers *together* with other societal actors are 'co-responsible' for determining the direction of scientific research and for the implementation of its results into society. What the 'inclusive' and 'participatory' model does not indicate, however, is how to resolve conflicts among the different views of the various stakeholders. Furthermore, private stakeholders *as well as* the public may be motivated by their own interests, they may have their own biases and, therefore, they may not contribute to scientific development just for the 'common good' (Block and Lemmens 2015). The demand of 'being responsive', but without an explanation of how researchers should respond to the competing or even conflicting interests and needs of different societal groups, is a too vague a concept to be operationalized and implemented at the midstream stage of research.

The problem is exacerbated in cases in which scientists engage with other stakeholders from a strong moral standpoint of their own: in cases, that is, in which researchers already possess a pretty strong disciplinary moral identity. As discussed in previous sections, while some scientists may consider their activity as rather 'value free', and to regard the production of a pure and unadulterated 'knowledge about facts' as their primary duty, a critical reflection about the potential implications of research is actually endemic in some scientific communities. As a consequence of the 'moral reinforcement' they have received through, for example, interdisciplinary collaborations with empirical bioethicists, biomedical researchers may end up imposing their (and their patients') views on the rest of the public. In other words, the moral reinforcement of biomedical researchers may end up increasing their disagreements with the other stakeholders, in situations where different stakeholders (for example, patients and non-patients) may already have different priorities and disagree with each other.

One may respond by saying that, from a biomedical perspective, research is always therapy oriented and that, therefore, patients' needs and views ought always to be prioritised. Such a response would be an invitation to "leave everything as it is" in biomedical research that already seeks the involvement of expert patient panels and patients' associations. As discussed earlier on, however, innovative biomedical research may have a variety of impacts, and not only on patients. It may transform the clinical practice and the structure of clinical panels, thus leading to more expenses in countries with a national health system; it may even lead to the production of ‘portable’ digital devices that may reshape society in unpredictable ways. Even assuming that patients’ needs must be prioritised, however, there still remains ample room for opening biomedical research to the inputs coming from other (non-patients) stakeholders. Once again, it remains unclear how to integrate such a plurality of views, values and stances in a kind of research that is traditionally defined as mainly (or even exclusively) patient-oriented.

7. Concluding remarks

Like many other contributions to the literature on the potential barriers to the institutionalisation of social responsibility in research, this is mainly a theoretical and speculative article. Its main point is to consider the possibility that, especially in the case of value-laden and socially oriented research fields, such as biomedicine, the proposed measures to make research socially responsible may lead to unexpected or even unwanted consequences. At the same time, the article encourages further social and psychological investigations into the disciplinary moral identity of biomedical researchers to complement the current literature in responsible biomedical research.

References

- Åm Heidrun, Solbu Gisle, Sørensen Knut Holtan (2021). The imagined scientist of science governance. *Social Studies of Science* 51:277–297.
- Antes Alison, Xiaoqian Wang, Michael Mumford et al 2010. “Evaluating the Effects That Existing Instruction on Responsible Conduct of Research Has on Ethical Decision Making.” *Academic Medicine* 85:519-526.
- Barbolosi Dominique, Joseph Ciccolini, Bruno Lacarelle, et al. 2016. “Computational oncology: mathematical modelling of drug regimens for precision medicine.” *Nature Review Clinical Oncology* 13: 242–254.
- Beauchamp Tom and James Childress. 1979. *Principles of Biomedical Ethics.* Oxford: Oxford University Press.
- Bensaude-Vincent Bernadette. 2014. “The politics of buzzwords at the interface of technoscience, market and society: the case of ‘public engagement in science.” *Public Understanding of Science* 23: 238–253.

- Block Vincent and Lemmens Pieter. 2015. “The emerging concept of responsible innovation. Three reasons why it is questionable and calls for a radical transformation of the concept of innovation.” In *Responsible Innovation 2: Concepts, Approaches, and Applications*, Koops B. (Ed.) Berlin: Springer, pp. 19-35.
- Borry Pascal, Paul Schotsmans, and Kris Dierickx. 2004. “What is the role of empirical research in bioethical reflection and decision-making? An ethical analysis.” *Medicine Healthcare Care and Philosophy* 7: 41–53.
- Borry Pascal, Paul Schotsmans, and Kris Dierickx. 2005. “The birth of the empirical turn in bioethics”, *Bioethics* 19: 49–71.
- Brownlee Shannon, Kalipso Chalkidou, Jenny Doust, et al.2017. “Evidence for overuse of medical services around the world.” *The Lancet* 390:156–168.
- Cabitza Federico. 2019. “Biases affecting human decision making in AI-supported second opinion settings.” In Torra V., Y. Narukawa, G. Pasi, and M. Viviani (Eds.) *Modelling Decisions for Artificial Intelligence*. London: Springer, pp. 283–294.
- Cabitza Federico, Raffaele Rasoini and Gian Franco Gensini. 2017. Unintended consequences of machine learning in medicine. *JAMA: The Journal of the American Medical Association* 318: 517–518.
- Clarke Brendan, Virginia Ghiara and Federica Russo. 2019. “Time to care: why the humanities and the social sciences belong in the science of health.” *BMJ Open* **9**: e030286. doi: 10.1136/bmjopen-2019-030286
- Delgado Ana, and Heidrum Åm. 2018. “Experiments in interdisciplinarity: Responsible research and innovation and the public good”, *PLoS Biol*ogy 16.
- de Melo-Martín Immaculada, and Kristen Intemann (2012) Interpreting evidence: why values can matter as much as science. *Perspectives in Biology and Medicine* 55:59–70.
- Deming Nicole, Kelly Fryer-Edwards, Denise Dudzinski et al. 2007. “Incorporating Principles and Practical Wisdom in Research Ethics Education: A Preliminary Study.” *Academic Medicine* 82:18–23.
- deSaille Stevienna, and Medvecky Fabien. 2016. “Innovation for a steady state: a case for responsible stagnation.” *Economy and Society* 45:1–23.
- Douglas Heather. 2009. *Science, Policy, and the Value-Free Ideal*, Pittsburgh: Pittsburgh University Press.
- Elliott Kevin and Daniel McKaughan. 2014. “Non-epistemic values and the multiple goals of science.” *Philosophy of Science* 81: 1-21.

- Esmaeilzadeh Pouyan. 2020. “Use of AI-based tools for healthcare purposes: a survey study from consumers’ perspectives.” *BMC Medical Informatics and Decision Making*: 17.
- Fisher Erik, Roop Mahajan, and Carl Mitcham. 2006. “Midstream modulation of technology: governance from within.” *Bulletin of Science Technology and Society* 26: 485-496.
- Fisher Erik and Daan Schuurbiers D. 2013. “Socio-technical Integration Research: collaborative inquiry at the midstream of research and development”. In Doorn N., D. Schuurbiers, I. van de Poel, M. Gorman (Eds.) *Early engagement and new technologies: Opening up the laboratory.* Dordrecht: Springer, pp. 97-110.
- Fisher Erik, Michael O’Rourke, Robert Evans, et al. 2015. “Mapping the integrative field:
- taking stock of socio-technical collaboration.” *Journal of Responsible Innovation* 2: 39–61.
- Forssén Annika, Eivind Meland, Irene Hetlevik, et al. 2011. “Rethinking scientific responsibility.” *Journal of Medical Ethics* 37:299–302.
- Funtowicz Silvio, Jerome Ravetz. 1993. “Science for the post-normal age,” *Futures* 31: 735–755.
- Gardner John, and Clare Williams. 2015. “Responsible research and innovation: a manifesto for empirical ethics?”, *Clinical Ethics* 10: 5–12.
- Gibbons Michael, Camille Limoges, Helga Nowotny, et al. 1994. *The New Production of Knowledge: the dynamics of science and research in contemporary societies*, London:
- SAGE.
- Grunwald Armin. 2018. *Technology assessment in practice and theory.* London: Routledge.
- Heitman Elizabeth. 2000. “Ethical values in the education of biomedical researchers”, *Hastings Center Report* 30: S40-S44.
- Hennen Leonhard, Julia Hahn, and Miklos Ladikas. 2021. “Competing modes of responsibility in research organisations: insights from an international comparison.” *Science and Public Policy* 48:54–65.
- Hoff Timothy. 2011. “Deskilling and adaptation among primary care physicians using two work innovations.” *Health Care Management Review* 36:338–348
- Iaccarino Maurizio. 2006. “Science and ethics”, *EMBO report* 2: 747–750.
- Ivani Silvia and Catarina Dutilh Novaes. 2022. “Public engagement and argumentation in science.” *European Journal for Philosophy of Science* 12:54.
- Joffe Steven and Franklin Miller. 2008. Bench to bedside: mapping the moral terrain of clinical research. *The Hastings Center Report* 38:30–42.

- Jasanoff Sheila. 2004. *States of Knowledge: the co-production of science and social order*, London: Routledge.
- Jasanoff Sheila. 2016. *The Ethics of Invention: technology and the human future*, New York: W. W. Norton.
- Jongsma Karin, and Annelien Bredenoord. 2020. "Ethics parallel research: an approach for (early) ethical guidance of biomedical innovation", *BMC Medical Ethics* 21:81–88.
- Keil Ode. 2017. 2017. "Will extensive use of artificial intelligence degrade human intelligence?" Journal of Clinical Engineering 42:161.
- Kincaid Harold, John Dupré, and Alison Wylie. 2007. *Value-Free Science: ideals and illusions?* Oxford: Oxford University Press.
- Kitcher Philip. 2001. *Science, Truth, and Democracy*. Oxford: Oxford University Press.
- Kitcher Philip. 2011. *Science in a Democratic Society.* New York: Prometheus.
- Koskinen Inkeri. 2020. "Defending a Risk Account of Scientific Objectivity". *British Journal for the Philosophy of Science* 71: 1187-1207.
- Koskinen Inkeri. 2022. "How institutional solutions meant to increase diversity in science fail." *Synthese* 200:483.
- Kourany Janet. 2010. *Philosophy of Science after Feminism.* Oxford: Oxford University
- Press.
- Kuhn Thomas. 1977. Objectivity, Value Judgement and Theory Choice: In Kuhn Thomas, *The Essential Tension: selected study in scientific tradition and change.* Chicago:
- Chicago University Press, pp. 320-339.
- Lacey Hugh. 1999. *Is Science Value-Free? Values and scientific understanding.* London:
- Routledge.
- Laudan Larry. 1984. *Science and Values.* Los Angeles: California University Press.
- Leuschner Anna. 2015. "Social exclusion in academia through biases in methodological quality evaluation: on the situation of women in science and philosophy." *Studies in History and Philosophy of Science* 54:56-63.
- Longino Helen. 1990. *Science as social knowledge: values and objectivity in scientific inquiry.* Princeton: Princeton University Press.
- Longino Helen. 2002. *The Fate of Knowledge.* Princeton: Princeton University Press.
- Lee Eun Ah, Nicholas Gans, Magdalena Grohman, et al., "Ethics as a rare bird: a challenge for situated studies of ethics in the engineering lab." *Journal of Responsible Innovation* 6:284–304.

- Lehoux Pascale, Hudson Pacifico Silva, Renata Pozelli Sabio, et al. 2018. "The Unexplored Contribution of Responsible Innovation in Health to Sustainable Development Goals". *Sustainability* 10:4015.
- Lehoux Pascale, Federico Roncarolo, Hudson Pacifico Silva, et al. 2019. "What Health System Challenges Should Responsible Innovation in Health Address? Insights From an International Scoping Review." *International journal of health policy and management* 8:63–75.
- Lezaun Javier and Linda Soneryd. 2007. "Consulting citizens: Technologies of elicitation and the mobility of publics." *Public Understanding of Science* 16:279–297.
- MacNaghten Philip. 2021. *The Making of Responsible Innovation*, Oxford: Oxford University Press.
- McMullin Ernan. 1983. "Values in Science". In Asquith Peter and Thomas Nickles, (Eds.).
- *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1982*, Volume 2: 3-28.
- Musschenga Albert. 2005. "Empirical ethics, context-sensitivity, and contextualism." *The Journal of Medicine and Philosophy* 30:467–490.
- Osbeck Lisa and Nancy Nersessian. 2017. Epistemic identities in interdisciplinary science. *Perspectives on Science* 25: 226–260.
- Owen Richard, Jack Stilgoe, Phil Macnaghten et al. 2013. "A framework for Responsible Innovation". In Owen R., J. Bessant, and M. Heintz, M. (Eds.) *Responsible Innovation*. New York: John Wiley, pp. 27–50.
- Pacifico Silva Hudson, Pascale Lehoux, Fiona Alice Miller, *et al.* 2018. "Introducing responsible innovation in health: a policy-oriented framework." *Health Research Policy and Systems* 16:90.
- Polanyi Michael. 1962. "The Republic of Science: Its Political and Economic Theory", *Minerva* 1:54–74.
- Povyakalo Andrey, Eugenio Alberdi, Lorenzo Strigini et al. 2013. "How to Discriminate between Computer-Aided and Computer-Hindered Decisions: A Case Study in Mammography." *Medical Decision Making* 33:98–107.
- Redberg Rita. 2016. "The medicalization of common conditions." *JAMA Internal Medicine* 176:1863–1863.
- Rooney Phyllis. 1992. "On Values in Science: Is the Epistemic/Non-Epistemic Distinction Useful?" *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1992:13–22.
- Rooney Phyllis. 2017. The borderland between epistemic and non-epistemic values. In Elliott K. and D. Steel (Eds.) *Current Controversies in Values and Science*. Routledge, pp. 31–45.

- Sigl Lisa, Ulrike Felt, and Maximilian Fochler. 2020. "'I am Primarily Paid for Publishing...': The narrative framing of societal responsibilities in academic life science research." *Science and Engineering Ethics* 26:1569–1593.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *The psychology of intergroup behavior* (pp. 7–24).
- Chicago: Nelson Hall.
- Turner, J. C., & Onorato, R. S. (1999). Social identity, personality, and the self-concept. In T. R. Tyler, R. M. Kramer, & O. P. John (Eds.), *The psychology of the social self* (pp. 11–46). Mahwah, NJ: Lawrence Erlbaum Associates.
- Thorpe Charles and Jane Gregory. 2010. "Producing the post-Fordist public: The political economy of public engagement with science." *Science as Culture* 19:273–301.
- van Baalen Sophie, Mieke Boon, Petra Verhoef. 2021."From clinical decision support to clinical reasoning support systems." *Journal of Evaluation in Clinical Practice* 28: 520–528.
- van der Bug Simone. 2009. "Taking the 'soft impacts' of technology into account: broadening the discourse in research practice." *Social Epistemology* 23: 301–16.
- van der Scheer Lieke, Elisa Garcia, Anna Laura van der Laan, et al. 2017. "The Benefits of Patient Involvement for Translational Research." *Health Care Analysis* 25: 225–241.
- Vogt Henrik, Sara Green, Claus Ekstrøm, et al. 2019. "How precision medicine and screening with big data could increase overdiagnosis." *BMJ* 13, doi: 10.1136/bmj.l5270, 2019.
- Wynne Brian. 2006. "Public engagement as a means of restoring public trust in science: hitting the notes, but missing the music?" *Community Genetics* 9:211–220.
- Woolf Steven. 2008. "The meaning of translational research and why it matters." *JAMA* 299: 211-213.
- Woolley Patrick, Michelle McGowan, Harriette Teare, et al. 2016. "Citizen science or scientific citizenship? Disentangling the uses of public engagement rhetoric in national research initiatives." *BMC Medical Ethics* 17.
- Wolpe Paul. 2006. "Reasons scientists avoid thinking about ethics." *Cell* 125: 1023–1025.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.