

Neuroethics and Philosophy in Responsible Research and Innovation: The Case of the Human Brain Project

Arleen Salles  · Kathinka Evers · Michele Farisco

Received: 16 March 2018 / Accepted: 27 May 2018
© The Author(s) 2018

Abstract Responsible Research and Innovation (RRI) is an important ethical, legal, and political theme for the European Commission. Although variously defined, it is generally understood as an interactive process that engages social actors, researchers, and innovators who must be mutually responsive and work towards the ethical permissibility of the relevant research and its products. The framework of RRI calls for contextually addressing not just research and innovation impact but also the background research process, specially the societal visions underlying it and the norms and priorities that shape scientific agendas. This requires the integration of anticipatory, inclusive, and responsive dimensions, and the nurturing of a certain type of reflexivity among a variety of stakeholders, from scientists to funders. In this paper, we do not address potential limitations but focus on the potential contribution of *philosophical reflection* to RRI in the context of the Ethics and Society subproject of the Human Brain Project (HBP). We show how the type of conceptual analysis provided by philosophically oriented approaches theoretically and ethically broadens research and innovation

within the HBP. We further suggest that overt inclusion of philosophical reflection can promote the aims and objectives of RRI.

Keywords Responsible research and innovation · Reflexivity · Conceptual analysis · Neuroethics · Human identity · Consciousness · Poverty

Background

Responsible Research and Innovation (RRI) is an important ethical, legal, and political theme for the European Commission. Although variously defined, it is generally understood as an interactive process that engages social actors, researchers, and innovators who must be mutually responsive and work towards the ethical permissibility of the relevant research and its products [1]. RRI arose in response to the pace of technological and scientific research and related applications. With the prospect of new discoveries came an increasing awareness of the profound global and inter-generational impact of innovations, and of the limits of any policy that focuses just on risk assessment and regulation. [2, 3]¹

Of course, the view that research needs to be responsible is hardly new; the importance of acknowledging one's responsibility (legal and moral) is and has long been implicit within the description of many roles,

A. Salles (✉) · K. Evers · M. Farisco
Centre for Research Ethics & Bioethics, Uppsala University, Box 564, 751 22 Uppsala, Sweden
e-mail: Arleen.salles@crb.uu.se

A. Salles
Programa de Neuroética, Centro de Investigaciones Filosóficas,
Buenos Aires, Argentina

M. Farisco
Biogem Genetic Research Centre, Ariano Irpino, AV, Italy

¹ This does not entail that there are no additional drivers for RRI. See Rip, 2016.

including the role of the scientist [4]. However, in general, there has been a tendency to understand responsibility in individualistic and atomistic terms. To illustrate, within the research context, scientists have often been considered responsible for advancing knowledge and doing so in compliance with basic ethical and legal norms, for example, but generally not considered ethically responsible for the social, political and cultural impact of their findings or their potential extra-scientific uses and misuses [5]. Policy makers, on the other hand, are expected to be responsible for impact assessment and devising regulations and guidelines [6]. The problematic nature of such a fragmented approach to responsibility in research and emerging technologies is highlighted by the discourse on RRI. Such discourse places particular weight on a collective notion of responsibility: for RRI, responsibility and irresponsibility are distributed throughout the research and innovation process and they directly involve researchers, innovators, funders, policy makers, and other stakeholders [1, 2, 4]. Thus, the need for a forward looking and collective notion of responsibility -that addresses people's ambivalences and concerns regarding the products of scientific advances- and for mechanisms that will promote it.

From a philosophical, and specifically ontologically-grounded reasoning, the need for a richer notion of responsibility was already suggested by Hans Jonas, who argued that responsibility arises not only from practical reasons but more fundamentally from the ontological nature of life itself, more specifically from its being an end in itself [7]. Jonas derives the ethical imperative to act so that the resulting effects of one's actions are compatible with the continuation of an authentic human life. Of course, the importance of a deeper and more encompassing understanding of responsibility can also be established without Jonas' ontological commitments, as illustrated by Martha Nussbaum's human capabilities approach suggesting that a collective notion of responsibility (to promote people's living a truly human life, for example) could be grounded on an Aristotelian-based notion of human beings as moral agents and members of a community of peers [8, 9].

At present, RRI policy narratives urge that science and technology be aligned with societal needs and that research be carried out for and with society [3]. This suggests some awareness that scientific research is a social enterprise and must be recognized as such not only by its practitioners, but also by those who are affected by it. A dualistic view of science and society

that fails to recognize that science has a social identity is inadequate to assess the complexity of the issues emerging from their interaction. Scientific research, even if very specialized and unique, is itself a social enterprise. Among other things, this implies that: a- other social stakeholders can and should improve their understanding of what scientists do; b- science (understood either as a collective enterprise or as the activity of individual scientists) should enhance inclusive and collaborative relationships with the rest of society, and c- neither the scientific methodology nor the scientific goals are neutral with respect to external societal influences, and they necessarily affect other social contexts.

Within this framework, the RRI discourse has an important aspirational dimension: research and innovation should be socially beneficial. Accordingly, it proposes that rather than addressing the legal, ethical, and social dimensions of research and innovation by focusing primarily on outcomes, a careful assessment of the diverse potential emergent issues should inform the trajectory of the scientific work and feed into the research agenda itself. In practice, this requires both engagement with a number of societal actors and multidisciplinary interactions particularly with the social sciences and the humanities. Because the idea is to open up a space for inquiry that acknowledges the inherently social and political aspects of research, a sharp division of labour between the scientific and the socio-cultural-ethical tasks is undesirable [10].

It is true that scientific work is importantly mediated by a community of peers that either confirm or challenge the relevant findings. However, this reflective system of "checks and balances" typically does not entail scientific self-reflection on goals and values. RRI suggests that critical and reflective approaches to science and technology should not be seen as contingently provided by external disciplines or as work to be done by an external discipline: they are essential to the scientific enterprise itself. Science and technology in themselves are socially, ethically, and legally relevant: if those who work in these fields are not aware of this intrinsic connection their activities will be correspondingly limited. A reflection about the complex nature of science and technology, as well as about their extra-scientific impact, should be part of a conceptually mature scientific enterprise.

Despite its promise, the concept of RRI is not without theoretical and practical challenges: its definition is open ended, some of its features have been considered rather conceptually obscure, and its implementation,

even when possible, is often complex [6, 10, 11]. Still, the group of ideas underlying RRI – including the call for diversity and the integration of a number of actors – are being proposed as an adequate theoretical and practical foundation for governance of all scientific and technological research carried out in Europe (<https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation>).

In this paper, we do not address potential limitations but focus on the potential contribution of *philosophical reflection* to RRI in the context of the Ethics and Society subproject of the Human Brain Project (HBP). We show how the type of conceptual analysis provided by philosophically oriented approaches such as, for example, fundamental neuroethics, theoretically and ethically broadens research and innovation within the HBP [12–15]. We further suggest that overt inclusion of philosophical reflection can promote the aims and objectives of RRI.

RRI, “Reflexivity,” and “Philosophical Reflection”

Despite the lack of clear definitions of the main features of RRI, there is some agreement that it entails a commitment to a number of activities [1, 3, 4, 16]. It starts with the activity of *anticipation* – concerned with the identification of potential ethical and social concerns at an early stage –; it further requires methodological *reflexivity* – on the motivation and direction of science, societal goals and values, and desired impacts – [1, 2, 4], as well as the *inclusion* of a broad set of stakeholders – in order to promote a debate with all those affected by the research and to empower social agency – [2, 16, 17], and *responsiveness*, – specifically, the capacity to respond and change course on the basis of the relevant stakeholders’ and society’s values, and in view of the appropriate circumstances [2, 4].

The issue of how to understand the notions of anticipation, engagement, and inclusion has received significant attention in the context of RRI [1, 3, 4]. Less has been said about what the notion of reflexivity entails in this context. [1, 3, 16] Considering the different meanings of ‘reflexivity’, and its potential for becoming redundant (after all, isn’t reflexivity part and parcel of any academic activity?) it is important to understand how the notion can be conceived within the RRI framework.

Jack Stilgoe and colleagues focus on the term “reflexivity” as related to institutional practice, and

explain it as a specific kind of deliberate and self-critical attitude towards one’s own activities, commitments, and assumptions while recognizing the limits of knowledge and the variability in the framing of issues [2]. They are interested in its capacity to enable self-awareness beyond the lab, on how research is formulated and how it responds to social challenges. Bernd Stahl, Ethics director in the Human Brain Project, points to the need for internal reflexivity in order to “explore the assumptions and consequences of research” [16] as do Christine Aicardi and colleagues (also from the HBP) when highlighting the need for scientists and other stakeholders to be reflexive about the commitments that drive them and that shape the outcomes of their research [18].

To the extent that RRI calls for a socially embedded understanding of the motivations and scientific agendas, awareness of own assumptions and biases, and identification and recognition of existing uncertainties, such methodological reflexivity needs the kind of reflection provided by social science scholars who can bring to light the diverse aspects of social life, politics, and culture included in the scientific space. This is not to suggest that there is no reflection in the scientific domain but rather to recognize that considering that science curricula typically do not include the kind of training that would help science students to discover hidden social values and assumptions that shape both the scientific questions asked and the interpretation of the evidence, neuroscientists on their own might be unprepared to uncover and assess them [12, 19]. To the extent that this is the case, if the goal is a more ethically aware and sustainable scientific research and innovation that does not neglect social and political aspects, the active involvement of social scientists in the scientific research process and in research agenda setting plays a key role.

However, there is an additional dimension of reflexivity that should be highlighted, the one provided by philosophical reflection that may contribute to RRI a much-needed analysis of basic scientific and ethical concepts and their possible interpretations. In what follows, we elaborate on this point by focusing on the work of the Neuroethics and Philosophy group and its contribution to the HBP.

RRI and Reflexivity in the Human Brain Project

The HBP (a European Community Flagship Project of Information and Computing Technologies) proposes

that to achieve a fuller and more integrated understanding of the brain it is necessary to identify, integrate, and take advantage of the massive volumes of both already available data and new data coming from labs around the world. The project involves the development of new supercomputing technologies to federate and manage the data, to integrate it in computer models and simulations of the brain, to identify patterns and organizational principles that only appear when the data is gathered, and to identify gaps to be filled by new experiments [20, 21]. Expected outcomes of the research include the creation and operation of an ICT infrastructure for neuroscience and brain related research in medicine and computing which will help us achieve a multilevel understanding of the brain (from genes to cognition); of its diseases and the effects of drugs (allowing early diagnoses and personalised treatments) and to capture the brain's computational capabilities [22].

The HBP is funded by the European Commission in the framework of the EU's Horizon 2020 research-funding programme, which actively promotes RRI actions (public engagement, reflection, anticipation) aimed at the ethical and social acceptability of the research process and its products [23]. Indeed, one of the HBP Subprojects, Ethics and Society, has been developed to broaden and enhance RRI into all HBP research. This subproject is structured around a number of activities such as initially presented by Stilgoe and Richard Owen [2, 4]: foresight analysis uses scenarios construction to identify at an early stage the ethical and social concerns raised both by potential HBP research developments and their implications. It also produces reports to be used as background information by HBP directors, researchers, and other stakeholders [18]. Citizens' engagement promotes involvement with different points of view and strengthens public dialogue with public and private stakeholders via organization of workshops, webinars, and a number of other outreach activities; and ethics management develops principles and mechanisms for their implementation, creates Standard Operating Procedure (SOPs) and ensures that the ethical issues raised by the different research subprojects are transparently communicated and managed and that HBP researchers comply with the relevant ethical codes and legal norms [17, 24].

It seems evident that deliberative and introspective processes should play an important role in this integration of social, scientific, and ethical inquiry. It is worth noting, though, that in addition to the sociological and

ethical reflection required by the activities described, the HBP Ethics and Society subproject includes an additional dimension often lacking in other research projects: philosophical reflection. This is a type of reflection that aims to offer more than assistance to neuroscientists and social scientists in identifying the social, political, and cultural components of the research [14, 24]. In the HBP, philosophical reflection is intended to open a different and productive space for examining the relevant issues, carrying out self-critical analysis, and contributing to the understanding of HBP research itself.

At the root of philosophical analysis is the idea that engaging at a purely conceptual level and examining and clarifying the core concepts and language used by neuroscience and its resulting knowledge enhances both self-critical analysis and the evaluation of ethical concerns. This does not mean that philosophy is completely autonomous or self-referential in this conceptual task. Other disciplines, like theoretical physics, history or developmental biology, importantly contribute to such endeavour. But the fact that they do should not obscure the role that philosophy plays in adding to our understanding of neuroscience, its conceptual assumptions, epistemic virtues and limitations.

Unfortunately, in the context of neuroscience and the discussion of its implications, philosophical reflection on the issues raised by science and technology has often been reduced to the identification of the potential practical implications of the products of science and the application of ethical theory to manage them. In other terms, philosophy has been reduced to a poorly understood ethics, and ethics reduced to an after the fact examination and management of scientific conduct (and misconduct) and of the effects of neuroscience's products on the basis of more or less objective principles. It is not surprising, then, that in a number of domains there has been a tendency to identify a philosophical approach to science with a misguided type of applied ethics understood as a merely procedural approach. Indeed, at times, it seems as if philosophy's contribution to neuroethics – an interdisciplinary field that addresses the ethical, legal, and social questions raised by brain research – is limited to mechanically providing a repertoire of ethical approaches to be used to address practical concerns. The problem, however, is that conceptualized in this way philosophy becomes a mere tool often limited to manage risks and thus insufficient for unpacking some important concerns that people have regarding brain research- i.e. data protection,

privacy, dual use-, and for furthering conceptual transparency [25, 26].

An applied ethical approach (what could be called a “neuro-bioethics”) can be reflexive: a philosophically reflexive neurobioethics plays an important role in the discussion of a number of normative issues raised by brain research. We would like to advance, however, that within brain research the role of philosophical analysis is not exhausted by its contribution to addressing normative issues. Although they are significant, normative issues do not fully capture what is at stake in the scientific enterprise [12, 13, 15]. There are other, non normative concerns that need to be addressed, notably, the role of neuroscientific research in addressing fundamental philosophical questions. If interpreting scientific data in social and historical context is important in order to gain understanding, so is a careful conceptual analysis of key scientific notions such as, for example, matter. Such analysis facilitates a more integrated picture of, and a legitimate connection between neuroscientific findings and philosophical notions and questions. In short, understanding the neuroscientific enterprise and the issues it raises requires also theoretical philosophical reflection. Such reflection aims to do two things: bring to the forefront dimensions typically unacknowledged thus stifling the tendency to interpret neuroscientific results in a simplistic fashion, and, in the process, offer different and possibly complementary approaches to the issues investigated by empirical science.

Some call this type of philosophical approach to neuroscientific research “fundamental neuroethics” [12, 13, 15]. Beyond the name, the important point is that this neuroethical approach uses conceptual analysis of some foundational notions (concepts and methods) of neuroscience to provide the necessary background in examining the potential impact of neuroscience on topics such as the mind/brain relationship, criteria for consciousness, the question of what sets beings human beings apart, personal responsibility, and freedom among others (cf. [12]). The proposed philosophical reflection does not exclusively focus on ethical applications and on the moral permissibility of the implementation of neuroscientific findings. Nor is it primarily concerned with the social embeddedness of the scientific enterprise. Rather, it takes as a starting point the view that in addition to its social and ethical dimensions, brain research has important ontological and epistemic dimensions that need to be addressed in themselves (and, of course in order to adequately identify and manage

ethical issues as well). In short, the full range of issues raised by neuroscience cannot be adequately dealt with without also focusing on epistemic and ontological aspects that play a major role in the quality of the research process (for example, in framing scientific questions) and the legitimacy of the various interpretations of relevant scientific findings. The ethical, ontological, and epistemological aspects are not independent from each other but rather interwoven; effective reflection needs to address them all.

Philosophical Reflection and RRI in Practice

Our point so far has been that philosophical reflection can play a key work in the background of science, foresight, engagement, and ethics management (all key dimensions of RRI). Next, we illustrate this role by focusing on three concrete examples relevant to HBP research: neuroscientific research and its impact on human identity, neuroscientific studies on the unconscious, and neuroscientific studies of poverty.

Human Identity

It is often suggested that by providing knowledge of the structures and functions of the brain, not only will neuroscience enable a richer understanding of the brain and its diseases, but it will further our understanding of what sort of beings we are [27–30]. Indeed, unveiling some of the components that make us human is an explicit goal of the HBP (project that is particularly interested in addressing foundational issues) and an implicit interest in other international brain initiatives. It must be noted that the issue of whether there are species typical properties and what they are is not just of theoretical interest but of practical interest as well. The same progress in brain research that might allow considerable insights into what human beings are can also afford means to manipulate and access the brain. Some fear that such manipulation and access could make a significant impact on what humans are, and quite likely alter how people understand themselves as human [31–33]. Thus, to explore what is that “something” that makes humans different from other beings -what we can loosely call “human identity”- and to identify the underlying assumptions when discussing such identity should be a substantial concern to anyone making claims about the extent to which neuroscience might

unveil what we are, or how neurotechnology will alter us.

But the notion of human identity generates deep questions: what does it specify? And does giving a key role to brain research in uncovering fundamental components of what human beings are betray a problematic neuroessentialism or braincentrism, (i.e. the view that the brain plays a unique role in human identity and that any meaningful approach to what we are must entail a focus on this organ) [34, 35]? It is clear that a productive approach to examining these topics requires a careful philosophical conceptual examination of both epistemological and ontological issues [15]. Epistemologically, core questions are: what can neuroscience tell us about human identity and why? What are the limits of neuroscientific knowledge when it comes to understanding human beings? And how can we bridge the gap between the knowledge provided by neuroscience and the knowledge provided by the social and natural sciences that have done work on the same notion? From an ontological perspective, the core issues appear to be whether there is a human identity, whether such identity is to be grounded on essential or non essential traits, or whether, instead, it is to be found in a particular kind of process. [36–40] The issue becomes more complex when we consider that for a number of historical and religious reasons, the possession of a human identity (often discussed in philosophy in terms of “human nature”) has often been taken to mean moral superiority.

And yet, the issue of whether such human traits exist or whether the idea itself of a human identity (often discussed in philosophy in terms of “human nature”) makes sense is still debated .

Conceptual analysis of recent empirical work on the brain gives support to the view that holds human identity to be based upon a particular process: the lengthy, constant, and complex interplay between human cerebral architecture and its diverse environments [22, 41–45]. Indeed, in recent years, neuroscientific discourse on the brain has developed a more nuanced understanding of this organ and its relational aspects, including its relationship with the body, its many environments, and the social contexts in which it is embedded. The view of the brain as a mechanistic input-output processing device has been consistently questioned and generally abandoned [29, 41, 42]. In particular, an alternative model that sees the brain as an «autonomously active, plastic, projective» and highly selective organ heavily affected by learning and experience has greater

explanatory potential [27, 42, 43, 46, 47]. The epigenetic model of neuronal development proposes that even if constrained by a genetic envelope, the human brain is able to adapt its neuronal connectivity by stabilizing or eliminating particular synapses in accordance with short- and long-term changes in its internal and external environment [41, 42]. The theory of neuronal epigenesis by selective stabilization has been used not just to explain the development of the brain, but also the acquisition of written and oral language, and the acceptance of and compliance with social and ethical rules [48, 49]. If correct, it provides grounds for endorsing a process oriented view of what humans are. Rather than looking for intrinsic universal human traits or presumptively confirming the importance of one or a group of specific behavioral or anatomical markers, it suggests that in our quest for humanity, we could focus on the constant interplay that allows for the coalescence of learning, experience, and genes and examine how dynamic interactions, and social environments impact synaptic connectivity and contribute to the formation of a variety of patterns of neural activity.

The general point is, however, that real progress in understanding what being human is and in discussing whether certain neurotechnologies will threaten humanity will only be made after rigorous conceptual analysis of the relevant philosophical and scientific notions is integrated into a comprehensive approach. In practice, expressions of fear regarding the potentially dehumanizing aspect of brain machine interfaces, robotics, or even DBS procedures can, at least partially, be explained by the prevalence of different, often muddled conceptions of human identity. Searching for conceptual clarity on what makes us human is then not just a valuable endeavour in itself but a way to responsibly address and potentially manage serious concerns regarding the neuroscientific agenda itself and the implications of the products of research.

Studies on Consciousness

The investigation of consciousness in the last few years increasingly reveals the inadequacy of a limited approach to the phenomenon. Accordingly, both empirical and conceptual efforts are devoted to consciousness research within the HBP.

Several scientific and conceptual models of consciousness have been suggested, and agreement is far from being reached [50–55]. Even if the debate about

their empirical and conceptual interpretation is still open, more agreement has been found regarding the role of so-called “neural correlates of consciousness” (NCC), i.e. a set of neuronal structures and functions correlating with conscious phenomena (such as wakefulness and arousal). Since their formal introduction in the scientific debate at the beginning of 1990s [56], the neural correlates of consciousness have been widely scrutinized from both conceptual and empirical points of view [57, 58]. Conceptually, David Chalmers defines NCC as minimal neuronal activations necessary for consciousness. [59]. Such a general definition has been widely accepted in both philosophical and empirical contexts, even though the need for a more accurate definition of NCC has recently been suggested [60].

More specifically, NCC can be described in two basic ways: either as referring to a general, global *state* of consciousness, i.e. as neural correlates that mark the difference between being and not being conscious, or as referring to particular *contents* of consciousness, i.e. as neural correlates that are sufficient for a specific object to enter consciousness [59, 61].

The empirical differentiation between understanding NCC as referring to a state and understanding NCC as referring to the content of consciousness is reflected in the clinical distinction between wakefulness and awareness, i.e. between the state of vigilance and the content of conscious processing [62]. This differentiation is also relevant to the description of the complexity of consciousness, which is not reducible to the processing of information coming from outside, but is also a sort of background state which allows processing that information [63].

Research on NCC provides important clues about the cerebral structures and functions involved in conscious phenomena. Yet, notwithstanding some progress in recent empirical investigations and conceptual clarifications of consciousness, we still lack an overarching theory providing a unitary picture of consciousness and related disorders. Michele Farisco and colleagues have recently formulated a new conceptual model [64], the Intrinsic Consciousness Theory (ICT), that starts from the predisposition of the brain to evaluate and to model the world [41, 47], i.e. from the brain’s ability to check the usefulness of the world to the satisfaction of its intrinsic needs and to develop a kind of map of the world in order to survive and thrive. Recent empirical investigation of the brain’s intrinsic activity (i.e., independent from external stimulation) and resting state activity (i.e., increasing in absence of

external stimulation) contribute in describing this organ as much more than an input-output machine but rather as spontaneously active [65, 66].

According to ICT, these intrinsic activities of the brain are identical with consciousness, even if at a very basic level (i.e., a level of consciousness the subject is not aware of). The distinction between consciousness and the unconscious is not discrete or binary: the ability of the brain to evaluate and model the world can occur in two modalities, implicit or explicit, unaware or aware, that correspond to what we usually refer to as the unconscious and consciousness, and both are multilevel configurations of the brain along a continuous and dynamic line. This means that consciousness can be depicted as an overarching brain characteristic, which the brain retains insofar as it is intrinsically active. Thus, starting from an empirical understanding of the brain as intrinsically active and plastic, ICT distinguishes between higher cognitive functions and basic phenomenal consciousness, suggesting that the latter might characterize the brain’s intrinsic activity as such, even if at a very basic level. The necessary and sufficient conditions for consciousness are that the brain have appropriate intrinsic and resting state activities.

This new conceptual model of consciousness is conceptually parsimonious and practically relevant, specifically with reference to the assessment and care of patients suffering from disorders of consciousness. It opens the possibility that what is usually described as the unconscious, which according to ICT is an unaware modality of consciousness possibly characterized by a very basic level of phenomenality, with specific abilities and needs, might be ethically relevant as well [67]. Moreover, at the ethical level ICT outlines the high level of elaboration and sophistication that the unaware brain exhibits, claiming for an appropriate treatment in clinical context.

Finally, the intrinsic consciousness theory is quite relevant to the RRI goals of promoting a more ethical process and practice. If we accept the wider model of consciousness suggested by this theory we can see that allocation of resources and research priorities in the clinical context of disorders of consciousness is often unjustifiably limited to cases in which it is possible to show that there is residual awareness, while unaware abilities potentially retained by affected patients are underestimated if not ignored. But this is not inevitable: a reconceptualization of consciousness that zooms in on the brain’s retained intrinsic activity rather than on its retained reactivity to external stimulation can start a

richer discussion with clear practical implications. In this way, ICT would make a positive contribution to RRI in allowing a more comprehensive ethical assessment of challenging cases like disorders of consciousness.

Neuroscientific Studies of Poverty

We focused above on two areas in which neuroscientific research may have an impact either by potentially altering our self-understanding *qua* humans or by enhancing our understanding of consciousness, which in turn could have significant implications on the care of patients with disorders of consciousness. Neuroscientific research may do more: it can also inspire systemic social change. The potential for this becomes evident when we focus on contemporary neuroscientific studies on the influences of poverty on cognitive, emotional, and stress regulation systems that propose to analyse how the different individual and contextual factors associated with material, emotional, and symbolic deprivation (i.e., lack of food, shelter, education, and health-care) influence neural development [68]. These studies have important ethical and public policy implications: they should play an important role in the discussion of a number of issues such as what are the structural conditions needed for the full exercise of human rights, the overt and covert ways in which citizens' rights can be violated, what respect for human dignity entails, the potential ways of depriving people from their identity as full citizens and of restoring such identity, and the determination of collective social responsibilities [68, 69]. The specific evidence that neuroscience brings to the analyses of poverty and its implications is then not only of great interest in itself but also relevant to policy-making. However, this evidence needs to be spelled out in detail and clarified conceptually, notably in terms of causes of and attitudes toward poverty, implications of poverty for brain development, and the possibilities to reduce and reverse these effects.

It is important to be cautious when interpreting the results of neuroscientific studies that consider contextual and cultural aspects to avoid misconceptions and stigmatization: there is notably a sizable difference between considering that neural and behavioural differences due to poverty are a deficit and considering them an adaptation. From an ethical perspective, the issue then becomes whether the consequences of poverty are related to circumstances in which no basic rights are satisfied

(e.g., inadequate nutrition, housing, or access to education and health services). This is an important point considering that social attitudes to poverty differ: some consider it a result of social irresponsibility, whereas others take a more individualistic approach and explain it as a personal failure of the person afflicted. In other words, poverty is not universally regarded as a consequence of social imbalances and unequal access to social benefits. The latter, individualistic views are common in North America and South America, where the problem of poverty is much more significant than in Western and Northern Europe, where social views on poverty are dominant. It is worth noting that countries and political systems that accept the rights of access to adequate nourishment, housing, education, and health care as a shared social responsibility are also among those who have been most successful in combatting poverty, social violence, and insecurity, e.g., the Scandinavian countries (cf. World Bank Global Poverty Overview <http://www.worldbank.org/en/topic/measuringpoverty#3>).

The relevance of neuroscientific evidence to policy-making and legislation can be illustrated by focusing on adolescent delinquency that arises more frequently in contexts of poverty. As pointed out elsewhere [49], this is frequently repressed through police and judiciary means, often resulting in incarceration. However, this approach to juvenile violence simply omits compelling findings that show that adolescence is a time of “neurodevelopmental crisis.” Evidence from anatomical and functional-imaging studies has highlighted major modifications of cortical circuits during adolescence including reductions of gyrification and grey matter, increases in the myelination of cortico-cortical connections, and changes in the architecture of large-scale cortical networks—including precentral, temporal, and frontal areas [70]. Uhlhaas and colleagues [71] have used Magneto Encephalography synchrony as an indicator of conscious access and cognitive performance [72]. Until early adolescence, developmental improvements in cognitive performance are accompanied by increases in neural MEG synchrony. This developmental phase is followed by an unexpected decrease in neural synchrony that occurs during late adolescence and is associated with reduced performance. This period of destabilization is followed by a reorganization of synchronization patterns that is accompanied by pronounced increases in gamma-band power and in theta and beta phase synchrony [71]. These remarkable changes in neural connectivity and performance in the adolescent are now being explored: awareness of their occurrence should lead to special

proactive care from society. The nature of this care may include a social educative environment adapted to adolescents' special needs, adequate physical exercise, or new kinds of therapies yet to be developed. The point is that a careful discussion of these issues is a moral priority, particularly considering that depending on the circumstances in some countries young people can be transferred to the adult system (e.g., Canada, United States, England, Wales) and in others there is a strong political will for them to be treated as adults (Brazil, Argentina). In view of the available evidence from neuroscience, social policies that treat and punish minors as adults may arguably be not merely ineffective, but also a clear breach of human rights. In this sense, neuroscientists can play an extremely useful role, providing and reinforcing the kind of evidence needed to understand minor delinquency and how to manage it in a way that promotes the wellbeing and respects the rights of all. In doing so, they could raise awareness of a point made before: scientific work is intrinsically socially and politically relevant.

The example above illustrates the need for careful unpacking and conceptual clarification of the specific evidence that neuroscience brings to analyses of poverty and of careful ethical analysis of its implications. There are additional related issues already suggested by Sebastian Lipina and Kathinka Evers [73]. One of them has to do with the causes and attitudes to poverty. For example, what does neuroscience concretely contribute to the debates over individualistic versus systemic or social explanations of poverty? On the other hand, what does considering the neural and behavioural differences due to poverty as a deficit rather than as an adaptation imply? A second one is related to the *Impacts of poverty on brain development*. The interpretation of evidence and the identification of ethical and social issues that arise might provide the means for reducing poverty's negative impacts. We have also *questions of reversibility*. Which impacts of poverty can be reversed, and how? What does the concept of "reversibility" entail? The evidence available in this area raises specific ethical challenges that should also be considered in the interpretation of results and the planning of future research. Finally, there is also the important question of how to communicate scientific findings in a responsible manner so as to avoid misconceptions, hypes, misuses, or the problem of stigmatization. Scientific findings can always be misused, but the risk is greater in an area permeated with values and norms, such as that of poverty. It is not self-evident that the science of poverty will be used for poverty-alleviation: it can also be used to increase

alienation from "the poor", and deepen the stigmatization of a group that is already disrespected in societies that have created such circumstances.

Conclusion

The framework of RRI calls for contextually addressing not just such impact -social benefits and drawbacks- but also the background research process, specially the societal visions underlying it and the norms and priorities that shape scientific agendas. This requires the integration of anticipatory, inclusive, and responsive dimensions, and the nurturing of a certain type of reflexivity among a variety of stakeholders, from scientists to funders. Such integration will hopefully result in a science and the innovation that comes from it that are more attuned to societal considerations and needs.

In this paper, illustrating with work on human identity, consciousness, and poverty and the brain, we propose that enhanced reflexivity within RRI benefits not only from interrogating entrenched political and cultural commitments, social contexts, and scientific motivations but also from conceptually analyzing the meanings of scientific terms and of the language used to define science and its products that frequently reinforces problematic assumptions about science itself and its role. Our claims are not intended then as a criticism of how RRI is typically put into practice. Rather, they should be seen as an opportunity: including philosophical analysis so as to generate a better and more sustainable research not only from the ethical but also the ontological and epistemological perspectives.

Acknowledgments Special thanks to Karen Rommelfanger and Tom Buller for their thoughtful comments on a previous draft of the paper and to Roland Nadler for the very relevant information he provided. We would also like to thank two anonymous readers for Neuroethics for their insightful critiques and feedback. The research is supported by funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 720270 (HBP SGA1) and 785907 (HBP SGA2).

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided

you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Von Schomberg, R. 2013. A vision of responsible research and innovation. In *Responsible Innovation*, ed. R. Owen, M. Heintz, and J. Bessant. Chichester, UK: Wiley & Son.
- Stilgoe, J., R. Owen, and P. Macnaghten. 2013. Developing a framework for responsible innovation. *Research Policy* 42: 1568–1580.
- Owen, R., P. Macnaghten, and J. Stilgoe. 2012. Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy* 39: 751–760.
- Owen, R., J. Stilgoe, P. Macnaghten, M. Gorman, E. Fisher, and D. Guston. 2013. A framework for responsible innovation. In *Responsible Innovation*, ed. R. Owen, J. Bessant, and M. Chichester Heintz. UK: John Wiley & son.
- Evers, K. 2009. Ethics in Science: A Socio Political Challenge. In *Universality: From Theory to Practice*, ed. B. Sitter-Liver. Fribourg: Academic Press.
- Fisher, E., and A. Rip. 2013. Responsible innovation: Multi-level dynamics and soft intervention practices. In *Responsible innovation*, ed. R. Owen, J. Bessant, and M. Heintz. Chichester UK: Wiley & Son.
- Jonas, Hans. 1984. *The imperative of responsibility : in search of an ethics for the technological age*. University of Chicago press: new Aufl. Chicago, IL.
- Sen, A. 1999. *Development as freedom*. Oxford, UK: Oxford University Press.
- Nussbaum, M. 2006. *Frontiers of justice: Disability, nationality, species membership*. Cambridge, MA: Harvard University Press.
- Rip, A. 2016. The clothes of the emperor. An essay on RRI in and around Brussels. *Journal of Responsible Innovation* 3 (3): 290–304.
- Weckert, J., H. Rodríguez Valdes, and S. Soltanzadeh. 2016. A problem with societal desirability as a component of responsible research and innovation: The “if we don’t somebody else will” argument. *Nanoethics* 10: 215–225.
- Farisco, M., Evers K., Salles, A. Forthcoming. Fundamental neuroethics ten years later: An overview of the field. *Cambridge Quarterly of Healthcare Ethics*.
- Evers, K., A. Salles, and M. Farisco. 2017. Theoretical framing of neuroethics: the need for a conceptual approach. In *Debates about Neuroethics: perspectives on its development, focus and future*, ed. E. Racine and J. Dordrecht Aspler. Springer International Publishing.
- Farisco, M., K. Evers, and A. Salles. 2016. Big science, brain simulation, and Neuroethics. *AJOB Neurosci* 7 (1): 28–29.
- Salles, A., and K. Evers. 2017. Social neuroscience and Neuroethics: A fruitful synergy. In *Social Neuroscience and social science: The Missing Link*, ed. A. Ibanez, L. Sedeno, and A. Garcia. Springer.
- Stahl, B. 2013. Responsible research and innovation: The role of privacy in an emerging framework. *Science and Public Policy* 40: 708–716.
- Stahl, B., S. Rainey, and M. Shaw. 2016. Managing ethics in the HBP: A reflective and dialogical approach. *AJOB Neurosci* 7 (1): 20–24.
- Aicardi, C., Reinsborough, M., Rose, N. 2017. The integrated ethics and society programme of the human brain project: Reflecting on an ongoing experience. *Journal of Responsible Innovation*.
- Sahakian, B.J., and S. Morein-Zamir. 2009. Neuroscientists need neuroethics teaching. *Science* 325 (5937): 147.
- Amunts, K., C. Ebell, J. Muller, M. Telefont, A. Knoll, and T. Lippert. 2016. The human brain project: Creating a European research infrastructure to decode the human brain. *Neuron* 92 (3): 574–581. <https://doi.org/10.1016/j.neuron.2016.10.046>.
- Markram, H., K. Meier, T. Lippert, S. Grillner, R. Frackowiak, S. Dehaene, A. Knoll, H. Sompolinsky, K. Verstreken, J. DeFelipe, S. Grant, J.-P. Changeux, and A. Saria. 2011. Introducing the human brain project. *Procedia Computer Science*. 7: 39–42.
- Changeux, J.P. 2017. Climbing brain levels of organisation from genes to consciousness. *Trends in Cognitive Sciences* 21 (3): 168–181. <https://doi.org/10.1016/j.tics.2017.01.004>.
- CEC. 2011. Proposal for a regulation establishing Horizon 2020—*The framework programme for research and innovation* (2014–2020).
- Evers, K. 2017. The contribution of neuroethics to international brain research initiatives. *Nature Reviews Neuroscience* 18: 1–2.
- Evers, K. 2005. Neuroethics: A philosophical challenge. *The American Journal of Bioethics* 5 (2): 31–33; discussion W33–34. <https://doi.org/10.1080/15265160590960302>.
- Evers, K. 2007. Towards a philosophy for neuroethics. An informed materialist view of the brain might help to develop theoretical frameworks for applied neuroethics. *EMBO Reports* 8 Spec No:S48–S51. <https://doi.org/10.1038/sj.embor.7401014>.
- Evers, K. 2009. *Quand la matière s'éveille*. Paris: Éditions Odile Jacob.
- Gazzaniga, Michael S. 2008. *Human : The science behind what makes us unique*. New York: Ecco.
- Le Doux, J. 2003. *Synaptic self: How our brains become who we are*. New York, NY: Penguin Books.
- Ramachandran, V.S. 2011. *The tell-tale brain*. New York, NY: W.W. Norton & Company Inc.
- Jottrand, F. 2014. Questioning the moral enhancement project. *The American Journal of Bioethics* 14 (4): 1–3. <https://doi.org/10.1080/15265161.2014.905031>.
- Kirmayer, L. 2012. The future of critical neuroscience. In *Critical Neuroscience. A handbook of the social and cultural contexts of neuroscience*, ed. J. Slaby and S. Choudhury. Chichester: Wiley&Sons Ltd.
- Clayden, Lisa. 2011. Lae, Neuroscience, and Criminal Culpability. In *Law and Neuroscience: Current Legal Issues*. Oxford University Press: Michael Freeman.
- Vidal, F. 2009. Brainhood, anthropological figure of modernity. *History of the Human Sciences* 22 (1): 5–36.
- Racine, E., O. Bar-Ilan, and J. Illes. 2005. fMRI in the public eye. *Nature Reviews. Neuroscience* 6 (2): 159–164. <https://doi.org/10.1038/nrn1609>.
- Hull, D. 1986. On human nature. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*. 1986: 3–13.
- Machery, E. 2008. A plea for human nature. *Philosophical Psychology* 21 (3): 321–329. <https://doi.org/10.1080/09515080802170119>.

38. Lewens, T. 2012. Human nature: The very idea. *Philosophy and Technology* 25: 459–474.
39. Linquist, S., E. Machery, P.E. Griffiths, and K. Stotz. 2011. Exploring the folkbiological conception of human nature. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366 (1563): 444–453. <https://doi.org/10.1098/rstb.2010.0224>.
40. Stotz, K. 2010. Human nature and cognitive-developmental niche construction. *Phenomenology and the Cognitive Sciences* 9 (4): 483–501. <https://doi.org/10.1007/s11097-010-9178-7>.
41. Changeux, J.P. 1985. *Neuronal Man*. Princeton NJ: Princeton University Press.
42. Changeux, Jean-Pierre. 2004. *The physiology of truth : Neuroscience and human knowledge*. Cambridge, Mass: Belknap Press of Harvard University Press.
43. Changeux, J. P. 2012. Synaptic epigenesis and the evolution of higher brain functions. In *Epigenetics, brain and behavior*, ed. P. Sassone-Corsi, Christen, Y. Dordrecht, NL: Springer.
44. Changeux, J.P. 2014. Genes, Brains, and Culture: From Monkey to Human. In *From Monkey Brain to Human Brain: A Fyssen Foundation Symposium*, ed. S. Dehaene, J.R. Duhamel, M. Hauser, and G. Rizzolatti. Cambridge, Mass: MIT Press.
45. Changeux, J.P., P. Courge, and A. Danchin. 1973. A theory of the epigenesis of neuronal networks by selective stabilization of synapses. *Proceedings of the National Academy of Sciences of the United States of America* 70 (10): 2974–2978.
46. Edelman, G. 2017. Building a Picture of the Brain. In *The Brain*, ed. G. Edelman and J.P. Changeux. New York: Routledge.
47. Edelman, G. 1987. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic books, Inc., Publishers.
48. Evers, K., and J.P. Changeux. 2016. Proactive epigenesis and ethical innovation: A neuronal hypothesis for the genesis of ethical rules. *EMBO Reports* 17 (10): 1361–1364. <https://doi.org/10.15252/embr.201642783>.
49. Evers, K. 2015. Can we be epigenetically proactive? In *Open Mind: Philosophy and the mind sciences in the 21st century*, ed. T. Metzinger, J. M. Windt, Cambridge, MA: MIT Press.
50. Block, N., O. Flanagan, and G. Güzeldere. 1997. *The nature of consciousness: Philosophical debates*. Cambridge: MIT Press.
51. Laureys, S. 2015. Un si brillant cerveau. Les états limites de consciences. Odile Jacob.
52. Dehaene, Stanislas. 2014. *Consciousness and the brain : deciphering how the brain codes our thoughts*. New York: Viking Adult.
53. Facco, E., C. Agrillo, and B. Greyson. 2015. Epistemological implications of near-death experiences and other non-ordinary mental expressions: Moving beyond the concept of altered state of consciousness. *Medical Hypotheses* 85 (1): 85–93. <https://doi.org/10.1016/j.mehy.2015.04.004>.
54. Chalmers, D. 1996. *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
55. Tononi, G., M. Boly, M. Massimini, and C. Koch. 2016. Integrated information theory: From consciousness to its physical substrate. *Nature Reviews. Neuroscience* 17 (7): 450–461. <https://doi.org/10.1038/nrn.2016.44>.
56. Crick, F., and K. Koch. 1990. Towards a neurobiological theory of consciousness. *Seminars in Neuroscience* 2: 263–275.
57. Koch, C., M. Massimini, M. Boly, and G. Tononi. 2016. Neural correlates of consciousness: Progress and problems. *Nature Reviews. Neuroscience* 17 (5): 307–321. <https://doi.org/10.1038/nrn.2016.22>.
58. Metzinger, T. 2000. *Neural correlates of consciousness: Empirical and conceptual issues*. Cambridge, MA: MIT Press.
59. Chalmers, D. 2000. What is a neural correlate of consciousness? In *Neural correlates of consciousness: Empirical and conceptual questions*, ed. T. Metzinger, 17–39. Cambridge, MA: MIT Press.
60. Fink, S.B. 2016. A deeper look at the "neural correlate of consciousness". *Frontiers in Psychology* 7: 1044. <https://doi.org/10.3389/fpsyg.2016.01044>.
61. Overgaard, M. 2017. The status and future of consciousness research. *Frontiers in Psychology* 8.
62. Laureys, S. 2005. The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences* 9 (12): 556–559. <https://doi.org/10.1016/j.tics.2005.10.010>.
63. Thompson, Evan. 2015. *Waking, dreaming, being : Self and consciousness in neuroscience, meditation, and philosophy*. New York: Columbia University Press.
64. Farisco, M., S. Laureys, and K. Evers. 2017. The intrinsic activity of the brain and its relation to levels and disorders of consciousness. *Mind and Matter* 15 (2): 197–219.
65. Dehaene, S., and J.P. Changeux. 2005. Ongoing spontaneous activity controls access to consciousness: A neuronal model for inattentive blindness. *PLoS Biology* 3 (5): e141. <https://doi.org/10.1371/journal.pbio.0030141>.
66. Raichle, M.E., A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, and G.L. Shulman. 2001. A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America* 98 (2): 676–682. <https://doi.org/10.1073/pnas.98.2.676>.
67. Farisco, M., and K. Evers. 2017. The ethical relevance of the unconscious. *Philosophy, Ethics and Humanities in Medicine* 12 (11): 11.
68. Lipina, S.J., and J.A. Colombo. 2009. *Poverty and brain development during childhood: An approach from cognitive psychology and neuroscience*. American Psychological Association.
69. Lipina, S.J., and M.I. Posner. 2012. The impact of poverty on the development of brain networks. *Frontiers in Human Neuroscience* 6: 238. <https://doi.org/10.3389/fnhum.2012.00238>.
70. Klein, D., A. Rotarska-Jagiela, E. Genc, S. Sritharan, H. Mohr, F. Roux, C.E. Han, M. Kaiser, W. Singer, and P.J. Uhlhaas. 2014. Adolescent brain maturation and cortical folding: Evidence for reductions in gyrification. *PLoS One* 9 (1): e84914. <https://doi.org/10.1371/journal.pone.0084914>.
71. Uhlhaas, P., F. Roux, W. Singer, C. Haenschel, R. Sireteanu, and E. Rodriguez. 2009. The development of neural synchrony reflects late maturation and restructuring of functional networks in humans. *Proceedings of the National Academy of Sciences* 106 (24).
72. Dehaene, S., and J.P. Changeux. 2011. Experimental and theoretical approaches to conscious processing. *Neuron* 70 (2): 200–227. <https://doi.org/10.1016/j.neuron.2011.03.018>.
73. Lipina, S.J., and K. Evers. 2017. Neuroscience of childhood poverty: Evidence of impacts and mechanisms as vehicles of dialog with ethics. *Frontiers in Psychology* 8: 61. <https://doi.org/10.3389/fpsyg.2017.00061>.