

# Pragmatism for a Digital Society: The (In)Significance of Artificial Intelligence and Neural Technology

Matthew Sample<sup>1,2</sup> and Eric Racine<sup>1-4</sup>

1. Pragmatic Health Ethics Research Unit, Institut de recherches cliniques de Montréal
2. Department of Neurology and Neurosurgery, McGill University
3. Department of Experimental Medicine (Biomedical Ethics Unit), McGill University
4. Department of Medicine and Department of Social and Preventive Medicine, Université de Montréal

## Abstract

Headlines in 2019 are inundated with claims about the “digital society,” making sweeping assertions of societal benefits and dangers caused by range of technologies. This situation would seem an ideal motivation for ethics research, and indeed much research on this topic is published, with more every day. However, ethics researchers may feel a sense of *déjà vu*, as they recall decades of other heavily promoted technological platforms, from genomics and nanotechnology to machine learning. How should ethics researchers respond to the waves of rhetoric and accompanying ethics research? What makes the digital society significant for ethics research? In this paper, we consider two examples of digital technologies (artificial intelligence and neural technologies), showing the pattern of societal and academic resources dedicated to them. This pattern, we argue, reveals the jointly sociological and ethical character of significance attributed to emerging technologies. By attending to insights from pragmatism and science and technology studies, ethics researchers can better understand how these features of significance effect their work and adjust their methods accordingly. In short, we argue that the significance driving ethics research should be grounded in public engagement, critical analysis of technology’s “vanguard visions”, and in a personal attitude of reflexivity.

Keywords: digital society, pragmatism, ethics, sociotechnical imaginaries

## 1. Introduction – Waves of Technology (Ethics)

In 2019, the prospect of a “digital society” seems to dominate the collective imagination, both in policy and research circles, as well as in popular media. How can it not, with recent high-profile scandals and media events centered on data and privacy, as when Mark Zuckerberg (Facebook CEO) was summoned to appear before multiple governing bodies including the US Congress and British parliament? This year’s shift in attention towards “the digital” did not happen spontaneously, however. Setting aside the longer (and deeply consequential) history of the Internet and other information technologies, governments around the world have already spent several years devising new initiatives and investing resources under this banner. In 2011, for example, EU member states each appointed “Digital Champions,” representatives who were given the mandate to “help every European become digital and benefit from an inclusive digital society” [1]. Now,

we see digital agendas everywhere, with an internationally booming private sector dedicated to information technology and the emergence of global collaborations for digital forms of governance; the European Commission, has announced a new funding program dubbed “Digital Europe.” Canada, with a similar rationale, recently signed onto “the Digital Seven” (D7), joining Uruguay and six other nations in pursuing new technological possibilities and in promoting “digital government.”

Throughout, the word “digital” might imply that there is some shared artifact or infrastructure being envisioned in all of these initiatives. The word has been used to refer to social media platforms and the Internet, and alternatively, as a label for any electronic devices that rely on programming. Nevertheless, the rhetoric being used here is more accurately summarized in terms of its massive and sweeping claims about the transformative impact of new technologies, rather than in terms of some particular object. Some claims stress the benefits to healthcare, economic productivity, and a whole range of social practices. Others highlight the tremendous dangers and risks of these technologies: enabling authoritarian practices, threatening privacy and equality, diluting the quality of social interaction, spreading misinformation, and so much more. Who is to be believed? There is certainly a hefty dose of social marketing promoting the value and benefits of digital technologies. But who is to say if those positive impacts will materialize or not in technological form and then in concrete social realities? More to the point of the present paper, what is an appropriate response to these developments from academic or policy researchers given these unknowns?

Faced with this quandary, many observers of sociotechnical change will experience a feeling of *déjà vu*. We have already observed multiple waves of technological rhetoric and corresponding societal and ethical worry, particularly among some researchers and in some segments of the media. Since just the 1990s, we have witnessed the rise (and occasional fall) of numerous big ideas: human genomics, nanotechnology, synthetic biology, neural engineering, big data, blockchain, personalized medicine, precision medicine, and most recently, artificial intelligence (AI). The momentary prominence of digital technologies has even led to combinations with previously promoted categories, like the use of machine learning in optimizing clinical care and in training brain-computer interfaces. Along with these waves, academics have assembled technology-centered specializations, most notably for the present paper, “neuroethics” and “AI ethics.” For each of these waves, the core ethical question is a simple one: “will this be a good technology or a good development?”

But answering this question and even just bringing clarity to it has proven difficult. Frequently and still today, ethics research on emerging technologies seems to be triggered by hyperbolic technological discourse, with limited critical scrutiny of both positive and negative speculative or promissory claims about emerging technologies [2-9]. Academic articles discussing artificial intelligence as an existential threat, for example, are likely to coincide with front page articles

about Google DeepMind, national leaders' speeches about digital innovation, and protests outside of Microsoft offices in Redmond, WA. Far from questioning the technological promises and worries of the day, ethicists may be among the first to reference and reinforce them through conferences, media appearances, and publications [10-12] as we have seen previously regarding other technologies [13]. A corollary is that these questions and worries have remained academic with limited genuine public engagement and concern for the impact of technology on everyday life [14-18]. As we will argue, such moments of cultural alignment – between ethics research and society – may create a false sense of significance regarding the objects of ethics inquiry. And captive to this skewed sense of significance and relevance, ethics research may become unfit to answer emerging ethical questions rigorously, let alone foundational questions in ethics. As Rayner (2004) has phrased the question, “why does institutional learning about new technologies seem so difficult?” [19].

In this paper, we aim to make progress on this front with respect to scholarship on emerging technology. We advance a discussion about the ethical *significance* of the digital society, in a way that neither dismisses nor naively embraces its societal prominence in 2019. We here understand *significance* as both the ethical importance currently granted to certain technologies and, more critically, the open-ended question of what significance these technologies should have within a broader view on human wellbeing and flourishing. To this end, we will present some recent developments in AI and neural technology as two case studies, creating a high-level picture of the digital society. We will then analyze this picture using insights from science and technology studies (STS) and from pragmatism. We suggest, ultimately, that the current significance of emerging or speculative digital technologies is underpinned by both sociological and ethical factors; that is to say, the digital society is both a product of self-interested technology promoters and something that can impact the well-being of individuals and society, positively or negatively. This has, we conclude, implications for the practice of ethics research on technology and dictates an interdisciplinary approach that focuses on impacts for individual life and democracy. It means, furthermore, that the significance of any given technology is not a foregone conclusion and that ethics researchers themselves have a role to play in foregrounding some problems over others and in handling critically the attribution of significance.

## **2. The Digital Society is Here, Again: Parallel Trends in Academia and in Society**

From social science and law to ethics and policy, a wide range of academic disciplines contribute to scholarship that is labeled as “ethics” or is taken to have an “ethical” dimension of inquiry. Part of the intuitive appeal of doing or applying this type of work, which we will refer to here as *ethics research*, is that it is more meaningful than the popular discourses associated with passing trends and national fads associated with technoscience. Unlike the undisciplined gaze of the public, the lone academic mind or research team can filter out meaning from mere noise. Or so we might think. Yet, there is reason to think that our scholarship on technology maps quite closely onto

broader societal trends. Focusing here on two exemplar cases (i.e. neural technology and AI), we can see general parallels between academic and societal attention given to emerging or speculative technologies, with only indirect links to actual harms and benefits as experienced by actual (as opposed to hypothetical) persons. As we will argue, the mirroring of academia and broader society is not necessarily or entirely problematic [20], but demands attention in order to properly understand and respond meaningfully to the ethical significance of the digital society.

### *Neural Technologies*

Consider first the case of BCIs and, more broadly, neural technology. First devised in the 1970s, neural devices now take many forms, either wearable or implanted, experimental or widely available, and span medical, commercial, security, and recreational uses [21-24]. The underlying logic of these devices is that the user's brain activity can be monitored and/or altered with digital hardware to create another means of interacting with the world. User-controlled BCIs are perhaps most well-known for allowing the user to communicate or control a prosthesis, despite complete bodily paralysis (e.g. due to amyotrophic lateral sclerosis) [21]. Other devices, like implanted deep-brain stimulators and wearable trans-cranial direct current stimulators, alter brain activity more directly and have been promised to improve the user's mood or, as widely documented and viewed on YouTube [25], reduce the symptoms of Parkinson's disease. Though these new affordances are covered in the media and have taken on symbolic importance [26], the framing is often primarily positive or promotional rather than critical. Coverage of neuroscientific technologies like BCIs often lacks a balanced consideration of their negative and positive impacts [27-28].

In the private sector, investment in neural technologies is passing 100 million USD per year, according to one estimate [29]. Meanwhile, Elon Musk (CEO of Tesla and SpaceX) has famously claimed that BCIs are the only way humans can remain relevant and productive in an increasingly automated economy; Musk even started his own BCI company, Neuralink, exhibiting his commitment to that envisioned future [30]. As farfetched and speculative as his reaction may seem, attitudes reported among some publics are not contradictory and indicate a shared measure of optimism when presented with the idea of brain-based devices. A Pew Center poll reports that respondents are simultaneously interested and quite worried about neural technologies that could be used for enhancement [31]. To the same effect, a recent study of public attitudes towards BCI ethics in Germany, Spain, and Canada reports that most respondents expressed moderate to high levels of worry about a wide range of potential ethical concerns of BCI use, but were nonetheless enthusiastic about using neural technology in medical applications [32]. Perhaps unsurprisingly, these trends in broader society are mirrored in research in several domains.

There are many indicators that neural technologies have become an object of concern in academia, with steady growth over the last few years thanks to specialized funding streams in many countries. The United States BRAIN initiative (explicitly oriented towards technology-driven discovery) and

the EU Human Brain Project each included dedicated (albeit proportionally small) funding for ethics-oriented research on neuroscience and its applications. Ethics researchers from various disciplines have reacted swiftly to these incentives. The existence of neural technologies has become a prominent, if not questionable, justification of a new field of research [33], “neuroethics,” which has become increasingly established since its first conference in May 2002. Non-governmental organizations and research centers, too, have been founded to collect and recognize ethical work related to neural technologies, like the International Neuroethics Society, the Neuroethics Network, Neuroethics Canada, and the Oxford Center for Neuroethics. The authors of this chapter have also contributed to this phenomenon.

Worries about BCIs reported in this literature range from lack of safety and cost to threats to the user’s self-understanding and responsibility; in most cases, these concerns are asserted on the basis of conceptual analysis or philosophical reflection [21]. Despite this tendency, the ethics of neural technology is not ‘all talk’. Underlying the transient and sometimes hyperbolic discourses in academic journals, conferences, press releases, news outlets, and social media, there seem to be some genuinely serious harms and benefits. It is thus imperative to remember that the experiences and well-being of some people have already been impacted, either in the course of BCI research or in its applications. Some participants in BCI research studies, for instance, may gain through participation a new capability to express themselves, using BCI devices to work around communication difficulties. But such research studies sometimes end with either no gain in communicative ability or, in the case of success, with a complete lack of technical and clinical support for continued use [34].

Other notable harms may be related to the interaction between the application of new technologies and political or cultural recognition. Some members of the Deaf community (i.e. individuals who collectively embrace a cultural identity linked to being deaf) have reported that cochlear implants, designed to augment hearing, actually reinforce systematic stigmatization of their bodies and lifestyle [35]. The hasty promotion of cochlear implants may ultimately preclude the acceptance of Deaf community members as simply different, undermining political obligations to make public infrastructure accessible to them. This dynamic has also been reported in reference to BCIs more generally. As part of a multi-stakeholder international deliberation in 2018, potential BCI users and patient advocates reported the ongoing disenfranchisement of individuals due to either the use of exclusionary technologies (that devalue certain types of bodies) or the failure to make enabling technologies widely available [36]. In sum, these examples show the stakes of effective and beneficent neural technology, despite media hype, are real and deeply consequential in some contexts.

### *Artificial Intelligence and Machine Learning*

Now in 2019, as major brain-oriented government funding initiatives are beginning to sunset, there is space for another platform to represent the digital society. At the moment, this alternative seems

to be AI, supplementing front-page media imagery of disembodied brains and wires with screenshots of friendly AI chatbots or schematic representations of neural networks. As neural technologies fade into our collective attentional background (temporarily or not), we should carefully attend to the way in which AI has newly been constructed as yet another significant technology: sudden and substantial media coverage (despite a longer history), dedicated government funding including ethics research, general awareness among publics, and of course serious stakes and harms for human experience. A schematic comparison with neural technologies reveals the same general pattern at work.

Looking back, AI also has a history that predates its current popularity. While the use of neural technologies often draws on the centuries-old belief that we *are* our brains [37], AI draws on a more recent permutation of the mind-body thesis. The first investments in AI research in the United States were primarily driven by a two-part justification: first, the academic belief that complex human cognition can be modeled and supplemented by computer systems (a form of digital “symbiosis”) and, second, the promise to meet Cold War demands for semi-automatic control for precise military decision-making (e.g., with cybernetics) [38]. AI proponents have, however, consistently promoted applications outside of military and defense purposes, applying the computer’s ability to solve problems heuristically to medical decision-making [39], logistics [40], translation, marketing, and a range of other uses. Notably, AI has even been applied to BCIs – in the form of adaptive algorithms for the interpretation of neural activity – at least since the early 1990s. These various applications have occasionally coincided with periodic surges in media attention.

In 1984, one commentator in *Science* lauds “exhilarating times” for AI research, and stresses the need to keep promises in line with “the science” [41]. So too, in recent years. At the unveiling of the Stanford Institute for Human-Centered AI, Bill Gates (former CEO of Microsoft) asserted that artificial intelligence is akin to nuclear technology, in terms of its promise and danger to humanity [42]. The *New York Times* has posed the question: “Is Ethical AI even possible?” [43]. These widely-reported opinions arrive after a global flurry of publications of national AI strategies, beginning with “A.I. Singapore” [44] in 2017 to the “American AI Initiative” of 2019, which aims to “develop AI in order to increase our Nation’s prosperity, enhance our national and economic security, and improve quality of life for the American” but does not directly earmark any additional resources for this task [45]. Reactions collected among some publics track this ambivalent media coverage. In an EU study, 61% of respondents reported positive attitudes towards robots and artificial intelligence, but 88% agree that this technology must be carefully managed to avoid negative impacts [46]. A similar study of Americans reported that approximately 40% of respondents supported the development of AI, with 82% agreeing with the need for careful management [47].

Again, as with neural technology, ethics research on technology has pivoted towards “AI ethics”

[48-50]. Part of this is made possible due to dedicated resources from funding agencies and institutions, with some controversial private-public partnerships. Amazon has recently co-funded a US National Science Foundation solicitation for AI and fairness research proposals, matching the government's 10 million USD [51]. Facebook garnered negative press when it announced the establishment of a \$7.5 million USD AI ethics institute at the Technical University of Munich, which appeared to some observers as a blatant form of "ethics-washing" for the company [52]. MIT will include AI ethics as a core focus in its newly founded Schwarzman College of Computing. For most of these projects, it is too early to evaluate their outcomes in terms of ethical progress, but there is already a large amount of theoretical content being produced by technology ethics researchers from a variety of disciplines.

Perhaps most remarkable about this global institutional shift to AI ethics is the sheer number of codes of ethics and guidelines that have been produced. In March 2019, a non-exhaustive search yields 9 unique sets of guidelines or principles [53-61], with many more reports and white papers not listed here. The tendency in this genre of document is not to document precise harms or benefits of AI; instead, signatories or authors agree on high-level aspirational goals that should guide the development and use AI (e.g. beneficence, fairness, democracy, or empowerment) [62]. Some guidelines, like the Montreal Declaration and Microsoft's "Approach to AI" are branded explicitly in terms of the private firm or polity in which they originated. Again, as with neural technologies, these transient societal discourses serve many purposes and are only indirectly tied to documented and concrete harms or benefits.

Noteworthy harms exist, nonetheless. Investigative journalists at ProPublica, for instance, report the use of algorithm-based risk assessment software that is more likely to falsely label black defendants as future criminals, at almost double the rate of white defendants. This error is consequential, they argue, because judges across the United States can and do cite risk scores during sentencing and at bond hearings [63]. This negative use of AI is sadly not unique and represents a range of practices that motivate the (much more abstract) calls for "fairness" and the avoidance of bias in AI applications [64-65]. In another high-profile case, an experimental autonomous car designed by Uber struck and killed Elaine Herzberg, a pedestrian in Tempe, AZ [66]. With the absence of regulatory guidance from local or national governments, safety, beneficence, and related ideals for AI applications demand more than the publication of principles or temporary media attention.

Taken together, where do these considerations leave us with respect to the ethical significance of the digital society? Studied as two exemplar digital technologies, neural technology and AI begin to shed light on some core features associated with ethical significance as it is commonly attributed. In brief these are: periodic surges in media coverage that belie the technology's longer history, dedicated government funding, and some cases of definite harms and benefits related (causally or otherwise) to digital technology. What lessons do we learn from the fact that this

pattern is repeated every few years for each new technological platform? What reaction is appropriate for researchers in ethics of technology, and what can we learn from this pattern?

### **3. Two Wrong Answers: Significance-as-Consensus and Reduction to Hype**

Faced with the pattern that we list in the previous section, there are two simplistic interpretations that need to be set aside: *significance-as-consensus* and *reduction to hype*. Though the two perspectives are presented here in their most extreme form and are perhaps not widely held, they facilitate the elaboration of a defensible middle ground (next section).

First, *significance-as-consensus*: it might be tempting to think that technology promoters, skeptical ethics researchers, and broader society are simply on the right track. That is to say, digital technologies are indeed ethically significant on account of a broad consensus on their importance, positive or negative. After all, why would so many individuals, states, and institutions invest so much energy into talking about them, developing them, and studying them? But this cannot be the whole story. Interpreting significance as consensus does not explain the constant shift in framing of technological challenges, from genomics to AI and back again. We believe that is unlikely that, as *significance-as-consensus* might imply, that AI was only societally meaningful in the 1950-70s and since approximately 2017 during its recent resurgence in popularity. We doubt that the meaningful effects of AI use on communities and individuals ceased in between these periods of attention, though as we will argue this is an empirical question.

At the other extreme, there is *reduction to hype*. More cynical observers will reduce academic-societal attention to technologies as the result of an inevitable boom-and-bust cycle, a predictable outcome of media hype and subsequent disappointment. And the content of hype, as the term's negative valence implies, can be and should be ignored. A commonly cited schematic version of this is the Gartner hype "cycle," which has been criticized for its lack of explanatory power [67]. But as we discuss above, there seem to be real issues present behind the waves of rhetoric and hype, even if the media eventually lose interest. Moreover, the examples we selectively list are just that, examples. Because of the unavoidably limited journalistic and empirical research on the impact of each technology, there remain many unanswered questions about the real utility and value, real-world impact of technology on people's lives, and what matters to people. Reducing technology discourses to mere hype inappropriately discounts this possibility.

### **4. Significant Technologies? Insights from STS and Pragmatism**

Ethics researchers can avoid the pitfalls of *significance-as-consensus* and *reduction to hype* to understand how technologies come and should come to be significant. In this section, we will briefly present two tools for critically understanding waves of technological hype and worry: *economies of promising* and the *problems of publics*. Researchers in the ethics of technology can benefit from learning both, because they are themselves implicated in these processes and these



tools can help them reflect critically on their own practices and their understanding of science and technology.

### *Economies of Promising: Transactions and Vanguard Visions*

The first theoretical tool is *economies of promising*. As we present it here, it is a sociological framework and describes how STEM (science, technology, engineering, and mathematics) researchers, funding agencies, and policy-makers use promises to create political legitimacy for their activities. In particular, recent literature in STS suggests that the mobilization of resources and the behavior of individual researchers is best understood in terms of high-level interactions between different societal actors. Joly (2010) suggests that the resources given to technoscience are administered according to what promises STEM researchers or tech developers are willing to make, whether to funding agencies or investors [67]. Technoscience is typically practiced in specialized spaces, inaccessible to most citizens, and there is a need to justify the work in terms of what impact it will have outside of the lab. This is a requirement even for researchers who have no intention or capacity to fulfill their promises. Joly labels this system of relationships as an “economy of technoscientific promising” [67].

This fact about the research ecosystem does not imply, however, that those making the promises are wholly dishonest or acting in bad faith. Some sociological research suggests that technoscientific practitioners will often take one of two performative positions with respect to promising [68]. In contexts close to the research (e.g. in a lab meeting), their knowledge of the expansive possibility space and of the myriad technical obstacles results in expressions of deep uncertainty and rejection of guarantees or promises. But when speaking with potential technology users, funders, or taking on the role of the user themselves, scientists and engineers must affirm the imminent utility of their work for concrete practical purposes. Promising in this way does more than just attract resources to one’s projects; it yields self-understanding and serves as a source of creativity.

Explained from a slightly different conceptual angle, Hilgartner (2015), suggests that technoscientific researchers must understand and describe their research in terms of an “imaginary” that engages with values present in the surrounding institutional, regional, or national culture [69]. Synthetic biologists in US agriculture, for example, might notice that previously funded projects in other areas of biology are framed in terms of helping fulfill the reputation of the US as a global leader in food production and food innovation. The synthetic biologists, in response, will frame their own work in this way, borrowing the ends of prior projects, but with a new means. Over time, if this particular project and “vanguard vision” for the technology gains sufficient traction, the discourses around it forms a new “sociotechnical imaginary” [70], a symbolic construct which can spread well beyond the lab-funder relationship. This can happen via university press releases, social media, technology blogs, funding agency reports, among other channels.

Both *economies of promising* and their associated vanguard visions are sociologically important because they help us to understand the cultural repertoires from which individual actors craft their behaviors and come to understand themselves. In other words, the action of shared imaginaries resolves the fundamental sociological antinomy of structure and agent [70]. But more to the point of this paper, they also have meaningful connections to ethics-oriented research on technology. First, they highlight that the ethical significance of a given technology is not always an intrinsic property of the objects or infrastructures themselves. The lesson of *economies of promising* is that technology can be made significant by the coordinated efforts of STEM researchers, developers, users, and whoever else has the means and motivation to formulate and promote a “vanguard vision.” As a result, we should expect imaginaries to be as numerous as their creators, with all the overlap, contradictions, and conceptual confusion that such multiplicity implies. It means too that we can empirically document imaginaries at various stages of uptake, from the first sentence of an unread tech ethics paper (e.g. “Digital technologies are poised to transform society and the lives of disabled people, but may be ethically worrying.”) to the internationally broadcasted speech of a prime minister or CEO.

The economy of promising has made its way directly into the work of ethics researchers when they build on existing imaginaries and speculate about the prospective benefits and harms of new technologies. It has been shown how, for example, ethics scholarship has replicated uncritically claims about the transformative impact of various technologies such as deep brain stimulation [71-72], 3D bioprinting [73], and cognitive enhancers [74-77]. Given the task of ethics, namely to analyze critically courses of actions and propose, via deliberation, paths and scenarios which are promissory of human flourishing (and identify paths which represent obstacles to flourishing), it is not surprising that ethics researchers are concerned about the future implications of technology. By its very nature, ethics is future-oriented although it can build on the past. However, it is in the manner by which claims about the future make their way into ethical analyses about harms and benefits which summons important theoretical, methodological, and practical concerns.

Second, these imaginaries necessarily feature valuations regarding what is the good life, the good community, etc. although these can be implicit in discourse and practices. These are all evaluable from the standpoint of ethics. Moreover, because imaginaries are the product of intensive work by particular actors, they may not represent the interests of the majority or of the most immediate stakeholders. In the face of historical inequality in society and in technoscientific practices, we cannot expect marginalized individuals or groups to have adequate resources to promote their own sociotechnical visions or to counter unacceptable “vanguard visions,” creating knock-on effect on the process by which technologies become significant. The US rhetoric of global AI leadership, for instance, may have reinforced the significance of AI in the United States without any meaningful input from individuals who have been subject to racial discrimination in the immigration or criminal justice systems. In this way, the sociological dynamics of significance can short circuit democratic hopes of representation and could negatively shape our choice of foci in

academic ethics research. More problematic is the power and influence of vanguard visions may imply that other voices be considered as enemies of the nation, of progress, of technoscientific development, and so on. In response, fostering deliberative approaches and dialogue is a clear strategy to offset the narrowness by which vanguard visions and imaginaries are conceived and deployed. This is the object of our next point.

### *Emerging Publics and Their Problems*

The second theoretical tool, *publics*, emphasizes the very real problematic experiences that can sometimes ground discussions of technology. Dewey identifies a public as something that coalesces around shared problems [78]. In his account, the formation of a public starts when a few individuals notice their chosen forms of life are being affected by some common factor, technoscientific or social, natural or of solely human origin. These individuals can choose to foster a group identity. Once this public has formed, its members can make rights claims or demand political representation in addressing the shared problem. This conception contrasts with the colloquial understanding of “public” as the rather neutral and insipient aggregate sum of individuals in a population. Rather than one singular entity representing the population, there can be multiple publics, just as there can be multiple problems, and with occasional overlap given the existence of multiple interests.

As we discuss above, “the digital society” and its many platforms create a circular or self-referential discourse in many ways, but they also present some very real harms. We know about these harms in part because of the publics that have formed around them. Notable publics of neural technology, for example, include disability activists rejecting the use of cochlear implants to augment hearing. As discussed above, this negative potential of neural technology can be understood in terms of personhood, specifically the experience of being disenfranchised and treated as less than a full political person. In this case, everyday experiences of disability stigma – as echoed by the promotion of the cochlear implant – have been sufficient to cause a variety of citizens to mobilize around this shared threat. Similarly in artificial intelligence, the experiences of unequal treatment in criminal justice, once documented and shared, have galvanized grassroots movements to remedy ongoing government negligence in the proper regulation of algorithm-assisted sentencing.

The assessment of these *publics and their problems* has a sociological component; in the case of emerging technologies, which publics are formed is predominantly an empirical question, requiring us to go out and look for ourselves and to understand the nature of problematic situations (real or foreseen) based on the experience of those who use/will use technology or be affected by it. We will likely find that some publics (ethicists experiencing a shared need for funding) may have less pressing problems than others (members of the Deaf community who experience stigmatization due to cochlear implants). However, for Dewey, reflection on the formation of publics is an essentially philosophical and political undertaking. Unlike *economies of promising*,

the concept of *publics* does not imply transactional, economic, or strategic functions.

Instead, Dewey's *publics* are sites of group intelligence and open-ended possibilities for human flourishing. He understands group deliberation to be a core component of democracy [79], a process of fostering dialogue to enrich our views about present-day problematic situations and, eventually, form a more comprehensive outlook. For pragmatists, this is the sign of a growth in ethical perspectives [80-81]. The concept of *public*, then, is a reminder that ethics research can and could be an instrument that empowers individuals (grouped as publics) to have their interests take a form and represented within open-ended (democratic) processes of political legitimacy formation and, in the ideal case, solve shared problems. Conversely, ethics researchers, unaware of the economy of promises and of its impact on them, their practices, their scholarship, their field, etc. (or sometimes deliberately complicit), can perpetuate discourse and practices that alienate everyday experience from the public gaze. The relationship of the ethics researcher with the public is often obfuscated by the servile role of ethics in the economy of promises or the use of ethics knowledge authoritatively to circumvent dialogue [7, 82] or bypass the understanding of perspectives of those concerned [33, 83].

## **5. Significance and Our Responsibility as Researchers**

The two tools discussed above – *economies of promising* and *publics and their problems* – reveal the dualistic character of ethical significance in this domain. Specifically, significance in the context of emerging technologies is jointly an ethical category, grounded in the experiences of real people, and a sociological dynamic, which we can embrace or resist. For this reason, we argue that the significance of “the digital society” – it is only the latest in a series of technological waves – demands a nuanced treatment from researchers in technology ethics. It is not enough to simply accept it *a priori* or to reject it out of hand as mere hype. Building on examples from emerging technologies and on theory, we stress the utility of an interdisciplinary approach to assessing significance. It involves, minimally, three overlapping strategies for ethics research: empirical and deliberative engagement, analysis of dominant imaginaries, and consistent reflexivity in our work.

First, as inspired by Dewey's account of publics, we prescribe engagement with the publics and the problems they face because of technology. Ethics researchers who position themselves as being in service of helping resolve problem situations need to remain true to the nature of the situation, notably that they are not themselves facing the situation and will not provide the answer themselves. That work is for those concerned by the situation. Accordingly, scholars can use empirical or social scientific methods to productively inquire into *which* publics are actually forming, beyond the self-serving discourses perpetuated by technology developers (in promotional roles) or by technology-oriented funding for ethics research. Furthermore, scholars can foster the productive and creative capacities of publics by convening deliberative exercises and events [84-86]. Such events can build on a rich literature on public engagement in sociotechnical change and,

as “technologies of humility”, can provide a much-needed counterbalance to “technologies of hubris” [16, 87-88].

The role of the ethics researcher with respect to publics is thus to accompany those concerned in making sense of their situation and helping them find the effective means to transform the situation via open inquiries and ethical deliberation. This does not prevent us from playing an important role, but it should be one in service of the situation so to speak rather than using different (e.g., academic) problem situations to build power and influence in specialist discourses or simply replicate existing valuations without questioning them.

Second, as inspired by work on the *economy of promises*, we recommend critical analysis of dominant sociotechnical imaginaries and vanguard visions motivating our work and implicit in our subject matter. A thorough analysis of imaginaries will require a keen philosophical eye for normative content, documenting and calling out ableist, classist, racist and other disenfranchising assumptions that animate imaginaries. It will also involve proactively inquiring after the sources of our imaginaries: where did we get them and whose interests does their distribution serve? Which voices and visions are absent? In sum, we propose that the vanguard visions used in ethics research should be recast as “grounded speculation,” [4, 89] according to which imaginaries are always situated either in broadly accessible forms of evidence or in deliberative exercises with diverse affected publics.

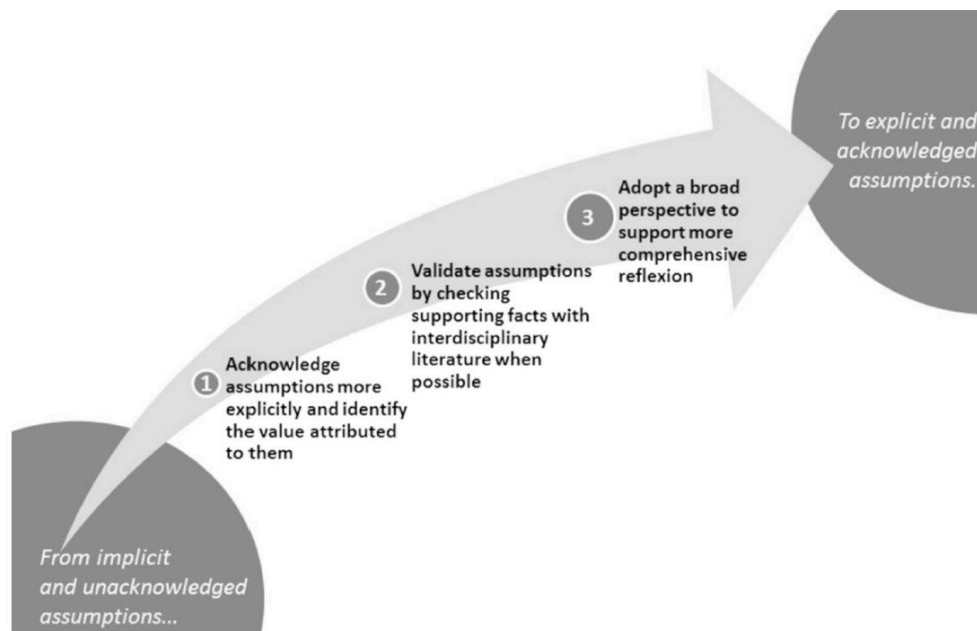


Fig. 1 Increasing objectivity and reflexivity in speculation

It has been proposed that greater transparency about different normative and factual (e.g., sociological, scientific) assumptions about speculative claims supporting ethical analyses be recognized as such, including the value attributed to these assumptions within ethical analyses [4]

Furthermore, ethics researchers can adopt methodological measures to cross-check speculative claims by, for example, validating assumptions with literature from different disciplines and adopting a broad perspective to support more comprehensive reflection (see Fig. 1) [4]. In this regard, comparing disciplinary frameworks, considering historical knowledge (e.g., evolution of a technology and attitudes toward it), and reflecting on the development of normative approaches toward a given technology are different methodological strategies to instill more objectivity, rigor, and comprehensiveness of ethical analyses which engage with new and evolving technologies.

Our two prescriptions for public engagement and for critical analysis of imaginaries can be recast as a unified intellectual virtue: reflexivity. The dualistic character of significance in this domain entails that responsible scholars are reflexive about their role in perpetuating some problems, some imaginaries, over others. We must be honest about the fact that we rarely conceive of a technological worry *de novo*, independent of our cultural moment or of front-page media. The authors acknowledge, of course, that this approach is not wholly new but combines the best features of recent sociological and ethical work on technology, including some of our own recent projects.

## 6. Conclusion

What is the significance of “the digital society” and related technologies such as BCIs and AI? How do we make sense of conflicting utopian, dismissive, and dystopian sentiments about its future? In this paper, we have considered strategic and sociological features of its significance, on the one hand, as well as features associated with actual impact on real world human experience, on the other. We have shown, by briefly considering two exemplar technologies, BCIS and AI, that these two aspects of significance are incompletely but meaningfully connected in society through overlapping technoscientific, social, and academic practices. We show, also, that this entanglement is likely not unique to any one technological platform but is a common formula across many waves of technological rhetoric.

This social phenomenon can be productively understood through pragmatist and STS resources. However, the understanding that we gain should be also translated into improving our own work on the ethics of technology. For this reason, responsible research in the face of endlessly emerging technologies requires a tailored methodology. Building on previous work, we emphasize the importance of “grounded speculation” and scholarly reflexivity. By taking these lessons to heart, researchers in technology ethics can address concrete problems facing actual communities and move beyond a vague and recurrent feeling of “haven’t we been through this before?”

## References

1. Digital champions joint mission statement. European Commission. 2014. <https://ec.europa.eu/digital-single-market/news/digital-champions-joint-mission-statement>. Accessed March 28, 2019.

2. Turner L. Bioethic\$ Inc. *Nat Biotech.* 2004; 22(8): 947-948.
3. Parens E, Johnston J, Moses J. Ethics. Do we need "synthetic bioethics"? *Science.* 2008; 321(5895): 1449.
4. Racine E, Martin Rubio T, Chandler J, Forlini C, Lucke J. The value and pitfalls of speculation about science and technology in bioethics: The case of cognitive enhancement. *Med Heal Care Phil.* 2014; 17(3): 325-327
5. Parens E, Johnston J. Against Hyphenated Ethics. *Bioethics Forum.* 2006. <http://www.bioethicsforum.org/genethics-neuroethics-nanoethics.asp>. Accessed March 28, 2019.
6. Parens E, Johnston J. Does it make sense to speak of neuroethics? Three problems with keying ethics to hot new science and technology. *EMBO Rep.* 2007; 8(1S): S61-64.
7. Evans JH. *Playing God?: Human genetic engineering and the rationalization of public bioethical debate.* Chicago: Chicago University Press; 2002.
8. De Vries R. Who will guard the guardians of neuroscience? Firing the neuroethical imagination. *EMBO Rep.* 2007; 8(S1): S65-69.
9. De Vries R. Framing neuroethics: A sociological assessment of the neuroethical imagination. *Am J Bioeth.* 2005; 5(2): 25-27.
10. Forlini C, Partridge B, Lucke, J, Racine E. Popular media and bioethics: Sharing responsibility for portrayals of cognitive enhancement with prescription medications. In: Clausen J, Levy N, editors. *Handbook on neuroethics.* Dordrecht: Springer; 2015. p. 1473-1486.
11. Caulfield T. The commercialisation of medical and scientific reporting. *PLoS Med* 2004; 1(3): e38.
12. Caulfield T. Biotechnology and the popular press: hype and the selling of science. *Trends Biotech.* 2004; 22(7): 337-339.
13. Hedgecoe A. Bioethics and the reinforcement of socio-technical expectations. *Soc Stud Sci.* 2010; 40(2): 163-186.
14. Forlini C, Racine E. Does the cognitive enhancement debate call for a renewal of the deliberative role of bioethics? In: Franke EHA, editor. *Cognitive enhancement: An interdisciplinary perspective.* New York: Springer; 2013. p. 173-86.
15. Racine E, Gareau I, Doucet H, Laudy D, Jobin G, Schraedley-Desmond P. Hyped biomedical science or uncritical reporting? Press coverage of genomics (1992-2001) in Québec. *Soc Sci Med.* 2006; 62(5): 1278-1290.
16. Doucet H. Imagining a neuroethics which would go further than genethics. *Am J Bioeth.* 2005; 5(2): 29-31.
17. Caulfield T, Condit C. Science and the sources of hype. *Pub Heal Genom* 2012; 15(3-4): 209-217.
18. Bubela T, Nisbet MC, Borchelt R, et al. Science communication reconsidered. *Nat Biotechnol.* 2009; 27(6): 514-8.
19. Rayner S. The novelty trap: why does institutional learning about new technologies seem so difficult? *Indus High Ed.* 2004 18(6), 349-355.
20. Simonson P. Bioethics and the rituals of media. *Hastings Center Rep.* 2002; 32(1): 32-39.
21. Burwell S, Sample M, Racine E. Ethical aspects of brain computer interfaces: A scoping review. *BMC Med Ethics.* 2017; 18(1): 60.

22. Dubljevic V, Saigle V, Racine E. The rising tide of tDCS in the media and academic literature. *Neuron*. 2014; 82(4): 731-36.
23. Wexler A. The social context of “do-it-yourself” brain stimulation: Neurohackers, biohackers, and lifehackers. *Front Hum Neurosci*. 2017. doi: 10.3389/fnhum.2017.00224
24. Clausen J. Man, machine and in between. *Nature*. 2009; 457(7233): 1080.
25. Gardner J, Warren N, Addison C, Samuel G. Persuasive bodies: Testimonies of deep brain stimulation and Parkinson's on YouTube. *Soc Sci Med*. 2019; 222: 44-51.
26. Jasanoff S. Perfecting the human: Posthuman imaginaries and technologies of reason. In Hurlbut JB, Tirosch-Samuels H, editors. *Perfecting human futures*. Wiesbaden: Springer; 2016. p. 73-95.
27. Cabrera LY, Bittlinger M, Lou H, Müller S, Illes J. The re-emergence of psychiatric neurosurgery: Insights from a cross-national study of newspaper and magazine coverage. *Acta Neurochir*. 2018;160(3): 625-635.
28. Racine E. Neuroscience and the media: Ethical challenges and opportunities. In Illes J, Sahakian B, editors. *Oxford handbook of neuroethics*. Oxford: Oxford University Press; 2011. p. 783-802.
29. Yuste R, Goering S, Bi G, et al. Four ethical priorities for neurotechnologies and AI. *Nat News*. 2017; 551(7679): 159-163.
30. Allen M, VandeHei J. Elon Musk: Humans must merge with machines. *Axios*. <https://www.axios.com/elon-musk-artificial-intelligence-neuralink-9d351dbb-987b-4b63-9fdc-617182922c33.html>. Accessed March 28, 2019.
31. Funk C, Kennedy B, Sciupac E. US public wary of biomedical technologies to “enhance” human abilities. Pew Research Center. 2016. <https://www.pewresearch.org/science/2016/07/26/u-s-public-wary-of-biomedical-technologies-to-enhance-human-abilities/>. Accessed March 28, 2019.
32. Sample M, Sattler S, Racine E, Blain-Moraes S, Rodriguez-Arias S. Do publics share experts' worries about neural technology? A trinational survey on the ethics of brain-computer interfaces. Under review. *Sci Tech Hum Val*.
33. Racine E, Sample M. Two problematic foundations of neuroethics and pragmatist reconstructions. *Camb Q Healthc Ethics*. 2018; 27(4): 566-577.
34. Klein E, Peters B, Higger M. Ethical considerations in ending exploratory brain-computer interface research studies in locked-in syndrome. *Camb Q Healthc Ethics*. 2018; 27(4): 660-674.
35. Sparrow R. Implants and ethnocide: Learning from the cochlear implant controversy. *Disabil Soc*. 2010; 25(4): 455-466.
36. Sample M, Aunos M, Blain-Moraes S, et al. Brain-computer interfaces and personhood: Interdisciplinary deliberations on neural technology. Under review. *J Neur Eng*.
37. Vidal F. Brainhood, anthropological figure of modernity. *Hist Human Sci*. 2009; 22(1); 5-36.
38. Edwards PN. *The closed world: Computers and the politics of discourse in Cold War America*. Cambridge: MIT Press; 1997.
39. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019; 25(1): 44-56.
40. AI-Powered Supply Chains Supercluster. Government of Canada. 2018. <https://www.ic.gc.ca/eic/site/093.nsf/eng/00009.html>. Accessed March 28, 2019.



41. Waldrop, MM. Artificial intelligence (I): into the world; AI has become a hot property in financial circles: but do the promises have anything to do with reality?. *Science*. 1984; 223: 802-806.
42. Paxton S, Yin W. Bill Gates, Gov. Gavin Newsom speak at unveiling of new human-centered artificial intelligence institute. 2019. *Stanford Daily*. <https://www.stanforddaily.com/2019/03/19/bill-gates-gov-gavin-newsom-speak-at-unveiling-of-new-human-centered-artificial-intelligence-institute/>. Accessed March 28, 2019.
43. Metz C. Is ethical AI even possible? *The New York Times*. 2019. <https://www.nytimes.com/2019/03/01/business/ethics-artificial-intelligence.html>. Accessed March 28, 2019.
44. Choudhury SR. Singapore to invest over \$100 million in A.I. in next five years in smart nation, innovation hub push. *CNBC*. 2017. <https://www.cnn.com/2017/05/03/singapores-national-research-foundation-to-invest-150-million-dollars-in-ai.html>. Accessed March 28, 2019.
45. Accelerating America's leadership in artificial intelligence. White House Office of Science and Technology Policy. 2019. <https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/>. Accessed March 28, 2019.
46. Special Eurobarometer 460: Attitudes towards the impact of digitisation and automation on daily life. European Commission. 2017. [https://data.europa.eu/euodp/data/dataset/S2160\\_87\\_1\\_460\\_ENG](https://data.europa.eu/euodp/data/dataset/S2160_87_1_460_ENG). Accessed March 28, 2019.
47. Zhang B, Dafoe A. Artificial intelligence: American attitudes and trends. Future of Humanity Institute. 2019. [https://governanceai.github.io/US-Public-Opinion-Report-Jan-2019/us\\_public\\_opinion\\_report\\_jan\\_2019.pdf](https://governanceai.github.io/US-Public-Opinion-Report-Jan-2019/us_public_opinion_report_jan_2019.pdf). Accessed March 28, 2019.
48. Anderson M, Anderson SL, editors. *Machine ethics*. Cambridge: Cambridge University Press; 2011.
49. The Ethics and Governance of Artificial Intelligence Initiative. 2017. <https://aiethicsinitiative.org>. Accessed March 28, 2019.
50. Bostrom N, Yudkowsky E. Ethics of artificial intelligence. In Frankish K, Ramsey W, editors. *New York: Cambridge University Press*; 2014. p. 316-334.
51. Natarajan P. Amazon and NSF collaborate to accelerate fairness in AI research. *Alexa Blogs*. 2019. <https://developer.amazon.com/blogs/alexa/post/1786ea03-2e55-4a93-9029-5df88c200ac1/amazon-and-nsf-collaborate-to-accelerate-fairness-in-ai-research>. Accessed March 28, 2019.
52. Facebook-funded AI ethics institute faces independence questions. *Times Higher Education*. <https://www.timeshighereducation.com/news/facebook-funded-ai-ethics-institute-faces-independence-questions>. Accessed March 28, 2019.
53. Floridi L, Cows J, Beltrametti M, et al. AI4People - An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations." *Minds Mach*. 2018; 28(4): 689-707.
54. Cutler A, Pribić M, Humphrey L. Everyday ethics for AI design. *IBM*. 2018. <https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf>. Accessed March 28, 2019.
55. Asilomar AI principles. Future of Life Institute. 2017. <https://futureoflife.org/ai-principles/>. Accessed March 28, 2019.

56. IEEE ethically aligned design. IEEE Standards Association. 2019. <https://standards.ieee.org/industry-connections/ec/autonomous-systems.html>. Accessed March 28, 2019.
57. Diakopoulos N, Friedler SA, Arenas M, et al. Principles for accountable algorithms and a social impact statement for algorithms. 2016. <http://www.fatml.org/resources/principles-for-accountable-algorithms>. Accessed March 28, 2019.
58. Montréal declaration for responsible development of artificial intelligence. 2019. <https://www.declarationmontreal-iaresponsable.com/la-declaration>. Accessed March 28, 2019.
59. The European Commission's High-Level Expert Group on artificial intelligence: Ethics guidelines for trustworthy AI. European Commission. 2018. <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>. Accessed March 28, 2019.
60. Artificial intelligence at Google: Our principles. Google AI. 2019. <https://ai.google/principles/>. Accessed March 28, 2019.
61. Microsoft AI principles. Microsoft Corporation. 2019. <https://www.microsoft.com/en-us/ai/our-approach-to-ai>. Accessed March 28, 2019.
62. Whittlestone J, Nyrup R, Alexandrova A, Dihal K, Cave S. Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. Nuffield Foundation. 2019. <http://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf>. Accessed March 28, 2019.
63. Angwin J, Larson J, Mattu S, Kirchner L. Machine Bias. ProPublica. 2016. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Accessed March 28, 2019.
64. Calo R. Artificial Intelligence policy: a primer and roadmap. *UCDL Rev.* 2017; 51: 399.
65. Courtland R. Bias detectives: the researchers striving to make algorithms fair. *Nature.* 2018; 558: 357-60.
66. Somerville H. Uber shuts Arizona self-driving program two months after fatal crash. Reuters. 2018. <https://www.reuters.com/article/us-autos-selfdriving-uber/uber-shuts-arizona-self-driving-program-two-months-after-fatal-crash-idUSKCN1I02SD>. Accessed March 28, 2019.
67. Joly PB. On the economics of techno-scientific promises. In Akrich M, Barthe Y, Muniesa F, Mustar P, editors. *Débordements. Mélanges offerts à Michel Callon*. Paris: Presses des Mines; 2010. p. 203-222.
68. Brown N, Michael M. A sociology of expectations: Retrospecting prospects and prospecting retrospects. *Technol Anal Strateg Manag.* 2003; 15(1): 3-18.
69. Hilgartner S. Capturing the imaginary: Vanguards, visions and the synthetic biology revolution In Miller C, Hagendijk R, Hilgartner S, editors. *Science and democracy: Making knowledge and making power in the biosciences and beyond*. London: Routledge; 2015. p. 51-73.
70. Jasanoff S, Kim SH, editors. *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. Chicago: University of Chicago Press; 2015.
71. Gilbert F, Viaña JNM, Ineichen C. Deflating the “DBS causes personality changes” bubble. *Neuroethics.* 2018. doi: 10.1007/s12152-018-9373-8.
72. Gilbert F, Ovidia D. Deep brain stimulation in the media: Over-optimistic portrayals call for a new strategy involving journalists and scientists in ethical debates. *Front Integr Neurosci.* 2011; 10(5): 16.

73. Gilbert F, Viaña JNM, O'Connell CD, Dodds S. Enthusiastic portrayal of 3D bioprinting in the media: Ethical side effects. *Bioethics*. 2018; 32(2): 94-102.
74. Wade L, Forlini C, Racine E. Generating genius: how an Alzheimer's drug became considered a 'cognitive enhancer' for healthy individuals. *BMC Med Ethics* 2014; 15: 37.
75. Forlini C, Racine E. Added stakeholders, added value(s) to the cognitive enhancement debate: Are academic discourse and professional policies sidestepping values of stakeholders? *AJOB Prim Res*. 2012; 3(1): 33-47.
76. Forlini C, Racine E. Disagreements with implications: Diverging discourses on the ethics of non-medical use of methylphenidate for performance enhancement. *BMC Med Ethics* 2009; 10(1): 9.
77. Racine E, Forlini C. Cognitive enhancement, lifestyle choice or misuse of prescription drugs? Ethics blind spots in current debates. *Neuroethics*. 2010; 3(1): 1-4.
78. Dewey J. *The public and its problems*. Denver: Swallow Press; 1927.
79. Pappas GF. *John Dewey's ethics: Democracy as experience*. Bloomington: Indiana University Press; 2008.
80. Pekarsky D. Dewey's conception of growth reconsidered. *Educ Theory*. 1990; 40(9): 283-94.
81. Gouinlock J. Dewey's theory of moral deliberation. *Ethics*. 1978; 88(1977-1978): 218-28.
82. Evans JH. A sociological account of the growth of principlism. *Hastings Cen Rep*. 2000; 30(5): 31-39.
83. Fiester AM. Weaponizing principles: clinical ethics consultations & the plight of the morally vulnerable. *Bioethics*. 2015; 29(5): 309-15.
84. Racine E. Éthique de la discussion et génomique des populations. *Éthique publique*. 2002; 4(1): 77-90.
85. Doucet H. Les méthodes empiriques, une nouveauté en bioéthique?. *Revista Colombiana de Bioética*. 2008; 3(2): 9-19.
86. Doucet H. Le développement des morales, des législations et des codes, garder le dialogue ouvert et la conscience inquiète. In Hébert A, Doré S, and deLafontaine I, editors. *Élargir les horizons: Perspectives scientifiques sur l'intégration sociale*. Sainte Foy: Éditions Multimondes; 1994. p. 135-41.
87. Jasanoff S. Technologies of humility: Citizen participation in governing science. *Minerva*. 2003;41(3): 223-44.
88. Doucet H. Le développement de la génétique: quelle tâche pour l'éthique? *Isuma*. 2001; 2(3): 38-45.
89. Voarino N, Dubljević V, Racine E. tDCS for memory enhancement: Analysis of the speculative aspects of ethical issues. *Front Hum Neurosci*. 2017. doi:10.3389/fnhum.2016.00678.