

European Journal of Engineering Education

ISSN: 0304-3797 (Print) 1469-5898 (Online) Journal homepage: www.tandfonline.com/journals/ceee20

Fostering responsible anticipation in engineering ethics education: how a multi-disciplinary enrichment of the responsible innovation framework can help

Janna van Grunsven, Taylor Stone & Lavinia Marin

To cite this article: Janna van Grunsven, Taylor Stone & Lavinia Marin (2024) Fostering responsible anticipation in engineering ethics education: how a multi-disciplinary enrichment of the responsible innovation framework can help, European Journal of Engineering Education, 49:2, 283-298, DOI: 10.1080/03043797.2023.2218275

To link to this article: https://doi.org/10.1080/03043797.2023.2218275

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 30 May 2023.

Submit your article to this journal

Article views: 3400

View related articles

View Crossmark data

Citing articles: 17 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=ceee20

EUROPEAN JOURNAL OF ENGINEERING EDUCATION
2024, VOL. 49, NO. 2, 283–298
https://doi.org/10.1080/03043797.2023.2218275

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Fostering responsible anticipation in engineering ethics education: how a multi-disciplinary enrichment of the responsible innovation framework can help

Janna van Grunsven [a], Taylor Stone [b] and Lavinia Marin [a]

[a]Department of Values, Technology and Innovation, TU Delft, Delft, The Netherlands; [b]Institut für Wissenschaft und Ethik (IWE), University of Bonn, Bonn, Germany

ABSTRACT

It is crucial for engineers to anticipate the socio-ethical impacts of emerging technologies. Such acts of anticipation are thoroughly normative and should be cultivated in engineering ethics education. In this paper we ask: 'how do we anticipate the socio-ethical implications of emerging technologies responsibly?' And 'how can such responsible anticipation be taught?' We offer a conceptual answer, building upon the framework of Responsible Innovation and its four core practices: anticipation, reflexivity, inclusion, and responsiveness. We forge a more explicit link between the practices of anticipation, reflexivity, and inclusion, while also enriching them with insights from disability studies, STS, design theory, and philosophy. On this basis we present responsible anticipation as an activity of reflective problem framing grounded in epistemic humility. Via the RI-practice of responsiveness we present responsible anticipation as a creative approach to engineering ethics, offering engineering students a critical yet productive perspective on how ethics may inform innovation.

ARTICLE HISTORY
Received 7 September 2022
Accepted 4 May 2023

KEYWORDS
Responsible innovation; anticipation; engineering ethics education; problem framing; epistemic humility; disability studies

1. Introduction

Through the innovation of emerging technologies, engineers can introduce mass-scale changes in the world with far-reaching socially and ethically disruptive consequences. It is crucial, then, for engineers to anticipate the socio-ethical impacts of their innovations. As many recognise, such acts of anticipation are not value-neutral but thoroughly normative; in anticipation we conjure up possible technological futures and imagine the ways in which an innovation may or may not bring about (un)desirable consequences. As such, anticipation is an activity that should be cultivated in engineering ethics education [EEE]. Indeed, most engineering ethics educators across the globe will likely already incorporate anticipatory activities in their curricula, exploring and debating the socio-ethical implications of emerging innovations such as self-driving cars, exoskeletons, and sex robots. Considered less frequently, though, are the normative criteria that such anticipatory activities themselves are beholden to. Presumably, not all anticipatory explorations are of equal worth. There are better and worse ways to foster anticipation, as an act of relating to our technological future, within our students (Amsler and Facer 2017). Therefore, there is such a thing as good (and bad) anticipation, with 'good' and 'bad' referring broadly to 'responsible' or 'irresponsible'.

CONTACT Janna van Grunsven J.B.vanGrunsven@tudelft.nl

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Taking these presumptions as a starting point, the main question we aim to answer in this paper is: 'what should responsible anticipation of emerging technologies and their socio-ethical implications look like?' Relatedly, we take up a second question: 'how can such anticipation be taught?' We offer a conceptual answer, building upon the field of Responsible Innovation [RI]. Over the past decade, RI has emerged as a guiding approach to technological innovation in the European context. RI understands innovation as a normative endeavour that can introduce, alter, or disrupt the socio-ethical values we hold dear (Stilgoe, Owen, and Macnaghten 2013; van den Hoven 2013). Given the recognition of the importance of RI by governments, industry, and funding agencies, it comes as no surprise that RI's concepts and tools are also increasingly being considered for educational purposes and incorporated into courses and programmes – in particular EEE – at various institutions across the globe (Bergen 2014; Hoople 2014; Robaey 2014; Spruit 2014; Margherita and Bernd 2018; Fischer, Guston, and Trindidad 2019; Mejlgaard et al. 2019; Richter, Hale, and Archambault 2019).[1] The HEIRRI project, for instance, has 'developed ten different programs to introduce RRI in Higher Education through active learning methodologies', highlighting RI's potential for developing creative skills in engineers (Rodriguez et al. 2018, 1257). Similarly, the Nottingham TERRAIN tool for teaching responsible research and innovation offers a rich collection of teaching exercises aimed at promoting proactive or 'upstream' engagement with 'societal issues and concerns to steer or shape innovation pathways' (Hartley et al. 2016). Broadly speaking, RI offers an opportunity to re-position the role of ethics training in engineering and design curricula (Stone, Marin, and van Grunsven 2020a). It moves away from ethics as an external constraint on or limitation to engineering, instead positioning ethical values as 'supra-functional' design requirements that foster creative solutions throughout innovation and design processes (van den Hoven 2013, 2017).

In this paper we contribute to this growing body of literature on RI's potential for EEE, focusing on how the established RI framework, when enriched with multidisciplinary insights, can help elucidate *responsible* anticipation in the context of innovating emerging technologies. Specifically, we will build upon the four dimensions of RI as proposed by Stilgoe, Owen, and Macnaghten (2013) and Owen et al. (2013). These dimensions are: *anticipation*, *reflexivity*, *inclusion*, and *responsiveness*. Focusing on the first three, we propose that an account of responsible anticipation can be arrived at by (1) forging a more explicit link between anticipation, reflexivity, and inclusion as interlocking activities and by (2) enriching the notions of anticipation, reflexivity and inclusion with multidisciplinary insights. We draw heavily from the field of disability studies but also incorporate notions from STS, design theory, and the philosophical sub-field of epistemic injustice. The account that we develop on this basis presents responsible anticipation as an *activity of reflective problem framing grounded in epistemic humility*. Finally, by appealing to the fourth dimension of RI, responsiveness, we present our notion of responsible anticipation as an activity that can be positioned within EEE as a creative forward-looking approach to engineering ethics, providing engineering students with a critical yet constructive perspective on how ethical issues may be incorporated into their innovative endeavours.

Two final points are useful before proceeding. First, there is a family resemblance between the notion of anticipation and the notion of moral imagination, which is frequently identified as a learning goal for EEE. Although the concept of moral imagination has been cashed out in a number of different ways, it arguably captures the ability to 'creatively explore and rehearse alternative courses of actions such that likely outcomes and impacts on others will guide moral decisions' (Narvaez and Mrkva 2014; see Johnson 2014 for a distinctively different account of moral imagination). By contrast, anticipation as presented within the RI framework is distinctive for its future-directedness and its role in coping with the ethics of emerging technologies under conditions of epistemic uncertainty. Admittedly, moral imagination could serve a similar role when applied to contexts of innovation and emerging technologies (Umbrello 2020). However, what we aim to contribute to this literature is a worked-out account of what a responsible usage of moral imagination – or in our language, responsible anticipation – in the context of emerging technologies necessitates. And since we develop this account via an enrichment of the framework of RI it is the language of

anticipation – not imagination – that we use primarily. Second, although the main drive of this paper is theoretical rather than practical, focusing on how the critical task of responsible anticipation can – and should – be conceptualised within the context of EEE, it began as a reflexive process by the authors, who have taught engineering ethics and RI for several years at an institution where EEE has been largely codified. Our institution actively works to refine and improve its education, making EEE a continued topic of critical reflection, positioned as a practice that requires continual re-evaluation within our world of rapid technological change (Van Grunsven et al. 2021). We hope that the reflections offered here provide other EEE practitioners with novel insights about what it means to foster responsible anticipation in their EEE curricula.

2. Responsible anticipation as reflective problem framing

We will begin with a closer look at the notion of anticipation itself, in order to work towards our account of responsible anticipation *as an activity of reflective problem framing grounded in epistemic humility*. Anticipation, as a central pillar of RI, is defined as

> "the forward-looking activity of asking 'what if … ' questions … to consider contingency, what is known, what is likely, what is plausible and what is possible" within the context of innovation processes and concerning innovation's products. (Stilgoe, Owen, and Macnaghten 2013, 19)

This forward-looking activity, in which we consider 'what is known, what is likely, what is plausible, and what is possible', poses a number of challenges. Current criticisms within RI scholarship reveal the inherent tension in the pursuit of forward-looking anticipatory analyses: between the need to formulate rich and robust evaluations of the downstream effects of new innovations, and the importance of avoiding assessments or forecasts that are overly speculative or too detached from reality (Brey 2012). Anticipatory analyses can run the risk of focusing on morally thrilling scenarios – for example the 'grey goo' apocalypse of nanotechnology – that overlook important, if more mundane and obscured, societal impacts (Van de Poel 2016). How, then, do we avoid a narrow focus on sensationalist or dramatic scenarios with a low likelihood (Van de Poel 2016; Van Grunsven 2022)? How do we find a balance between excessively optimistic and pessimistic forms of anticipation (cf. Vallor 2016)? How do we differentiate between anticipations informed by hype and ideology, as is common in Big Tech, versus anticipations that help reveal genuine socio-ethical implications? Making such a differentiation is problematised by an ineluctable epistemic challenge: anticipation is aimed at bringing in view a future that is inherently unpredictable – indeed, any future anticipated may itself undergo radical transformation precisely as a result of the emerging technology under anticipation. Alfred Nordmann worries that 'Trying too hard to imagine possible or plausible futures may diminish our ability to see what is happening' (2014). To this worry we can add that 'what is happening' and what is decided in the here-now can itself be shaped by anticipatory visions of emerging technology, which, when influential enough, affect whether an innovation is embraced or rejected by society (Hilgartner 2015; van de Poel 2020).

Despite these challenges, the demand to anticipate remains. Technological innovation is pervasive in our world and giving up on normative-ethical anticipation in the context of innovation seems like giving up on ethics as a project of aligning the technological world with our values. Thus, when it comes to anticipating the socio-ethical consequences of emerging technologies (such as self-driving cars, CRISPR_Cas9 gene-editing, sex robots, or exoskeletons) we are, on the one hand, confronted with RI's demand to anticipate their societal consequences early in the innovation-process; and on the other hand, with the well-documented epistemic limits of such anticipatory acts.[2] What is needed, then, is an elucidation of the normative criteria for anticipation or benchmarks for success – what, exactly, does good or responsible anticipation look like, and why does it deserve this qualification?

The well-documented epistemic challenges of predicting the socio-ethical consequences of innovative emerging technologies highlight that EEE should *not* position anticipation as a predictive

activity.[3] While anticipation, as described within the RI framework, involves directing our attention to 'what is known, what is likely, what is plausible and what is possible', Owen et al. (2013) simultaneously add the following qualification: 'Tempered by the need for plausibility, such methods do not aim to predict, but are useful as a space to surface issues and explore possible impacts and implications that may otherwise remain uncovered and little discussed' (38). Adding to this, we would like to emphasize that *how* an emerging technology is anticipated (the imagery, values, and desires that are appealed to in sketching the kind of future a technology might enable) are not just non-committal 'what-if' reflections that neutrally 'surface issues', but value-laden activities that can have genuine practical consequences. In Van de Poel's words: 'the different modes of thinking about technology and society are … not innocent: they help to determine not only how we interpret technology and its relation to society but also what we see as possible and desirable' (van de Poel 2020, 500).

At issue here is how anticipation itself *frames* a (proposed) technological innovation and the associated ethical issues at stake. STS scholar Sheila Jasanoff (2003, 240) notes that 'It has become an article of faith in the policy literature that the quality of solutions to perceived social problems depends on the way they are framed. If a problem is framed too narrowly, too broadly, or wrongly, the solution will suffer the same defects'. In *Frame Innovation* (2015), Kees Dorst focuses on the importance of problem framing from the perspective of design thinking (as a broadly construed approach to problem-solving). Dorst explains that any frame relies on concepts with their own meaning – these are not neutral, but will steer explorations and perceptions in certain directions. Crucially, successful frames quickly fade into the backdrop of routine behaviour. In doing so, they become 'limiting rationalities' that may hold back new developments or alternative perspectives (Dorst 2015, 65).

To expand upon this point, consider two brief examples of emerging technologies that might get debated in engineering ethics courses: sex robots and self-driving cars. Sex robots are currently a fringe technology that, according to some, will soon become proliferated into mainstream society (Cf. Levy 2009). The common way of anticipating our future with sex-robots focuses on the potential societal good or harm of sex-robots understood *as quasi-human embodied agents* who capable of (some degree of) self-movement, expressivity, and interaction (Van Grunsven 2022). When seen through this anticipatory lens, the debate often circles around the question whether these quasi-human embodied agents can and will meet the criteria to function as 'good companions' to people who, for a variety of reasons, have trouble finding a human mate to share their life with. Sex-robot companies, such as *Realbotix*, highlight the benefits of an 'infinitely patient embodied robot' for 'the elderly', 'the disabled', and those struggling with trauma (Cf. Coursey et al. 2019). This way of anticipating our future with sex robots can undoubtedly draw students into lively debates in which 'we can ask whether we do, indeed, find it desirable' if sex-robots are introduced as companions for those who may otherwise be excluded from the intimate bond of romantic partnership. But does such a debate instantiate good anticipation? Are we teaching our students the right skills and dispositions needed to take up the task of anticipation critically by inviting them to participate in this debate? Should we not instead encourage our students to reflect on how this way of framing the technology may affect its public perception? And should we not invite students to anticipate how this perception might change under different framing-conditions, for instance, when sex robots are anticipated as data-mongering privacy-sensitive nodes that are inevitably, for their functioning, connected to a wider system of smart objects, sensors and data-driven technologies?

Likewise, the dominant focus on trolley problem-esque scenarios in assessing the ethics of autonomous vehicles has spurred a great deal of discussion and debate over dilemmatic crash scenarios. Yet this way of framing the socio-ethical upshot of autonomous vehicles sets up a limiting rationality that risks obscuring other – and arguably more impactful – ethical issues related to the introduction of autonomous vehicles into society. Decisions made regarding technical developments (and associated policy) such as level of automation and private ownership may have a wide range of impacts on

(sub)urban planning and design (e.g. Duarte and Ratti 2018), ramifications on systemic issues such as transportation justice (e.g. Epting 2019), and can influence seemingly tangential urban problems that *could* be addressed depending on the chosen development path (e.g. Stone, Santoni de Sio, and Vermaas 2020b).

We must therefore help students to cultivate a critical look at these presumed futures, including the effects that acts of anticipation themselves can have on shaping technological and policy development in the here-now. Of course, it is unlikely that anticipatory analyses within the classroom will directly contribute to the societal dominance of a particular way of framing an emerging technology. Still, we maintain that reflecting on the work that anticipation qua framing does is integral to cultivating and habituating responsible anticipation and should be taught to those who may, one day, contribute to the innovations that give shape to society. With this in mind, we can provide a tentative answer to the question, 'when do we anticipate responsibly?' by establishing a boundary condition: *failures* of good anticipation are – at least in part – failures of reflective problem framing. To rephrase this in a more positive sense: responsible anticipation is not about knowing how an emerging technology will affect the future (an impossible epistemic task), but about reflecting upon:

(1) how we are framing the future now,
(2) how those acts of framing are both reflective and formative of socio-ethical saliences, and
(3) how those acts of anticipation have at least the potential of contributing to the actualisation of a certain socio-technical future.

3. Problem framing as grounded in epistemic humility

In trying to identify what it means to promote responsible anticipation in the EEE classroom, we have pointed to the importance of instilling reflective awareness with regards to the problem-framing dimensions of anticipation and the way in which problem frames enact limiting rationalities. What we recognise as morally relevant in our anticipatory endeavours about emerging technologies reflects how we have framed the problems that those technologies purportedly generate or help solve. But this immediately raises a further question: How can we foster such awareness regarding the underlying assumptions, intentions, and ramifications of problem framing, such that it can be cultivated within educational contexts? In this section, we propose that if our anticipations ineluctably enact *limiting rationalities*, as Dorst (2015) argues, then responsible anticipation as reflective problem framing demands a sense of humility regarding the limits of one's anticipatory perspective. We will refer to this as *epistemic humility*. Though underdeveloped, the idea that promoting epistemic humility is much needed in the engineering context is not new. Jasanoff (2003), for instance, argues that to counter-act the uncertainty of the downstream effects as well as the power relations that may arise or be reinforced through processes of innovation, we need to complement existing 'technologies of hubris' with a new anticipatory approach that is explicitly humble in outlook: 'methods … that try to come to grips with the ragged fringes of human understanding – the unknown, the uncertain, the ambiguous, and the uncontrollable. Acknowledging the limits of prediction and control, technologies of humility confront "head-on" the normative implications of our lack of perfect foresight' (Jasanoff 2003, 227).[4] Aligning ourselves with this insight, our main aim in the remainder of this paper is to give this notion of epistemic humility more content. To do so, we draw heavily upon insights from disability studies, while also turning to the philosophical subfield of epistemic injustice.

3.1. Reflexivity as a feature of epistemic humility

In its current format there is already a central, albeit somewhat tacit, role for epistemic humility in the RI framework. Specifically, the dimension of reflexivity can be invoked to counter epistemic hubris. As

Stilgoe et al. characterise it, reflexivity refers to the activity of 'holding a mirror up to one's own activities, commitments and assumptions, being aware of the limits of knowledge and being mindful that a particular framing of an issue may not be universally held' (2013, 20). The main focus here is on *institutional* reflexivity, where 'the value systems and theories that shape science, innovation and their governance are themselves scrutinized' (Stilgoe, Owen, and Macnaghten 2013, 20). To be sure, operationalising such institutional reflexivity is essential in engineering and innovation contexts, where the activities of an individual engineer are embedded within larger institutional systems. An account of what it means to innovate responsibly that places too great of an emphasis on an engineer's individual's responsibility without attending to the wider systemic and institutional structures within which individual engineers functions arguably misses the mark.[5] That said, unless one believes that the relationship between systems, organisations, and institutions and the individual engineers of which they are comprised is one-directional and deterministic, the question of what we can expect from individual engineers and which moral competencies we want them to develop remains. At the very least, we believe that the project of promoting institutional reflexivity stands to benefit from epistemically humble individuals who are disposed to 'holding up a mirror to [their] own activities, commitments and assumptions' and who are 'aware of the limits of knowledge'. Thus, we propose that it is equally important to spell out what epistemic humility might require of individuals and, relevant for EEE, what it would mean to foster epistemic humility in individuals.

As described in the introduction, RI understands innovation as a normative endeavour that can introduce, alter, or disrupt the socio-ethical values we hold dear (Stilgoe, Owen, and Macnaghten 2013; van den Hoven 2013). A commitment to reflexivity, however, goes beyond the safeguarding and operationalising of those values through innovation. Reflexivity demands a readiness to scrutinise presupposed values and to expose their effects on how we anticipate the impacts of a given technology. The central importance of a reflexive scrutinisation of our values is powerfully argued for by philosopher and disability studies scholar Eva Kittay. She warns that:

> 'we need to be alert to the possibility that the values we hold dear blinker us and allow us to presume that these values *must* have the same importance for others … When we pay little heed to what others have to say about what they believe to be important, create hierarchies in which our own values always trump those of another, or unreflectively rely on such hierarchies when we appeal to 'what is evident' or what is 'surely' the case, then we act out of hubris. While we cannot help but make appeal to our own values and perspectives, we need to pay close attention to the role these are playing and not presume our … argumentation is untouched by the importation of such values'. (2008, 231)

Kittay is addressing a certain type of moral philosopher (or bioethicist) here. She is concerned with philosophers who are inclined to believe that they, as skilled experts in detached observation and rational argumentation, are in a privileged position to identify what lives are worth living and who deserves (or doesn't deserve) our moral regard. For Kittay, the stakes for exposing and battling the epistemic hubris operative in this type of philosopher are deeply personal. As a mother of a severely cognitively disabled woman she experiences philosophy as a 'battleground' on which she is fighting for the moral visibility of her daughter, and others like her, who, from the philosophical armchair, are frequently categorised as beings who fail to meet a particular threshold for moral personhood (Kittay 2009).

Kittay argues that this dehumanising stance stems directly from epistemic hubris or arrogance. Among the philosophers with whom she is in battle, Kittay notes a pervasive resistance towards the idea that opening oneself up to genuine interaction with a severely cognitively disabled person could reveal something new and worthwhile about one's (hierarchical) value-commitments and one's views on what makes a person's life worth living. This is a case of intellectual myopia: a particular conception of philosophical expertise, namely expertise in detached observation and rational argumentation, dismisses out of hand the idea that one can obtain philosophically relevant knowledge through emotionally engaged interactions such as the ones Kittay has with her daughter.[6] By denying the epistemic relevance of engaged interaction, this intellectual myopia is reinforced; for it is precisely in the arena of genuine unpredictable interaction that one's epistemic

assumptions are most likely to be dislodged – that reflexivity is triggered and epistemic humility is experienced. Kittay describes how, for instance, in her interactions with her daughter:

> I am often surprised to find out that Sesha has understood something or is capable of something I did not expect. These surprises can only keep coming when she and her friends are treated in a manner based not on the limitations we know they have *but on our understanding that our knowledge is limited*. (2009, 619, our italics)

To advocate for the epistemic value of such humility, and to underscore that it is plainly false that a stance of detached observation and rational argumentation is epistemically superior, Kittay bolsters her argument with examples from the natural sciences. Highlighting 'the close personal attentiveness' and 'feeling' with which nobel-winning scientist Barbara McClintock and primatologist Jane Goodall attended to 'the entities that [they] studied', Kittay argues that 'the value of … interaction with the individuals studied' is in part that it gives 'rise to perceptual capabilities that are not shared by those who have at best a glancing acquaintance' and 'often fail to get a glimpse into the lives of these [beings]' (Kittay 2009, 406–407). It was in virtue of this attentiveness that McClintock 'made startling discoveries concerning the transmission of genetic material in maize', and that Goodall profoundly deepened our understanding of the capabilities and behaviours of chimpanzees (2009, 406).

Examples such as these, which help expose the epistemic hubris and myopia that can follow from certain views about detached observation, rationality and their presumed epistemic superiority, are useful in the context of EEE. It is not uncommon for engineering students to reason that their (developing) technical expertise provides them with a more 'rational' and detached, *and therefore* more legitimate, perspective on the socio-ethical values and implications at stake in innovation than lay-persons. As Sabine Roeser has argued, this often stems from engineering students' self-conception as 'the archetype of people who make decisions in a rational and quantitative way', and who, by reasoning as 'unemotional calculators' are free from the emotional biases that allegedly cloud the evaluative abilities of laypeople (2012). It is true that having a grounded understanding of the technological facts *can* help identify relevant ethical issues, as it will play a role in anticipation processes. For instance, to harken back to an earlier example, a solid technological understanding of state-of-the-art robotics and AI might offer critical ammunition against the dramatic anticipation that sex robots will soon be able to function in a manner analogous to human partners, such that they could help solve the socio-ethical problem of mass-scale loneliness (McClelland 2017). Yet, EEE should help students identify that they commit a technocratic 'is/ought' fallacy by assuming that their understanding of the technological facts (themselves framed by discipline-specific norms and practices) translates into a privileged appreciation of the normative upshots of those facts (see also Roeser 2010; Van Grunsven et al. 2023). For example, understanding the scientific mechanisms underlying CRISPR_Cas9 gene-editing technology does not by itself translate into better judgments about when it is appropriate to apply the technology in order to further patients' health. Health is a normative concept, the meaning of which is as much if not more illuminated via the lived perspectives of experts-by-experience than it is from the perspective of clinical experts and technologists. If left unchallenged, this kind of technocratic 'is/ought' fallacy could lead to a sense of justified epistemic hubris, where students 'do not know what they do not know, nor do they appear to take any concrete steps to rectify the situation, because they presume that they have nothing to learn that is of moral significance' (Kittay 2009, 619).

In the context at issue here, which concerns the fostering of responsible anticipation in engineering students, the form of epistemic hubris just sketched can be characterised as an explicit one, where engineering students explicitly take their (developing) expert knowledge as providing them with a superior anticipatory understanding of the socio-ethical values (presumably) implicated by a given innovation. In addition to this explicit form, epistemic hubris can also come in an implicit form. In this form, the meaning of the ethical values that are targeted in, or that frame, acts of anticipation are unreflectively taken by engineering students as unproblematically given. To give an

example: in highlighting the desirability of a newly emerging technology, there is a recurring pattern of appealing to how technology X can help promote the health and well-being of disabled people. Such appeals are made in sex robot and gene-editing anticipations, but they are equally at play in discourses surrounding exoskeletons, care robots, self-driving cars and other emerging technologies. STS and Disability Studies scholar Ashley Shew has termed this tendency *technoableism*, which, she writes, 'describe[s] a particular strain of ableism that I have witnessed in the context of imagination, technology, and bodies. … Technoableism … at once talks about empowering disabled people at the same time reinforcing ableist tropes about what body-minds are good to have and who counts as worthy' (2020, 43) Shew emphasizes that these technologies and the anticipatory narratives in which they are embedded are often developed by well-meaning designers and engineers who are entirely unaware of their ableist assumptions and biases: 'technoableists usually think they have the good of disabled people in mind. They do not see how their work reinscribes ableist tropes and ideas on disabled bodies and minds' (2020, 43). When it is implicitly taken that uprightness is essential for a person's physical well-being and the preferred human way of moving through the world, then exoskeletons will likely be anticipated as a noble innovation aimed at alleviating the alleged burdens associated with certain mobile disabilities. Never mind that many wheelchair users do not see exoskeletons as a desirable solution to some of the mobility-challenges they confront:

> the full story of exoskeletons has yet to be written, but the existing narratives around paralysis (framed as an awful fate) and walking (assumed to be the best way of moving through the world) almost seem designed to end with exoskeletons as solutions to 'the problem' of mobility impairment – a technoableist's dream solution. Yet exoskeletons aren't the 'cure-all' imagined by the news stories in which they are featured. And actual wheelchair users – the people for whom the devices were ostensibly developed – point out that non-disabled people are often wrong about what is bad about being a wheelchair user. (Shew 2020, 46)

The assumptions at work here are less overt than in explicit instances of epistemic hubris. In implicit instances of epistemic hubris, it is not the case that engineers (in training) *take* their expertise to place them in a position of superiority with respect to the relevant socio-ethical implications of an innovation; rather, it is that they accept, unreflectively, their interpretation of the socio-ethical implications and values at stake in the acts of anticipation animating their innovative endeavours. In the explicit form, alternative perspectives, particularly of laypeople and experts-by-experience, are explicitly diminished or dismissed; in the implicit form the need to seek out alternative perspectives is overlooked, because the meaning of the values at stake is taken as uncontroversial and requiring no reflection. In terms of consequences, this implicit epistemic hubris is no less pernicious; when innovations are designed on the basis of (for instance) ableist assumptions these innovations can end up reinforcing precisely the inequities that they were meant to help mitigate.[7] Both forms of epistemic hubris, the dislodging of which is the target of reflexivity, point to the importance of practices and interactions that promote *inclusion*.

3.2. Inclusion as a feature of epistemic humility

In the RI framework that we are building upon, inclusion refers to

> processes of democratization in innovative endeavors; processes that call into question traditional hierarchical division of power between experts and laypersons and that make room 'for public and stakeholder voices to question the framing assumptions' guiding innovation processes. (Stilgoe, Owen, and Macnaghten 2013, 21)

How such inclusion ought to be operationalised in innovation processes is a real question. As Carolyn ten Holter has recently argued, RI ought to pay more attention to the political power dynamics that affect successful inclusion:

> The question '*who participates?*' … is significant in terms of questions of power, … The explicit concern with dominance patterns could be centralised by RRI to better engage with the democratic principles that are necessarily a part of participatory methodologies. … If participatory methodologies of innovation have an inherently

> political position that reserves the right to challenge the status quo, then RRI is necessarily … political at its core … but this is little discussed. RRI may need to foreground this political challenge in a more conscious way and create a means to address these political challenges more directly. (2022)

As we saw with Shew, one dominant pattern particularly rampant in contexts of innovation and technology development is that of technoableism. Technoableism, we saw, is marked by a tendency to frame the appeal of hyped newly emerging technologies around their alleged benefits for 'the disabled'. Strikingly, as Shew but also Kittay and many other disability studies scholars have shown, these claims are often made without any genuine inclusion of the perspectives of disabled people. The obvious first step towards inclusion is for technologists to stop working on the basis of what their imagination tells them disabled people need and instead to 'Simply ask disabled people what kinds of technologies we want and need' (Shew 2020). We propose that this call applies equally to anticipatory activities in the classroom, helping to instil a humble awareness in developing professionals with respect to the limiting rationality at play in their anticipatory stance and the ways in which they might implicitly accept potentially pernicious value-assumptions. This can mean inviting guest speakers, encouraging students to look for first-personal testimonials of marginalised direct and indirect stakeholders, and, if done with the right sensitivity and supervision, creating opportunities for students to engage in real-life sustained interactions with under-represented stakeholders.

It is important to highlight, however, that getting the right stakeholders around the table is a necessary but not sufficient step towards promoting genuine inclusion. This is the case for several reasons. Firstly, it is always possible –and possibly tempting–for (developing) professionals to collaborate with experts-by-experience who will confirm their implicit value assumptions. As Shew points out: 'ableism is a system of value that all of us participate in, including individual disabled people' (2020, 47). While this would result in collaborations that count as co-anticipations on the surface, they likely fail to promote reflexivity and to challenge attitudes of epistemic hubris.

Secondly, there are numerous ways in which people can be undervalued as knowers and communicators even while they are included in research and innovation processes. To explicate this, it can be helpful to bring in insights from the field of Epistemic Injustice [EI]. As EI has revealed, marginalised stakeholders seeking inclusion from within a system characterised by asymmetrical power relations–such as an ableist value system–tend to live under conditions of testimonial and hermeneutic injustice (Fricker 2007). *Testimonial injustice* or oppression occurs when, as a result of pernicious stereotyping, someone is discredited as a reliable giver of testimony–as someone whose voice matters and whose experiences ought to be taken seriously. Such stereotyping often operates in unintentional habitual ways, with implicit ableist, sexist, racist and other biases quietly affecting a hearer's openness to a speaker's testimony (Dotson 2011). *Hermeneutic injustice* occurs when dominant discourses and concepts make the articulation of one's experiences as a member of a marginalised group difficult or impossible. For instance, for many autistic people, 'self-stimulatory' or 'stimming' behaviours–such as flapping, rocking, or humming–are richly meaningful and expressive ways of engaging their environment (Kapp, Steward, and Crane 2019). However, the dominant clinical way of framing stimming has been as pathological non-expressive problem behaviour (Van Grunsven and Roeser 2022). This dominant discourse around stimming has hindered autistic people's access to concepts that can help them articulate and convey to others the rich experiential meaning that stimming has for them. As emphasized by the field of EI, these forms of epistemic injustice can occur without the presence of explicit intentional harm caused by a conscious agent. Testimonial injustice often stems from implicit biases and hermeneutic injustices from institutionalised practices and interpretative schemes, making it harder to identify their negative impact on efforts towards inclusion.

Thus, in order to give the voices of direct stakeholders or experts-by- experience a prime position in innovation and anticipation processes, we stress that these processes must involve a continual reflection on (1) the criteria we use to select representatives of marginalised stakeholders and (2)

the pervasive ways that overarching pernicious value-systems, such as ableism, sexism, racism, and ageism, impact on those processes, as these hierarchical value-systems contribute to testimonial and hermeneutic forms of injustice that can limit the extent to which stakeholders can meaningfully participate and express their experiences.

3.3. Interlocking and operationalizing anticipation, inclusion, and reflexivity

Having discussed anticipation, reflexivity, and inclusion separately, we will now discuss how they operate as *interlocking* activities for promoting responsible anticipation as reflexive problem framing grounded in epistemic humility. An awareness of anticipation as a framing activity that inevitably enacts a limiting rationality motivates the need for reflexivity – for considering our own beliefs, commitments, and values, as well as feeling a sense of humility with the limits of our own perspective. Reflexivity, in turn, thrives under conditions of inclusion. As Kittay showed us, genuine inclusion and interactive engagement with (marginalised) others are powerful, perhaps even necessary, means for interrogating the value-commitments animating our acts of anticipation. But genuine inclusion, as the EI field underscores, is difficult if not impossible to achieve without reflexively holding up a mirror to our own beliefs, commitments, and values, which may be permeated by pernicious biases and hierarchies.

It is by moving back and forth between the interlocking dimensions of RI, enriched with insights from disability studies, EI, STS and design research, that technology anticipations can be critically dissected and re-evaluated. To use the example of sex-robots, rather than delving straight into a debate centred around their alleged ability to solve problems of societal loneliness by functioning as quasi-autonomous responsive human-esque agents, EEE educators could foster good anticipation as follows:

(1) *Stage 1: Set up Anticipatory exercises of reflective problem framing*. The focus would be on attending to the way in which sex robots are predominantly framed in academic and popular literature. Salient images, metaphors and arguments would be identified as well as the ways in which these images, metaphors and arguments are prone to impact on our thinking about the socio-ethical (un)desirability of this emerging innovation. Once this is articulated, EEE educators would
(2) *Stage 2: Encourage reflexivity*. Whereas the aim of stage 1 is for students to attend to the framing of a given innovation, the aim in stage 2 is for students to attend to their own intuitions, beliefs, and value-commitments. Students would explore, for instance, how their situated perspective reflects or helps to disrupt explicit and implicit forms of epistemic hubris operative within the framed anticipation that is under examination.
(3) *Stage 3 Support processes of Inclusion*. Stage 3 moves past the inward self-reflexive turn of stage 2, with students being encouraged to seek out alternative voices, particularly of marginalised direct stakeholders who have a stake in how a particular emerging innovation is anticipated. In doing so, students would widen, challenge, and nuance their own situated perspective, scrutinised via reflexivity, through processes of inclusion. All the while, they would be encouraged to bear in mind the dangers of epistemic forms of injustice that negatively impact upon these processes.

Thus, insights gathered through stage 3 circle back into students' reflexivity, while the reflexive awareness of epistemic hubris, obtained through stage 2, would stand in the service of genuine processes of inclusion in stage 3. Throughout, the outcome of these stages is looped back into how students evaluate the technology under anticipation, as illustrated below in Figure 1.

As such, the three stages described above are best understood as iteratively interwoven. At the same time, our linear presentation of the stages of responsible anticipation (as reflexive problem framing grounded in epistemic humility) has the benefit of lending itself well to the design of an EEE course or activity centred around an emerging technology.

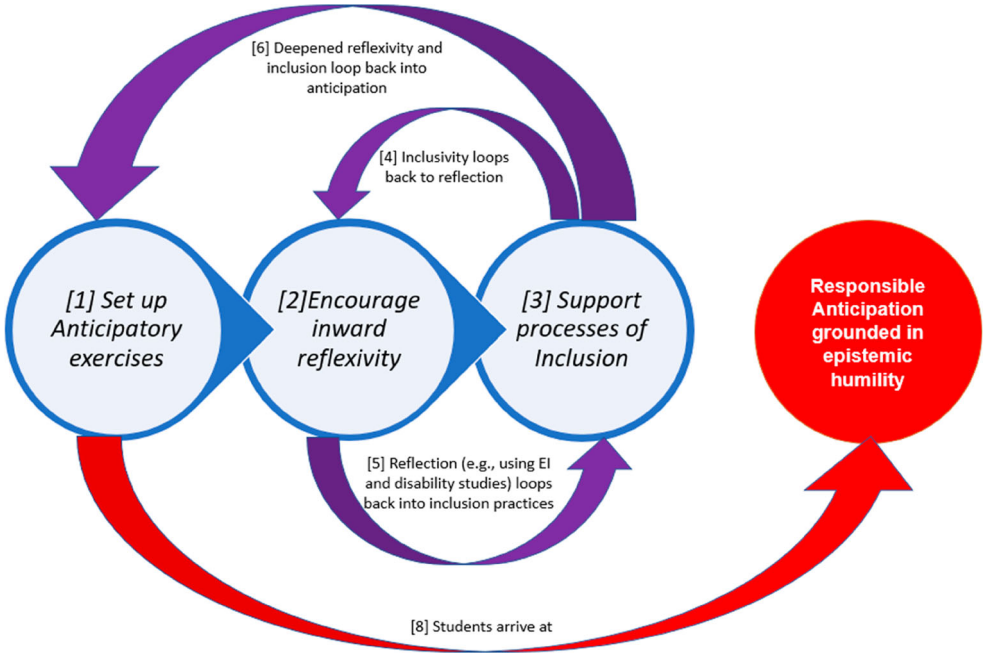


Figure 1. Diagram of how the three stages, iteratively interlocking, contribute to a process of responsible anticipation grounded in epistemic humility.

4. Responsiveness: reflectively and humbly moving ahead

At this point, one might worry that the emphasis on epistemic humility as promoted through reflexivity and inclusion undermines engineering student's confidence to innovate. Students might feel that, if epistemic hubris is constantly lurking around the corner, leading them to entirely misguided solutions to problems they have potentially misconstrued ('exoskeletons as [misguided] solutions to "the [misguided] problem" of mobility impairment', to recite Shew), they may feel it is better not to take on the messy, challenging, epistemically fraught endeavour of anticipation and, correspondingly, innovation.

On the one hand we, as ethics educators, embrace the possibility that responsible anticipation can foster a humble disposition not only with respect to how we frame particular innovations but also to how we frame innovation as such; a disposition that includes a willingness to critically reflect on the blind faith that new technologies progressively lead towards a better, more ethical world (Cf. Blok and Lemmens 2015; Wisnioski, Hintz, and Kleine 2019; Marin and Steinert 2022). One the other hand, if one accepts that many of today's grand challenges require technology-driven innovations, it is equally important to situate ethics in engineering education not as a limiting obstacle, but as a motivational source for moving ahead. To this end, RI can be fruitful in two ways. Firstly, when processes of inclusion are taken seriously as a dimension of good anticipation, then taking up the epistemically and ethically messy challenge of anticipation is never one that should be undertaken alone. The burden of responsible anticipation is, by our account, something that is constitutively participatory and shared. We all occupy a finite situated perspective that enacts limiting rationalities, but under conditions of inclusion we can see more and arrive at better ways to give material shape to our world.

Secondly, we argue for the importance of 'framing' the outcomes of epistemic humility as opportunities for novel anticipations that can lead to creative new solutions. It is at this stage that responsiveness as a dimension of RI can be brought in. In the field of RI, responsiveness

refers to the 'capacity to change shape or direction in response to stakeholder and public values and changing circumstances. … Responsiveness involves responding to new knowledge as this emerges and to emerging perspectives, views and norms' (Stilgoe, Owen, and Macnaghten 2013, 21). Critically, then, the cultivation of epistemic humility not only serves as a countermeasure to hubris; it can also serve as a basis for creativity and exploration. A humble responsive outlook can open us up to new possibilities and perspectives, through what Shannon Vallor (2016, 127) describes as 'a reasoned, critical but *hopeful* assessment of our abilities and creative powers, combined with a healthy respect for the unfathomed complexities of the environments in which our powers operate and to which we will always be vulnerable'. Similarly, in their proposed educational approach of *Creative Anticipatory Ethical Reasoning*, York and Conley (2020) see the development of a critical – but imaginative – outlook as crucial. It allows students to confront their values and assumptions, shifting from the 'ideal or promise of a technology to a technology-in-context that might raise unexpected but plausible additional scenarios which can then be the focus of ethical reasoning' (York and Conley 2020, 7). Thus, epistemic humility in the context of RI plays the role of balancing a critical and reflexive acknowledgement of one's limited perspective with a hopeful, creative, and exploratory outlook regarding if (and how) identified problems can be addressed.

At our institution, we increasingly see that many of our engineering students embrace the idea that society's most urgent ethical challenges can and should be solved via newly emerging technological innovations. This arguably signifies a cultural shift in how engineering students understand themselves and their role in society, moving away from an understanding of engineering as an apolitical value-neutral activity (Cf. Cech 2013). Though this shift may vary across institutions (Jamison, Kolmas, and Holgaard 2014), it correlates with a wider tendency in Big Tech that has grown in recent years, namely to frame innovation as emphatically socio-ethically motivated. In the words of Jill Lepore (2021, Episode 1):

> Tech companies started to talk about their mission, and their mission was always magnificently inflated: 'transforming the future of work,' 'connecting all of humanity,' 'making the world a better place,' … Companies worry very publicly and quite feverishly about planetary disaster, about the all too real catastrophy of climate change, … and all sorts of existential risks to the future of the human race … so that they can save us all.

As indicated by Lepore's tone in this passage, and echoing worries captured in Shew's notion of inflated technoableist narratives, anticipations surrounding the socio-ethical implications of innovation are likely to function as a double-edged sword in EEE. On the one hand, pedagogical activities built around anticipation can tap into an existing trend among engineering students to embrace the socio-ethical potential of innovation. Thus, it can be utilised to 'meet students where they are' (Sunderland 2014), serving as a jumping board for refining students' grasp of the socioethical dimensions of innovation while doing so in a solution-oriented, creative, forward-looking way. At the same time, it is always tempting, in our anticipations, to be gripped by riveting, hyped, sensationalised future scenarios – scenarios that frame the socio-ethical potential and dangers of an emerging technology that are inevitably reflective of a limiting rationality. Too often, these limited rationalities harbour taken-for-granted hierarchical value-systems that steer our thinking about the value of a certain innovation in ways that are ethically problematic. As such, EEE must confront the challenge of fostering not just any form of anticipation, but *responsible* anticipation. As we have argued, that means fostering anticipation as reflective problem framing grounded in epistemic humility.

Notes

1. At our own institution, Delft University of Technology, RI is increasingly incorporated into ethics training at the MSc and Ph.D. levels through stand-alone courses, workshops, and seminars. At the BSc level, RI is embedded in ethics learning lines across numerous engineering and design curricula as well as offered as a minor specialisation.

2. Another response is to this challenge is to abandon prediction or anticipation as a primary goal, and seek out alternative frameworks – for example, to see new technologies as social experiments and adopt a precautionary, incremental approach (Van de Poel 2016). Such cautionary approaches can provide a powerful counterweight to technological enthusiasm, and thus are useful resources for teaching EEE. Still, for reasons we discuss throughout the paper, we believe anticipation (and the question of what makes it good or bad) should play a central role in EEE.
3. As Amsler and Facer (2017) discuss, there are not just epistemic but also ethical and political reasons for resisting notions of anticipation as prediction.
4. In *Technology and the Virtues* (2016) Shannon Vallor posits humility as a key technomoral virtue for the twenty-first century. Similar to Jasanoff (2003), Vallor defines humility as 'a recognition of the real limits of our techno-social knowledge and ability; reverence and wonder at the universe's retained power to surprise and confound us; and renunciation of the blind faith that new technologies inevitably lead to human mastery and control of our environment' (126–127). As such, epistemic humility also promotes a more critical and reflective engagement with the unbridled optimism that often animates technological innovation.[4]
5. Thanks to an anonymous reviewer or encouraging us to emphasize this point.
6. See McGeer (2009) for a similar accusation of intellectual myopia, this time with respect to how a certain (dominant) way of framing autism dismisses, on the basis of its own theoretical assumptions, precisely those first-person testimonials from autistic people that directly challenge those assumptions.
7. It is a consequence of our emphasis on reflexive epistemically humble anticipation that it cannot be fully settled in advance which substantive values (and which conception of the meaning of those values) are the most relevant one's to uncover and prioritise within a given innovation context. Presuming, for instance, that health, or well-being, or safety, are *the* values to focus on (and presuming that we already have a clear handle on how those values ought to be cashed out) sets one on a path toward epistemic hubris. That said, our account does postulate inclusivity and humility as cross-contextual procedural values that must animate any context-sensitive activity of retrieving and anticipating substantive values.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the 4TU Centre for Engineering Education [grant number TBM_ERE_2021_02_4TU.CEE.-TUD] & the NWO Dutch Research Council, which funded the projects: Mattering Minds: Understanding the Ethical Lives of Technologically Embedded Beings with 4E Cognition [grant number VI.Veni.211F.055] & the Gravitation research programme Ethics of Socially Disruptive Technologies [grant number 024.004.031].

Notes on contributors

Janna van Grunsven is an assistant professor in TU Delft's ethics and philosophy of technology section. She conducts research at the intersection of embodied cognition, philosophy of technology, engineering ethics (education), and disability studies. In a project funded by the Dutch Research Council (NWO), entitled Mattering Minds: Understanding the Ethical Lives of Technologically Embedded Beings with 4E Cognition, she examines how different theoretical accounts of the mind and different technological developments can have decisive ethical implications for how disabled people are brought in view in a moral sense. Additionally, she is involved in COMET; a project on experiential engineering ethics education funded by the 4TU Centre for Engineering Education. Her work has appeared in journals such as Advances in Engineering Education, Social Epistemology, Ethics and Information Technology, and Techné: Research in Philosophy and Technology.

Taylor Stone is a Senior Researcher at the Institute for Science and Ethics (IWE), University of Bonn. He received his PhD in Ethics of Technology from TU Delft (2019), and has held postdoctoral and lecturer positions in industrial design, ethics of technology, and engineering ethics education. His research focuses on the ethics of (urban) technologies and how to incorporate environmental values into design processes and built artifacts.

Lavinia Marin is an assistant professor at the Ethics and Philosophy of Technology Section, TU Delft, the Netherlands. Her current research investigates the conditions of possibility for epistemic and moral agency (both at the individual and group level) for users of social networking platforms using approaches from ethics, social epistemology, and situated cognition. She is involved in several collective research projects, such as ESDIT and COMET - a project on experiential engineering ethics education funded by the 4TU Centre for Engineering Education. She is a member of the Delft Digital Ethics Centre.

ORCID

Janna van Grunsven http://orcid.org/0000-0002-8751-1289
Taylor Stone http://orcid.org/0000-0002-3546-2577
Lavinia Marin http://orcid.org/0000-0002-8283-947X

References

Amsler, S., and K. Facer. 2017. "Contesting Anticipatory Regimes in Education: Exploring Alternative Educational Orientations to the Future." *Futures* 94: 6–14. doi:10.1016/j.futures.2017.01.001.

Bergen, Jan Peter. 2014. "On Engineers Engaging Ethics Through Dis-location and Reconnection." *Journal of Responsible Innovation* 1 (2): 242–244. doi:10.1080/23299460.2014.922342.

Blok, Vincent, and Pieter Lemmens. 2015. "The Emerging Concept of Responsible Innovation. Three Reasons Why it is Questionable and Calls for a Radical Transformation of the Concept of Innovation." In *Responsible Innovation 2*, edited by B. J. Koops, I. Oosterlaken, H. Romijn, T. Swierstra, and J. van den Hoven. Springer.

Brey, Philip. 2012. "Anticipatory Ethics for Emerging Technologies." *Nanoethics* 6 (1): 1–13. doi:10.1007/s11569-012-0141-7.

Cech, Erin A. 2013. "The (mis) Framing of Social Justice: Why Ideologies of Depoliticization and Meritocracy Hinder Engineers' Ability to Think About Social Injustices." In *Engineering Education for Social Justice*, edited by Juan Lucena, 67–84. Dordrecth: Springer.

Coursey, Kino, Susan Pirzchalski, Matt McMullen, Guile Lindroth, and Yuri Furuushi. 2019. "Living with Harmony: A Personal Companion System by Realbotix™." In *AI Love You*, edited by Yuefang Zhou and Martin H. Fischer, 77–95. Cham: Springer.

Dotson, K. 2011. "Tracking Epistemic Violence, Tracking Practices of Silencing." *Hypatia* 26 (2): 236–257. doi:10.1111/j.1527-2001.2011.01177.x.

Dorst, Kees. 2015. *Frame Innovation: Create New Thinking by Design*. Cambridge: The MIT Press.

Duarte, Fábio, and Carlo Ratti. 2018. "The Impact of Autonomous Vehicles on Cities: A Review." *Journal of Urban Technology* 25 (4): 3–18. doi:10.1080/10630732.2018.1493883.

Epting, Shane. 2019. "Automated Vehicles and Transportation Justice." *Philosophy & Technology* 32 (3): 389–403. doi:10.1007/s13347-018-0307-5.

Fischer, Erik, David Guston, and Brenda Trindidad. 2019. "Making Reponsible Innovators." In *Does America Need More Innovators?*, edited by Matthew H. Wisnioski, Eric S. Hintz and Marie S. Kleine, 345–366. Lemelson Center Studies in Invention and Innovation Series. Cambridge: The MIT Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. New York: Oxford University Press.

Hartley, S., W. Pearce, C. McLeod, B. Gibbs, S. Connelly, J. Couto, T. Moreira, et al. 2016. *The TERRAIN Tool for Teaching Responsible Research and Innovation*. Nottingham: University of Nottingham.

Hilgartner, Stephen. 2015Science and Democracy Making Knowledge and Making Power in the Biosciences and Beyond." In *Science and Democracy Making Knowledge and Making Power in the Biosciences and Beyond*, edited by Stephen Hilgarten, Clark Miller, and Rob Hagendijk, 33–55. New York: Routledge.

Hoople, Gordon. 2014. "Engineering Ethics in Every Decision." *Journal of Responsible Innovation* 1 (2): 241–242. doi:10.1080/23299460.2014.922341.

Jamison, A., A. Kolmas, and J. E. Holgaard. 2014. "Hybrid Learning: An Integrative Approach to Engineering Education." *Journal of Engineering Education* 103: 253–273. doi:10.1002/jee.20041.

Jasanoff, Sheila. 2003. "Technologies of Humility: Citizen Participation in Governing Science." *Minerva* 41: 223–244. doi:10.1023/A:1025557512320.

Johnson, M. 2014. *Moral Imagination: Implications of Cognitive Science for Ethics*. Chicago: University of Chicago Press.

Kapp, K. S., R. Steward, and L. Crane. 2019. "'People Should be Allowed to Do What They Like': Autistic Adults' Views and Experiences of Stimming." *Autism* 23 (7): 1782–1792. doi:10.1177/1362361319829628.

Kittay, Eva Feder. 2008. "Ideal Theory Bioethics and the Exclusion of People with Severe Cognitive Disabilties." In *Naturalized – Toward Responsible Knowing and Practice*, edited by H. Lindemann, M. Verkerk and M. Urban Walker. Cambridge: Camrbidge University Press.

Kittay, Eva Feder. 2009. "The Personal is Philosophical is Political: A Philosopher and Mother of a Cognitively Disabled Person Sends Notes from the Battlefield." *Metaphilosophy* 40 (3-4): 606–627. doi:10.1111/j.1467-9973.2009.01600.x.

Lepore, Jill. 2021. "Elon Musk: The Evening Rocket." *Season One Episode One*, November 29. Pushkin. Accessed August 31, 2022. https://www.pushkin.fm/podcasts/elon-musk-the-evening-rocket.

Levy, David. 2009. *Love and Sex with Robots: The Evolution of Human-Robot Relationships*. New York: HarperCollins.

Margherita, Nulli, and Stahl Bernd. 2018. "RRI in Higher Education." *The ORBIT Journal* 1 (4): 1–8. doi:10.29297/orbit.v1i4.78.

Marin, Lavinia, and Steffen Steinert. 2022. "Twisted Thinking: Technology, Values and Critical Thinking." *Prometheus* 38 (1). doi:10.13169/prometheus.38.1.0124.

McClelland, Richard T. 2017. "Confronting Emerging New Technology." *The Journal of Mind and Behavior* Summer and Autumn 2017, Volume 38, Numbers 3 and 4: 247–270.

McGeer, Victoria. 2009. "The Thought and Talk of Individuals with Autism: Reflections on Ian Hacking." *Metaphilosophy* 40 (3-4): 517–530. doi:10.1111/j.1467-9973.2009.01601.x.

Mejlgaard, Niels, Malene Vinther Christensen, Roger Strand, Ivan Buljan, Mar Carrió, Erich Griessler Marta Cayetano i Giralt, et al. 2019. "Teaching Responsible Research and Innovation: A Phronetic Perspective." *Science and Engineering Ethics* 25 (2): 597–615. doi:10.1007/s11948-018-0029-1.

Narvaez, D., and K. Mrkva. 2014. "The Development of Moral Imagination." In *The Ethics of Creativity*, edited by S. Moran, D. Cropley, and J. C. Kaufman, 25–45. Palgrave Macmillan/Springer Nature. doi:10.1057/9781137333544.0007.

Owen, Richard, Jack Stilgoe, Phil Macnaghten, Mike Gorman, Erik Fisher, and Dave Guston. 2013. "A Framework for Responsible Innovation." In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, Vol. 31, edited by Richard Owen, John Bessant, and Maggy Heitz, 27–50. Wiley.

Richter, Jennifer, Annie E. Hale, and Leanna M. Archambault. 2019. "Responsible Innovation and Education: Integrating Values and Technology in the Classroom." *Journal of Responsible Innovation* 6 (1): 98–103. doi:10.1080/23299460.2018.1510713.

Robaey, Zoë. 2014. "A Commentary on Engineering Ethics Education, or How to Bring About Change Without Needing Scandals." *Journal of Responsible Innovation* 1 (2): 248–249. doi:10.1080/23299460.2014.922345.

Rodriguez, Gemma, Núria Saladie, Gema Revuelta, Clara Vizuete, Carolina Llorente, and Mar Carriò. 2018. "Responsible Research and Innovation: An Opportunity to Develop Creative Skills at Higher Education." In *4th International Conference on Higher Education Advances (HEAd'18)*. València: Universitat Polite`cnica de València. doi:10.4995/HEAd18.2018.8187.

Roeser, Sabine. 2010. "Intuitions, Emotions and Gut Reactions in Decisions About Risks: Towards a Different Interpretation of 'Neuroethics'." *Journal of Risk Research* 13 (2): 175–190. doi:10.1080/13669870903126275.

Shew, Ashley. 2020. "Ableism, Technoableism, and Future AI." *IEEE Technology and Society Magazine* 39 (1): 40–85. doi:10.1109/MTS.2020.2967492.

Spruit, Shannon. 2014. "Responsible Innovation Through Ethics Education: Educating to Change Research Practice." *Journal of Responsible Innovation* 1 (2): 246–247. doi:10.1080/23299460.2014.922344.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. "Developing a Framework for Responsible Innovation." *Research Policy* 42 (9): 1568–1580. doi:10.1016/j.respol.2013.05.008.

Stone, T., L. Marin, and J. van Grunsven. 2020a. "Before Responsible Innovation: Teaching Anticipation as an Intellectual Virtue for Engineers." In *Engaging Engineering Education: Proceedings of the 48th Annual Conference of the European Society or Engineering Education (SEFI)*, edited by J. van der Veen, and H.-M. Järvinen, 1401–1408. ISBN: 978-2-87352-020-5.

Stone, T., F. Santoni de Sio, and P. Vermaas. 2020b. "Driving in the Dark: Designing Autonomous Vehicles for Reducing Light Pollution." *Science and Engineering Ethics* 26 (1): 387–403. doi:10.1007/s11948-019-00101-7.

Sunderland, M. E. 2014. "Taking Emotion Seriously: Meeting Students Where They Are." *Science and Engineering Ethics* 20 (1): 183–195. doi:10.1007/s11948-012-9427-y.

Ten Holter, Carolyn. 2022. "Participatory Design: Lessons and Directions for Responsible Research and Innovation." *Journal of Responsible Innovation* 9 (2): 275–290. doi:10.1080/23299460.2022.2041801.

Umbrello, S. 2020. "Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design." *Science and Engineering Ethics* 26 (2): 575–595. doi:10.1007/s11948-019-00104-4.

Vallor, Shannon. 2016. *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford: Oxford University Press.

van den Hoven, Jeroen. 2013. "Value Sensitive Design and Responsible Innovation." In *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, edited by R. Owen, J. Bessant and M. Heintz, 75–83. West Sussex: Wiley.

van den Hoven, Jeroen. 2017. "The Design Turn in Applied Ethics." In *Designing in Ethcis*, edited by J. van den Hoven, S. Miller and T. Pogge, 11–31. Cambridge: Cambridge University Press.

Van de Poel, Ibo. 2016. "An Ethical Framework for Evaluating Experimental Technology." *Science and Engineering Ethics* 22 (3): 667–686. doi:10.1007/s11948-015-9724-3.

van de Poel, Ibo. 2020. "Three Philosophical Perspectives on the Relation Between Technology and Society, and How They Affect the Current Debate About Artificial Intelligence." *Human Affairs* 30 (4): 499–451. doi:10.1515/humaff-2020-0042.

Van Grunsven, J. 2022. "Anticipating Sex Robots: A Critique of the Sociotechnical Vanguard Vision of Sex Robots as 'Good Companions'." In *Being and Value in Technology*, edited by Enrico Terrone and Vera Tripodi, 63–91. Cham: Springer International Publishing.

Van Grunsven, J., and S. Roeser. 2022. "AAC Technology, Autism, and the Empathic Turn." *Social Epistemology* 36 (1): 95–110. doi:10.1080/02691728.2021.1897189.

Van Grunsven, J. B., L. Marin, T. W. Stone, S. Roeser, and N. Doorn. 2021. "How to Teach Engineering Ethics? A Retrospective and Prospective Sketch of TU Delft's Approach to Engineering Ethics Education." *Advances in Engineering Education* 9 (4). doi:10.18260/3-1-1153-25254.

Van Grunsven, J., L. Marin, T. Stone, S. Roeser, and N. Doorn. 2023. "How Engineers Can Care from a Distance." In *Thinking Through Science and Technology: Philosophy, Religion, and Politics in an Engineered World*, edited by Glen Miller, Helena Mateus Jeronimo and Qin Zhu, 141–163. Rowman & Littlefield International.

Wisnioski, Matthew, Eric S. Hintz, and Marie Stettler Kleine, eds. 2019. *Does America Need More Innovators?* Cambridge: The MIT Press.

York, Emily, and Shannon N. Conley. 2020. "Creative Anticipatory Ethical Reasoning with Scenario Analysis and Design Fiction." *Science and Engineering Ethics* 26 (6): 2985–3016. doi:10.1007/s11948-020-00253-x.