Imaginative Value Sensitive Design: How Moral Imagination Exceeds Moral Law Theories

in Informing Responsible Innovation

Exam Number: B098426

Word Count: 7936

MSc Epistemology, Ethics and Mind

The University of Edinburgh

2018

Abstract

Safe-by-Design (SBD) frameworks for the development of emerging technologies have become an ever more popular means by which scholars argue that transformative emerging technologies can safely incorporate human values. One such popular SBD methodology is called Value Sensitive Design (VSD). A central tenet of this design methodology is to investigate stakeholder values and design those values into technologies during early stage research and development (R&D). To accomplish this, the VSD framework mandates that designers consult the philosophical and ethical literature to best determine how to weigh moral trade-offs. However, the VSD framework also concedes the universalism of moral values, particularly the values of freedom, autonomy, equality trust and privacy justice. This paper argues that the VSD methodology, particularly applied to nano-bio-info-cogno (NBIC) technologies, has an insufficient grounding for the determination of moral values. As such, an exploration of the value-investigations of VSD are deconstructed to illustrate both its strengths and weaknesses. This paper also provides possible modalities for the strengthening of the VSD methodology, particularly through the application of moral imagination and how moral imagination exceed the boundaries of moral intuitions in the development of novel technologies.

Keywords: value sensitive design, design psychology, moral imagination, applied ethics

1. Introduction

Developed for software engineering and systems development, Value Sensitive Design (VSD) was conceived by its founders Batya Friedman and others in the decade prior to the twentieth century (Friedman 1996). The methodology's intent begins with the premise that technology, and its design, is not value-neutral; building on the ethical and political ladenness of artifacts within the field of science and technology studies (STS) which argues along similar threads (Pinch and Bijker 1987). The social and moral abstracts of design, particularly within that of the human-computer interaction (HCI) community, are what have been of particular interest to Friedman and colleagues. (Friedman, Kahn Jr., and Borinng 2008; Borning and Muller 2012).

Similar communities interested in value-based technological development have, and still are, in development and use such as participatory design, universal design and inclusive design. These different research communities, although focusing on the incorporation of values, differ from the VSD framework given that the former methodologies tend to prize instrumental and functional values such as user-friendliness (van den Hoven and Manders-Huits 2009; Newell et al. 2011). VSD, on the other hand, aims to primarily investigate the stakeholder values of moral, rather than functional or pragmatic importance. These values most centrally include freedom,

autonomy, equality, trust, and privacy justice (Friedman 1997). Although the instrumental values that may be enrolled in the design and development of the technology can be translated as values important to VSD, this framework's primary attention concerns moral values.

Beyond moral epistemology, VSD methodology assumes that upon close analysis of values, some may be determined to be universal among different cultures and societies. Although these values may manifest themselves in varying ways, upon close analysis one can decide that in fact, those manifestations are merely a variation of a universal value. This is the contention that is asserted by Friedman (Friedman 1997). As such, the goal of this paper is to argue that the universal values that VSD instantiates are built on contentious foundations. I contend that, in light of recent developments in cognitive sciences, a revaluation of moral theory and epistemology is necessary if VSD is to remain effective across diverse societies and cultures. To this end, this dissertation provides a thoroughgoing analysis of what the philosopher Mark Johnson calls 'moral imagination' and how moral imagination theory can reinforce the ethical investigations in VSD (Johnson 1993). This paper's core thesis can be formalized as follows:

1. VSD draws from Moral Law Theories to inform responsible innovation (RI).
2. Moral Law Theories fail to capture an accurate understanding of human cognition.
3. An accurate mapping of human cognition is necessary for salient human design.
4. ∴ VSD, as it currently exists, is insufficient for informing RI.
5. Moral Imagination theory accurately maps on to human cognitive architectures.
6. ∴ An adoption of moral imagination theory in VSD will make it more accurate in informing RI.

To the best of my knowledge, this paper is the first to explore the value-investigations that are central to the VSD framework in isolation. Prior literature on VSD has focused on judging the applicability of moral values to technological systems (Brey, 2010; J. Davis & Nathan, 2015), VSD methodology (van den Hoven and Weckert 2008), applications of VSD to current technological innovations (Briggs & Thomas, 2015) as well as to novel technologies (Timmermans et al., 2011; van Wynsberghe, 2013, 2016). Although all these studies provide useful information, they are built upon the existent assumptions regarding the validity of the

moral values enrolled in VSD's value-investigations. This paper's deconstruction of the enrolled moral values is comparatively unique; the intention being to spur continued discussions on some of the most pertinent ethical issues regarding the design of emerging and converging technologies.

To successfully tackle these arguments, this article is organized into the following sections: the first presents an in-depth account of the VSD framework. The goal of doing so is to lay out the methodology envisioned by VSD's founders as well as the framework's evolution to its current state. Particular attention is given to what VSD calls 'conceptual investigations', one of its tripartite constituents which investigates stakeholder values. As this is the part that explicitly focuses on, and is dedicated to, the exploration of moral values, it is of the most significance for this paper's analysis. The second section deconstructs conceptual investigations, mainly looking at the normative ethics on which it takes for granted. In doing so, I evaluate its normative foundations both independently and within the framework itself. The final section of this paper provides a thorough overview of moral imagination theory, why it crucial to conceptualize morality within this framework as well as how it can inform the VSD approach.

2. Value Sensitive Design (VSD)

VSD is a design methodology founded for technology design with the aim to incorporate human values[1] into early design phases in an encompassing and formalized manner (Friedman, Kahn Jr., et al. 2013). Because VSD was birthed from the needs of informational and computational systems design, a design framework was necessary to encompass the value focuses of the field. Specific research emphasis within the field found that the values of privacy, ownership, property, physical welfare, freedom from bias, universal usability, autonomy, informed consent, and trust were each of substantial value to stakeholders to warrant a principled methodology for their incorporation into informational and computation systems (Friedman, Kahn Jr., et al. 2013). As such, the VSD framework was developed in the effort to provide a principled way for designers to account for these values in early design phases (Friedman, Hendry, and Borning 2017). VSD has since been adopted as the potential design methodology for the applications of technologies

---

[1] The recent literature in technology studies have argued for the inseparability and against the bifurcation of the nature/culture dichotomy, and similarly against the privileging of human values over the values of nonhuman animals (see for example Harman, 2018; Morton, 2016).

outside of the ICT and HCI communities (e.g., Umbrello & De Bellis, 2018; van den Hoven, 2013).

According to VSD founders, although rich with different meaning and definitions, values take on something that is beyond the factual distinction of which they may sometimes be equated. Thus, the fact/value separation is of critical importance here. One cannot be conflated with the other, and as such a fact does not logically entail that it is of value. Thus, what is of 'value' in VSD are those "interests and desires of human beings within a cultural milieu" (Friedman et al., 2013, 71), yet the reduction to universalism[2] and absolutism is nonetheless affirmed by VSD theorists and practitioners (e.g., Friedman & Kahn Jr, 2003; van Wynsberghe, 2012). Section 2 provides a more in-depth discussion of what values are of ethical importance and implicated in technology design. The framework, then is situated within the broad range of VSD literature from its founding values of import to more contemporary value investigations.

2.1 A Tripartite Framework

The design of technologies and the artifacts and techniques that emerge from their highly collaborative network of assemblages enroll a host of processes that form an implosive whole, that is it is greater than the sum of its parts (for a more in-depth analysis of this see Harman, 2016; Morton, 2018). VSD, in and of itself, is similar in this way too. VSD is iterative, it builds upon existing design practices and traditions as well as on itself as new information, techniques, issues, and advancements emerge throughout the design process. As such, VSD is composed of three distinct, yet necessarily interrelated parts or stages: conceptual investigations, empirical investigations and technical investigations.

VSD's method sets it apart from competing approaches because it is markedly self-reflexive, falsifiable, and recursively self-improving. Similarly, VSD is anticipatory and does not impute values or responsibility to individuals or groups solely on an ex post facto basis (van den Hoven 2017). VSD is thus predictive, aiming to ascertain potentially implicating values before they

---

[2] Universalism is not equivocal with absolutism. The former argues that moral values are independent of socio-cultural differences, whereas the latter is concerned with intrinsically right/wrong actions.

materialize and in response "influences the design of technology early in and throughout the design process" (Friedman, Kahn Jr., and Borning 2006, 12). I outline only the first part of the methodology given that it is the one that is most philosophically based. The latter two, empirical and technical investigations, although feeding into the first, come after the stakeholders are delimited and values discerned.

2.1.1   Conceptual Investigations

This is considered the most theoretical among the tripartite investigations. Conceptual investigations are philosophically informed and require consultation with the philosophical literature. To this end, it aims to answer questions of limitations and constraints both prior to and throughout the design process such as: Who are the stakeholders? What are the situationally relevant values? Where are the boundaries when discussing often inversely relational values like convenience vs. safety? When do the limits of accepted practice come under question when conflicting with relevant values? What are the opportunity costs of one design over another? These questions and other theoretical areas are explored in conceptual investigations (see Denning, Kohno, and Levy 2013 who do this with smart homes). Research is thus undertaken that set the stage for the remaining two investigations, however, those proceeding investigations reflexively update the conceptual parameters that were initially considered.

3.   What Are Values?

Because it is the goal of this dissertation to articulate a novel framework from which VSD can conceptualize values, it is vital to articulate first what values are and how they are designed into artifacts. To begin, when entering into discussions about values, the lack of agreement between scholars and the existent literature makes any absolutist conception of 'value' - whether ontic or epistemic - highly contentious. To this end, this paper does not aim to argue for any such certainty, instead, this section will opt to discuss how VSD *per se* conceptualizes values.

In the VSD scholarship, a value is considered that attribute with which an individual or group consigns significance (see more specifically Friedman & Kahn Jr, 2003, 2). Throughout their discussion of values, Friedman & Kahn (2003) explore a genealogical account of values, and how, historically, they have been articulated. To this end, the authors' eventually opt for a

deontic account of values that they believe can be distilled from a historical investigation of value conceptualization, particularly those that are of direct relevance to technological design and development (specifically Friedman & Kahn Jr., 2003, 6). The decision to formalize this set of values is due to their apparent trans-societal universality. However, the authors do recognize the incommensurability of these values, their dynamic character depending on what technology is under consideration and the socio-cultural context in which the design is being undertaken. Still, the underlying characteristic of these chosen values, despite their varying relevance, are taken as universal regardless of the socio-cultural hermeneutics that is employed to distill them.

With this in mind, how are values – at least in the VSD sense – designed into artifacts? Technology philosopher Philip Brey provides a salient way of conceptualizing this integration:

> The idea of embedded values is best understood as a claim that technological artifacts (and in particular computer systems and software) have built-in tendencies to promote or demote the realization of particular values. Defined in this way, a built-in value is a special sort of built-in consequence…Then tendencies for the promotion of values are identified as special kinds of built-in consequences of technological artifacts (P. A. E. Brey, 2012, 3).

Because of this, values are instantiated and expressed via their use. To this end, some values that may not have been manifest prima facie may express themselves when implemented by users. Dangers that may result are easy to envision, particularly when cognitive biases are implicated in technological design (Caviola et al. 2014). Although some work has been undertaken to limit the effects of cognitive biases in technological design, the danger that implicit or biased values may be embedded lingers nonetheless and calls for a more holistic conceptualization of values that is in line with developments in the cognitive sciences and their findings on the basic structure of human cognition and moral reasoning (an ad hoc functionalist means of de-biasing VSD has already been undertaken, see Umbrello, 2018b).

For this reason, I argue that moral imagination provides an effective framework for conceptualizing the 'design turn in applied ethics', and in particular VSD. The following section presents a thorough layout of what moral imagination theory is, why it is preferable over the Enlightenment tradition of moral theorizing as well as how VSD can be modified to better account for values in design.

4. Moral Imagination and Imagining Possible Futures

From Plato onwards, the continued variance between the existent theories of morality is, in itself, an indicator of the often conflicting commitments, values and aims that individuals find themselves in their day-to-day lives. Ultimately, distilling any of this 'moral overload' (see van den Hoven et al., 2012) to a single, universal or absolutist principle for grading diverging commitments falls short; something is always sacrificed, albeit even if intuitively, when adopting any moral law tradition (e.g., Stocker 1976a).

This process calls for, as Mark Johnson indicates, "moral imagination…in all its various manifestations, as a means to both knowledge and criticisms" (Johnson, 1993, 187). What this means is a reflexive understanding of the self and the imaginative structure of moral deliberation including its principles and constraints. This also applies to those who reside and ascribe to traditions of morality that differ from our own. Imagining available possibilities that can translate, inform and modify our behaviours and actions and how those imaginaries affect our moral deliberations is what is involved. Moral growth is thus attained via an *imaginative rationality* aimed towards the understanding of cultures, institutions, systems and the self (Johnson 1993).

Moral Law Theories (MLTs) such as deontology, utilitarianism and their various offshoots rarely, if at all, result in moral growth. What they do is instead argue, mistakenly, that there are a constrained set of rules, with definitive actions for making moral decisions. This is typically undertaken in prototypical cases that are themselves constrained and unambiguous. Instead, how they should be viewed is as a distillation of a common set of past collective practices, situated within a particular moral habitat that applies to unproblematic cases. However, these MLTs become ineffable when confronted with the moral overload of day-to-day life and even more so with morally ambiguous technologies, of which no rooted rules or laws can clearly resolve (e.g., Caviola et al., 2014).

Moving beyond resolute theories on how to behave, or that they *should* give us rules is difficult, and particularly contentious within the analytic tradition. To this end, to formulate a strong

argument for why MLTs, as they are envisioned within safe-by-design (SBD) approaches to technological development, this section outlines moral imagination theory. Naturally, given the length constraints of this paper, the account detailed here will be skeletal, and the lack of multiple examples, case studies, and narrative modalities should not be construed as an argument against the strength of moral imagination theory. Readers should make careful note of the full-length arguments existent within the scholarship that adequately buttress moral imagination despite its current lack of popularity and dissemination (Johnson 2014; Fesmire 2003).

4.1 Situating Moral Imagination

The crux of moral imagination theory is founded on the consideration that moral reasoning is fundamentally imaginative and is constituted, given advances in the cognitive sciences, by shared essential imaginative elements. Here I will give a brief account of each of these proposed elements.

4.1.1 Concepts as being Prototypical

The classical understanding of basic concepts as being accountable through a sufficient/necessary framework does not take into consideration that even the most foundational notions of cognition have an architecture that is not exhaustible by this traditional account of concepts. For example, the concept of *love* is prototypical in the sense that there are undoubtedly categorical constituents that are paramount – both cognitively and experientially – but there are also peripheral constituents which, although not central to the concept itself, are more fundamental to critical examples.

To this end, Johnson (1993) argues that MLTs must be abandoned if, in fact, our basic moral concepts (e.g., harm, love, justice, etc.) are prototypical in constitution. This is based on MLTs erroneous account of concepts and their resulted view of moral judgments founded on the concept-understanding landscape. The ways that MLTs seeks to apply absolute judgments[3] to

---

[3] Here 'absolute' refers to 'moral absolutism' in which there is an objective basis on which ethical beliefs can be evaluated.

samples by designating features via a sufficient/necessary approach is itself insufficient; many of these concepts are not ontic in the real world as MLTs characterize them.

This does not mean that MLTs are not functional. Prototypical cases can, usually, be accounted by MLTs without much controversy within a given moral framework. However, the existence of these very moral principles is predicated on their ability to address the prototypical, fundamental constituents of moral concepts, rather than their peripheries.

In contrast, quotidian moral issues surround, instead, the aprototypical situations that individuals find themselves in. However, it is precisely located here, on the edges of those prototypical concepts where the rules prescribed by MLTs fail to function. This is because MLTs cannot, in and of themselves, be extended to encompass these types of cases. Why? Because MLTs are themselves constituted by strict rule-judgments that do not include the capacity for imaginative amplification. Nonetheless, it is this very imaginative ability to address day-to-day moral concerns that is necessary, something MLTs are not equipped, and never were equipped, to provide.

Still, we do possess the capacity for making informed moral judgments. Precepts that allow for application from central prototype cases to peripheral aprototypical cases exist. Metaphor is perhaps one of the most potent tools available to individuals for making this center-periphery leap. Our socio-cultural embodied experience is what permits us to attribute differentiated levels of significance to some instances rather than others. We learn about love, not as a formalized set of features/rules that constitute the central concept, but from prototypical experiences of dating, spatiotemporal closeness (and separation) as well as physical affection; however, they are not exhausted by these experiences (Lakoff and Johnson 2003).

Similarly, the experience of these prototypes is situationally contextual, and this situatedness comprises dynamic and incommensurable affect. This means that various emotional and cognitive states can emerge as a result of this situatedness, all of which are different catalysts for action. To this end, the essential moral concepts that we possess cannot be reduced to disembodied, purely rationalistic features, but are fundamentally co-constituted by our variated

emotional states of embodiment. This means that prototypical models are necessarily dynamic and account for various value-changes. Challenging, often multi-faceted situations that emerge in continually changing societies require greater encompassing power that move far beyond the more primordial prototype from which they began. The co-construction of the prototype-aprototype amplification means that the features of central prototypes will similarly be modified across time.

In sum, because of MLTs' inability to account for non-abstracted contextuality, their potency is contingent on their applicability to the prototype itself, MLTs cannot remain potent independently. Similarly, moral growth requires the extension of prototypes to aprototypical cases via imagination amplification. This results in a cathartic dynamism of prototypes and is precisely what permits moral flourishing.

4.1.2   *Prejudgments*

Like contextualization described in the preceding subsection, the manner in which various cases are conceptually confronted is of critical import. Like metaphors and prototypes which are central to the way individuals approach and conceive circumstances. The way these confronted circumstances are approached and understood is fundamental to the actions taken (or not taken) in response. Consider an imagined case where a terminally-ill individual is given experimental nanorobotics to deliver targeted pharmaceuticals without consent and dies as a result. Conceived as an unethical application of power on behalf of medical professionals, we may respond in one way, however conceptualizing it as an empathetic, last-ditch attempt by doctors to save the already-dying patient's life may result in a very different response.

Language and semiotic structures often implicate imagination, mostly through the enrollment of prototypes and metaphors, among others. What these structures do then is not reflect some indisputable, external concept/case, but translate it. Because the prejudgments that we bring with us into situations where moral evaluation and action is undertaken changes the way we act, a reflexive self-analysis of what prejudgments we are in possession of is necessary if we wish to arrive at an authentic understanding that human moral judgment is contingent upon.

4.1.3 Metaphor

I have consistently brought the concept of metaphor into accounting for what moral imagination theory consists of, and because of its central importance to the theory it warrants further detail.

The basic contention of moral imagination theory (MIT) is that moral deliberations hinge on the use of at least more than one metaphor and not the universal application of moral laws onto external cases (Johnson 2014). This is a profound shift away from traditional conceptions of moral epistemology and the sources of moral knowledge. Novel issues emerge out of this metaphorical extension and conceptualization of moral epistemology that could not have been devised by MLTs because metaphor has no place within those frameworks of morality at all.

MLT, and post-Enlightenment Western philosophy, have typically conceived of metaphor as lacking any rational basis to be considered part of moral epistemology (i.e., the reason/desire bifurcation) (Wood 1999; Borges 2004).[4] The cognitive/emotive dichotomy is typically asserted as a means to avoid the claim that metaphor is implicated in situational reasoning. Similarly, dissent regarding metaphor comes in another form, that being that they lack any objective determinacy to be sufficient for moral knowledge. As such, their content cannot map sufficiently onto cases nor give clear direction on correct action. The precarity that is often associated with extreme moral relativism is cited when conceptions of the metaphoric nature of human cognition, particularly in moral understanding is discussed.

However, the fear of extreme relativism can be averted once there is a genuine understanding of metaphor as a universal process of comprehension, one that is curtailed by common biological, sociocultural, and institutional effects. What must be acknowledged, however, is that a plurality of possible actions can be taken in any situation, none of which are the deduced offspring of univocal universal laws. What MIT is then is a very restricted and limited form of moral relativism (Johnson 1993).

---

[4] See also Critique of Pure Reason where Kant states along these lines that; "Inclination is blind and servile, whether it is kindly or not; and when morality is in question, reason must not play the part of mere guardian to inclination but, disregarding it altogether, must attend solely to its own interest as pure practical reason" (Kant, 2015, 5:118; see also KpV 5:71–72)

Here I think it would be useful to cite one of the specific examples that Johnson uses to explain the metaphoric nature of cognition. He identifies a set of particularly salient moral metaphors that are morally significant. One such metaphor is what he calls the *moral accounting* metaphor that aims to account for the increase/decrease of well-being. He divides this metaphor into five possible instantiations/interpretations of which he calls 'schema'. Due to space limitations, I list one here solely as an illustration:

> *Schema 1:* RECIPROCATION: '*one good turn deserves another'*
> EVENT: A gives *something good*[5] to B.
> JUDGEMENT: B owes *something good* to A.
> EXPECTATION: B should give *something good* to A.
> MORAL INFERENCES: B has an obligation to give *something good* to A. A has a right to receive *something good* from**.**
> MONETARY INFERENCE**:** B pays A for getting *something good* by giving *something good* (of equal price).
> EXAMPLES: I *owe you* a favour for that good deed.
>> You saved my life! How could I ever *repay* you?
>> You've been so kind; I'm deeply *indebted* to you.
>> Your generous acts have *earned* you my respect.
>> I *owe* you more than you'll ever know, for what you've done for me**.**
>> She *bought* his respect with her constant good will. (Johnson, 1993, 47.)

What is happening here is a linguistic illustration of the structure of our moral reasoning, one of the innumerable examples within Western culture of how we speak and reason about morality (Taub 2012; Ervas, Gola, and Rossi 2018). The cognitive sciences have pushed beyond the inherited Enlightenment tradition of faculty psychology regarding the strict bifurcation of reason/desire resulting in a radical re-envisioning of the structure of human cognition. Examples of this are the concepts of *genetically transmitted response strategies* and the *epigenetic rules* of sociobiology (De Tiège et al. 2018; Segerstrale 2016). The once thought natural split between the two selves (in itself a profoundly theological construct) is reformulated as co-constituting one another, intermingled in constant evaluation and self-analysis that makes their a priori separation a fruitless, and philosophically erroneous endeavour (see Kahn, 1987; Tilmouth, 2007).

---

[5] Here 'something good' is defined intra-metaphor as a "valuable commodity (or its monetary equivalent) …" (Johnson, 1993, 47).

In reevaluating the nature of moral knowledge, primarily by acknowledging its metaphorical structure, Johnson lists some of the implications for moral understanding that follow:

- Because our moral understanding is metaphorical, understanding the nuances and specifics of that structure is essential to arriving at a more general understanding of self. In doing so, we can arrive at an understanding of our own prejudgments and similar values (Johnson 2013).

- Perhaps the most important implication of metaphoric evaluation is that investigating metaphor allows us to determine potential contenders for univocal morals. Cross-culture experimental investigations can be undertaken to determine if specific metaphors, like the *moral accounting* metaphor, are essential to human practice. Methods have already been outlined for how this can be undertaken (Gibbs Jr 2017).

- The universal understanding of the metaphoric structure of cognition allows us to ascertain, more specifically, where sociocultural influences come into play (such as the concept of transperspectivity outlined by Putnam, 1981, particularly at 168). Still, although similar metaphoric structures may be universally grounded, their instantiations and definition of values may nonetheless be markedly varied.

- A general understanding of the metaphoric structure of cognition may show that the specific metaphors adopted by individuals, those who define their value structures, serve as the nail in the coffin for any ontological status of absolute moral values. What may constitute 'absolute' in this sense, general or not, may be so divorced from its individual instantiation that its employment as a moral law would be absurd (Johnson 2014). However, the sociocultural instantiation of any particular metaphor may be so potent, despite being domain-specific, that it directs the actions within the group to a determinate degree (Goldberg 2016).

- An emergent issue determines how well established is a metaphor, or network of metaphors (for more on embeddedness see Fauconnier and Turner, 2008a, 2008b). Naturally, the more central, or essential a metaphor is, the more distributed it will be as a node in a network of other metaphors. The essential nature of a metaphor is of concern, especially when discourses of change and critique come into play. The more established a metaphor is in cognition, the lower is its ability to shift. This does not entail the impossibility of change, but such change often comes with devastating sociocultural

expenditures; i.e., European/North American colonialism is a good example of change viz. coercion.

- Aprototypical cases continually present themselves within the dynamic experience of being. Prototypical concepts must then be expanded to confront these. Metaphor is one of the cognitive structures that is employed in making this extension. Many of the moral issues that we confront when exposed to an original case is not the case *per se* but the difficulty in mapping prototypes on to aprototypical entities. Technology, in particular, presents continually dynamic and novel cases that stretch our moral prototypes in ways that enlightenment structures of moral deliberations simply did not, and which could not be accounted for essentially (Ervas, Gola, and Rossi 2018).

- Moral growth is gained by metaphorical extrapolation from similar, yet different, past experiences to new ones. Metaphorical accommodation/assimilation can only be possible because of the plasticity of metaphors.

- One of the critical points about the ontological nature of metaphors is that we can only have limited epistemic access to them. Because they are entities/objects that are consistently networked, interconnected, and enmeshed with other metaphors and cognitive structures, total understanding is impossible and any attempt to exhaust a metaphor through literal or relational strategies fails to capture it sufficiently (for a further exploration of the 'withdrawn' nature of objects/entities see Harman 2018a). One of the primary failures of the analytic moral tradition is its ultimately failed attempt to exhaust moral identity into a rule-based system that sidelines the very foundations that give this identity value.

In sum, moral epistemology hinges on an intimate comprehension of how metaphors structure and are continually structured by human moral thought and action. In doing so, the varied – although potentially universally based – instantiations of moral metaphors force us to accept that cases of moral overload have a plurality of potential ways of being addressed, rather than a single, universal rule. This does not mean that radical subjectivism/relativism are affirmed, as a consequence (that would be a false dilemma). Biopsychosocial restrictions strongly delimit the contentiousness of subjectivism by limiting the foundations of specific metaphors within cultural bounds (e.g., like the psychobiological response patterns, see Kemeny, 2009). The latter is

avoided because of the cognitive sciences which have shown the universality of some metaphors as being the consequence of embodied being (Varela, Thompson, and Rosch 2017; Gibbs Jr 2006).

4.1.4   Narrative

The function of the narrative structure of moral understanding and growth is often overlooked within the analytic tradition of moral theorizing. It has been noted that individuals seeking moral growth turn towards narratives (in the physical or performed modes) over the formulaic moral texts of philosophers (Rorty 1989; Peterson 1999; Mordini 2007).

The ultimate motivations for why narratives have such moral importance, and thus impact on moral growth lie in the narrative structure of day-to-day human-*being*. Our embodied, lived experiences within a structured narrative permit agents to make discrete discernments between what is of moral import or not in any given situation (Krakowiak and Tsay-Vogel 2015; Kramp 2003). What we do when we engage with performed narrative (books, plays, etc.) is that we engage in a vicarious projection of ourselves and embody the character's thoughts, actions and moral anxiety (i.e., Sarbin 2001). In doing so, we engage in a reflexive dialogue and critique of their chosen paths and their actions which, as a function of our very engagement in this behavior, impacts and develops our own moral understanding (Haidt and Joseph 2007).

Although the empirical literature on the narrative structure of moral growth (and similar cognitive modules) has been the product of the cognitive sciences, there has been philosophical work grounding the value of a narrative understanding of morality in philosophy as well. The vanguard of this tact has been by Nussbaum in her critical exploration of classical Greek tragedy, *Fragility and Goodness,* where she perceptibly states:

> a whole tragic drama, unlike a schematic philosophical example making use of a similar story, is capable of tracing the history of a complex pattern of deliberation, showing its roots in a way of life and looking forward to its consequences in that life. As it does all of this, it lays open to view the complexity, the indeterminacy, the sheer difficulty of actual human deliberation…*A tragedy does not display the dilemmas of its characters as pre-articulated; it shows them searching for the morally salient; and it forces us, as interpreters, to be similarly active* (Nussbaum, 2001, 14, *emphasis mine*).

Narrative, like life then, is not so determinate as the arguments moral philosophers are so apt to construe. Like Kant's impenetrable *das Ding-an-sich*, no amount of re-contextualizing and forceful determinacy can exhaust the nuances and contextual contingencies of any given situation that an individual can, and will, inevitably find themselves. What narrative does then, at the very least, is bring a genuine colour to the portrait of which the rules of moral philosophers have only captured a tiny, but still important outline.

Narrative, to this end, allows for a more holistic and situated evaluation of our diverse moral landscapes *in time*. Our various, and sometimes conflicting moral enmeshments can be engaged within a continually developing, rather than ad hoc, way (van den Hoven, Lokhorst, and van de Poel 2012). Narratives are a mode towards enrichment that cannot be captured in any meaningful sense, nor exhausted by moral laws (for more on exhuastion see Harman, 2016, 2018b).

Even when we consider moral laws in themselves, their force – if that force becomes manifest – only emerge through a collective and experiential understanding of how these laws are embodied and function over time. Their salience is given force through narrative, not viz. some independent objectivity (D. J. Haraway 2016).

In sum, the narrative structure of human experience and cognition must be given its due role in moral philosophy. This becomes particularly apparent as the continued development of transformative[6] technologies shifts sociocultural, political and economic norms. Any sufficiently potent anticipatory design approach adopted for the development of these technologies must embody an imaginative architecture that more authentically maps our moral cognitive structures to our situated realities.

4.2 *Understanding* the Moral Imagination

Too move towards the application of moral imagination into the practical domain of VSD, here I will briefly summarize this section's arguments regarding MIT.

---

[6] Transformative being the operative term here, it's particularly within a narrative and linguistic evaluation where substantial change occurs.

Firstly, the central thesis underlying MIT is that rather than the overly literal and law-like conceptions of morality that are envisioned by MLTs, MIT instead argues that *understanding* is what underpins morality, and any sufficiently apt moral theory must be itself understood as an experiential and progressive venture. Similarly, this understanding must reflexively be recognised as being imaginatively constituted by structures that are narrative, prototypical, semantically constrained by prejudgments and metaphorical (Johnson 2014). Moral growth is attained gradually through the realization of these imaginative structures, their limits and constraints, the possible open futures of discrete potentials and the consequences that these varied implications can have.

This re-envisioning of morality as an act of understanding requires the disposal of any conception of objective moral absolutes as well as extreme relativism and subjectivism. The conception of the bifurcation and faculty psychology of the reason/desire dichotomy must be similarly discarded as illusory and the product of Enlightenment heritage thinking *sans* modern advances in the cognitive sciences which show otherwise (Gaesser, Keeler, and Young 2018; M. H. Davis 2018).

Another condition follows this re-envisioning: the conception of moral understanding shifts the telos of moral inquiry away from the application of strict moral principles to cases and instead towards the goal of cultivating our abilities to discriminate the relevant moral dimensions of scenes and their actors as well as maturing our abilities to empathically extend our imagination toward envisioning ourselves in the experience of others (Koehn 2012; Oxley 2011). This is one of the fundamental constituents of the human moral makeup that has generally been sidelined in MLTs and is required for reflexive moral evaluation to take place (Stocker 1976b; Llewellyn et al. 2016).

Finally, to better avoid confusion and to address possible rebuttals to the MIT, I briefly address the assumptions that may be levied against imagination as a concept itself. The analytic core at the centre of MLTs typically regard the distilled laws as the product of human rationality and to this end, consider imagination as antithetical to this conception. Imagination is thus relegated to an aesthetic realm within the Enlightenment understanding of cognition (i.e., faculty psychology)

(Johnson 2014). To this end, the aesthetic is considered independent and displaced from rationality and the reason-derived principles of moral theorizing. This tradition is based on the concept of determinacy towards application in specific instances, of which reason is understood as being the acme of determinacy whereas imagination is indeterminate (Rundell 2016; Kant 2015). This is a false assumption.

The rigid conceptions that were thought to demarcate reason from imagination and allow one faculty (reason) to distill rules of morality are based on unfounded assumptions of how human cognitive architecture is structured. The continued assumption of this reason/imagination bifurcation is the only conceptual route one can assume to exclude imagination from moral philosophy, however, the cognitive sciences have shown this dualism to be unfounded (Narvaez 2016). Nevertheless, imagination, given its essentiality to human cognitive architecture, morality only being a single – yet interconnected – facet of this network consists of the distribution of the aesthetic in experience. Thus, explanatory modes that are linguistically characterized as 'scientific' or as 'moral' can be construed as diverse instantiations of the aesthetic that differ in their priority of specific features or qualities. John Dewy best accounts this expanded conceptualization of the aesthetic structure of experience (Dewey, 1934, particularly at 46,55).

Similarly, the MIT abates any criticism of extreme subjectivism and moral relativism by discussing morality within its necessarily embodied state, something that is collectively shared by the species; rather than the disembodied and 'gods-eye-view' perspective (D. Haraway 1988). Studies have shown the shared, yet diverse instantiations of universal archetypes across cultural divides (Varela, Thompson, and Rosch 2017; Goldsmith and Huck 2013). This enables a discourse between what are often divergent and conflicting moral traditions (MacIntyre, 1988 especially at 144 best accounts the dialectic of moral growth between moral traditions).

This section has given a cursory account of what MIT consists of and how MLTs that have dominated the western tradition of morality are, in themselves, insufficient for moral progress; particularly in aprototypical cases that individuals are confronted with in their quotidian lives. Because these day-to-day moral decisions are influenced in strange and often counter-intuitive ways by new and emerging technologies, the salience of design hinges on accounting for values

in a way that aligns with the actual structure of our moral cognition. To this end, the following section will assemble the VSD methodology by reformulating its conceptual investigations to be imaginatively structured rather than 'principled' as its founders envisioned (Friedman, Kahn, et al. 2013).

5. Assembling a *Dynamic* Value-Sensitive Design

This project thus far has demonstrated two things: 1) it has given an outline of the motivations for VSD's (and other SBD approaches) emergence as well as a brief account of its structure and methodology; and 2) I have given a cursory argument of and for why MIT theory maps on to human reality in ways that MLTs are simply not equipped to do. The primary motivation behind this is that VSD and similar methodologies aim to address the value-ladenness of technology design through value investigations that essentially draw from the existent philosophical literature. The dominant analytic modes of moral inquiry have been the traditional sources of philosophical interest, but the 'ecological turn' in science and technology studies (STS) has resulted in a greater tapping of the continental tradition, mostly at the intersection of cultural anthropology (Morton 2018; Latour 2017).

Still, much of the VSD literature has, and is currently, drawing from the inherited Enlightenment tradition based on faculty psychology, primarily the deeply anthropocentric post-Kantian tradition of correlationism[7]. As I have argued, this is untenable, mostly because the moral laws distilled from these traditions are not reflective of the constitution of human cognition, primarily its failure to account for the imaginative structures such as metonymy, metaphor and narrative that have been shown as essential to moral deliberation and development. The issues are exacerbated with the introduction and co-construction of transformative technologies, primarily NBIC artifacts that push the limits of our moral intuitions (Umbrello, 2018b) and stretch any embedded deontic ethics to inoperability beyond any strictly delimited and prototypical cases.

---

[7] Correlationism was first coined by Quentin Meillassoux (2008) to describe the post-Kantian anthropocentric conceit that "we only ever have access to the correlation between thinking and being, and never to either term considered apart from the other" (Meillassoux, 2009, 5). This idea has been a common ground for criticism that has united the speculative realist philosophers.

Autonomous vehicles (AVs) prove to be a timely example, and a predictive precursor to more advanced autonomous technologies, particularly in the socio-ethical issues that emerge with their introduction (Etzioni and Etzioni 2016). Primary concerns with the ubiquitous rollout of AVs are when they are confronted with trolley-like 'dilemmas' of life and death as well as what designers and programmers can do to address these issues (Gogoll and Müller 2017). Primary recourse has been to ad-hoc implementations of hardware that can transfer the action-potential and consequential responsibility onto the driver (Contissa, Lagioia, and Sartor 2017). A similar strategy has been taken to offset the inevitable inclusion of cognitive biases when enrolling stakeholders to extract values for design, mostly by applying de-biasing heuristic tools (Umbrello 2018b). Still, these approaches to design are severely restricted, mostly because it requires the programming of prototypical cases and strict responses. The embedding of MLT approaches to decision-making inherently leaves some degree of moral overload, meaning that the option that the programmers embed do not exhaust the moral tension that the proposed 'dilemma' appear to evoke. This is not to say that the MLT approach does not have some functional value, the reality remains that these technologies are being designed, rolled out, and integrated within the sociocultural milieu in which they are introduced. Ethical issues arise and must be addressed, this is an excellent preliminary avenue that can be taken, but further conceptual, technical, and empirical research must be undertaken for future technology assessment.

5.1 Shifting Focus Towards Designers

One of the most substantial progressions in applied technology ethics is one that emerged as the backbone of the *design turn*, primarily the shift of focus away from the artifacts themselves and towards the activities and values of stakeholders (both direct stakeholder such as designers and indirect stakeholders such as publics). Still, substantial emphasis in technology assessment and speculative techno-ethics discourses is placed on the moral agency of technologies per se. This is seen more often than naught in the literature on artificial intelligence, primarily artificial general intelligence (AGI) and artificial superintelligence (ASI). Although there is considerable value in speculative ethics (see especially Roache, 2008), care must be taken not to commensurate current forms of AI and those of a futuristic, although arguably inevitable forms of AI (Müller and Bostrom 2016). The operative point here is that the scholarship has argued that the latter

(AGI/ASI) constitute moral agents and similarly initiates discussions on personhood and rights attribution; the former, however, do not (Rolf and Crook 2016; Hughes 2014).

To this end, designers and ethicists have, and still are, considering which 'human values' to incorporate in design, how we align technologies – AI for example – towards pathways that are beneficial for human civilization (Umbrello and Baum 2018; Baum 2016). Aside from the blatant anthropocentrism and dominant Western moral ideology that are usually uncritically assumed in these discussions[8], particularly when totalizing conceptions of 'the world' and 'human civilization' are invoked, the modern deontic theories of morality (utilitarianism and deontology) are usually assumed as the philosophical solution, shifting the issues and attention towards technical means to operationalize them.

What is needed then is a fundamental re-evaluation of the strategies adopted by designers and ethicists in technology assessment and SBD approaches that shift away from the static moral conceptions that are presently being considered and employed and instead towards an *imaginative* technology assessment and design methodology that begins with conceptual investigations built on the premises of an imaginative constitution.

The following section offers a preliminary example of how this reassembly of VSD, in particular, can be undertaken. Naturally, the limits of a single paper of this type restrict this praxis to only a cursory and speculative level. However, it should be the work of future research projects to explore how, if at all, the imaginative structure of the human mind, with its use of metonymy, metaphor, and narrative can be implemented in design strategies without subverting itself to strict principles.

5.2 A Praxis for Imaginative Value Sensitive Design (IVSD)

Because the emphasis of this paper is on the moral approach that VSD adopts, it leaves out a full discussion of VSD's restructuring in favour of emphasizing conceptual investigations alone.

---

[8] I have elsewhere criticised this Western moral dominance towards a preliminary course of adopting an approach of intersubjectivity (Umbrello 2018b).

At its core, VSD's conceptual investigations ideally begin with researchers exploring and determining various 'value scenarios' that can be envisioned and the potential tensions and insufficiencies that may arise (Friedman, Hendry, and Borning 2017). This is done by 1) determining the stakeholder population group and 2) beginning with a core value (e.g., safety, autonomy, privacy, etc.). This also involves determining the project values (sources of funding, industry interests, etc.) (Borning and Muller 2012).

An excellent way to begin rethinking the potential and nature of VSD is to:
1. Abandon the notion of a core set of root values. Instead of thinking of values as culturally instantiated, and objectively distilled, instead think of basic cognitive modules as universally shared.
2. Similarly, the varied conception of 'value' as such must be explored as socioculturally situated, metaphorically conceptualized, narratively embodied, and developed through metonymy. VSD then must humble itself away from embedding objective values in design and towards a conception of design as embodying a dynamic value development. This is directly in line with VSD's explicit goals of flexibility and reflexivity.

A means of approaching a salient IVSD could then be understood as injecting an imaginative tool-analysis alongside the basic VSD approach. Here I lay out the basic steps of conceptual investigations paired with a possible imaginative tool (iTool) that can be used to begin to rethink how VSD is practiced:

Step 1: Determine project and designer values
iTool: Identify prototypical structures implicated in the design project

Step 2: Identify Direct and Indirect Stakeholders
iTool: Identify the plurality of moral traditions that these populations may be part of.

Step 3: Determine Benefits and Harms for Stakeholders

iTool: Determine a definition of benefits/harms *a posteriori*, that, at least functionally, is acceptable to all the enrolled stakeholder communities.

Step 4: Identify Potential Values
iTool: Identify prototypical structures

Step 5: Distill Working Definition of Identified Values
iTool: Expand how sociocultural aprototypical values are instantiated from prototypes.

Step 6: Identify Value Conflict
iTool: Identify if the conflicts emerge from opposing values within a single moral tradition or because of cross-tradition incomparability?

Step 7: Use Heuristic Tools in Stakeholder Value-Elicitation
iTool: Identify if heuristic tools are trans-socially applicable. Ask if they aim at prototype structures or aprototypical expansions.

6. Conclusion

What this paper has attempted to do is to provide an argument for why current design-for-values methodologies, particularly VSD are insufficient for informing responsible innovation. The primary motivation for this is they rely exclusively on the MLTs of the Anglo-American tradition of philosophy to design technologies across cultures. The crux of the argument proposed has been to show how these MLTs do not accurately map on to innate human cognitive architecture that employs metaphor, narrative, metonymy, and imagination when engaging in moral deliberation. To this end, this paper has offered a way to incorporate an imaginative rationality in the VSD framework that may authentically map on to human cognition. Future research projects should empirically investigate which prototypical structures exists across socio-cultural divides and how those structures can be incorporated to form the core of VSD's conceptual investigations. In lieu of these empirical finding, the proposition of this paper remains conceptual and extrapolative. Yet, the observations from the cognitive sciences makes a strong case for re-envisioning moral deliberations as an imaginative process nonetheless.

References

Aad Correljé, Eefje Cuppen, Marloes Dignum, Udo Pesch & Behnam Taebi. 2015. "Responsible Innovation in Energy Projects: Values in the Design of Technologies, Institutions and Stakeholder Interactions 1 (Draft Version for Forthcoming Book) Aad Correljé, Eefje Cuppen, Marloes Dignum, Udo Pesch & Behnam Taebi." In *Responsible Innovation 2*, edited by Bert-Jaap Koops, Ilse Oosterlaken, Henny Romijn, Tsjalling Swierstra, and Jeroen van den Hoven, 183–200. Springer International Publishing. https://link.springer.com/chapter/10.1007%2F978-3-319-17308-5_10.

Baum, Seth D. 2016. "On the Promotion of Safe and Socially Beneficial Artificial Intelligence." *AI and Society*, no. July: 1–9. doi:10.1007/s00146-016-0677-0.

Borges, Maria. 2004. "What Can Kant Teach Us about Emotions?" *The Journal of Philosophy* 101 (3). Journal of Philosophy, Inc.: 140–58.

Borning, Alan, and Michael Muller. 2012. "Next Steps for Value Sensitive Design." *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12*, 1125. doi:10.1145/2207676.2208560.

Brey, Philip A. E. 2010. "Philosophy of Technology after the Empirical Turn." *Techné: Research in Philosophy and Technology* 14 (1): 36–48. doi:10.5840/techne20101416.

———. 2012. "Values in Technology and Disclosive Computer Ethics." In *The Cambridge Handbook of Information and Computer Ethics*, edited by Luciano Floridi, 41–58. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511845239.004.

Briggs, Pam, and Lisa Thomas. 2015. "An Inclusive, Value Sensitive Design Perspective on Future Identity Technologies." *ACM Transactions on Computer-Human Interaction* 22 (5): 1–28. doi:10.1145/2778972.

Caviola, Lucius, Adriano Mannino, Julian Savulescu, and Nadira Faulmuller. 2014. "Cognitive Biases Can Affect Moral Intuitions about Cognitive Enhancement." *Frontiers in Systems Neuroscience* 8 (October): 1–5. doi:10.3389/fnsys.2014.00195.

Contissa, Giuseppe, Francesca Lagioia, and Giovanni Sartor. 2017. "The Ethical Knob: Ethically-Customisable Automated Vehicles and the Law." *Artificial Intelligence and Law* 25 (3): 365–78. doi:10.1007/s10506-017-9211-z.

Davis, Janet, and Lisa P. Nathan. 2015. "Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains." In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, edited by Jeroen van den Hoven, Pieter E. Vermaas, and Ibo van de Poel, 12–40. doi:10.1007/978-94-007-6970-0.

Davis, Mark H. 2018. *Empathy: A Social Psychological Approach*. New York, NY: Routledge.

Dechesne, Francien, Martijn Warnier, and Jeroen van den Hoven. 2013. "Ethical Requirements for Reconfigurable Sensor Technology: A Challenge for Value Sensitive Design." *Ethics and Information Technology* 15 (3): 173–81. doi:10.1007/s10676-013-9326-1.

Denning, Tamara, Tadayoshi Kohno, and Henry M Levy. 2013. "A Framework for Evaluating Security Risks Associated with Technologies Used at Home." *Communications of the ACM* 56 (1). doi:10.1145/2398356.2398377.

Dewey, John. 1934. *Art as Experience*. New York, NY: G.P. Putnam's Sons.

Ervas, Francesca, Elisabetta Gola, and Maria Grazia Rossi. 2018. "Argumentation as a Bridge Between Metaphor and Reasoning." In *Argumentation and Language—Linguistic,*

Cognitive and Discursive Explorations, 153–70. Springer.

Etzioni, Amitai, and Oren Etzioni. 2016. "AI Assisted Ethics." *Ethics and Information Technology* 18 (2). Springer Netherlands: 149–56. doi:10.1007/s10676-016-9400-6.

Fauconnier, Gilles, and Mark Turner. 2008a. "Rethinking Metaphor." In *CAMBRIDGE HANDBOOK OF METAPHOR AND THOUGHT*, edited by Raymond W Gibbs. Cambridge University Press. doi:10.2139/ssrn.1290862.

Fauconnier, Gilles, and Mark B Turner. 2008b. "Conceptual Projection and Middle Spaces." *SSRN Electronic Journal*. doi:10.2139/ssrn.1290862.

Fesmire, Steven. 2003. *John Dewey and Moral Imagination: Pragmatism in Ethics*. Bloomington, IN: Indiana University Press. http://www.iupress.indiana.edu/product_info.php?products_id=21877.

Flanagan, Mary, Daniel C. Howe, and Helen Nissenbaum. 2008. "Embodying Values in Technology: Theory and Practice." In *Information Technology and Moral Philosophy*, edited by Jeroen van den Hoven and John Weckert, 322–53. New York, NY: Cambridge University Press. http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521855495&ss=cop.

Friedman, Batya. 1996. "Value-Sensitive Design." *Interactions* 3 (6): 16–23. doi:10.1145/242485.242493.

———. 1997. *Human Values and the Design of Computer Technology*. Edited by Batya Friedman. CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/site/1575860805.shtml#.

Friedman, Batya, David G. Hendry, and Alan Borning. 2017. "A Survey of Value Sensitive Design Methods." *Foundations and Trends® in Human–Computer Interaction* 11 (2): 63–125. doi:10.1561/1100000015.

Friedman, Batya, Peter H. Kahn Jr., and A. Borinng. 2008. "Value Sensitive Design and Information Systems." In *The Handbook of Information and Computer Ethics*, edited by K. E. Himma and H. T. Tavani. Hoboken, NJ, USA: John Wiley & Sons, Inc. 10.1002/9780470281819.ch4.

Friedman, Batya, Peter H. Kahn Jr., and Alan Borning. 2006. "Value Sensitive Design and Information Systems (PREPRINT)." *Human-Computer Interaction and Management Information Systems: Foundations*, 1–27. doi:10.1145/242485.242493.

Friedman, Batya, Peter H. Kahn Jr., Alan Borning, and Alina Huldtgren. 2013. "Value Sensitive Design and Information Systems." In *Early Engagement and New Technologies: Opening up the Laboratory*, edited by Neelke Doorn, Daan Schuurbiers, Ibo van de Poel, and Michael E Gorman, 55–95. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-7844-3_4.

Friedman, Batya, and Peter H Kahn Jr. 2003. "Human Values, Ethics, and Design." In *The Human-Computer Interaction Handbook*, edited by Julie A Jacko and Andrew Sears, 1177–1201. Hillsdale, NJ, USA: L. Erlbaum Associates Inc. http://dl.acm.org/citation.cfm?id=772072.772147.

Friedman, Batya, and Peter H Kahn Jr. 2003. "Human Values, Ethics, and Design." *The Human-Computer Interaction Handbook*. Lawrence Erlbaum Associates Mahwah, NJ, 1177–1201.

Friedman, Batya, Peter H Kahn, Alan Borning, and Alina Huldtgren. 2013. "Value Sensitive Design and Information Systems." In *Early Engagement and New Technologies: Opening up the Laboratory*, edited by Neelke Doorn, Daan Schuurbiers, Ibo van de Poel, and Michael E Gorman, 55–95. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-

7844-3_4.

Gaesser, Brendan, Kerri Keeler, and Liane Young. 2018. "Moral Imagination: Facilitating Prosocial Decision-Making through Scene Imagery and Theory of Mind." *Cognition* 171 (November 2017). Elsevier: 180–93. doi:10.1016/j.cognition.2017.11.004.

Gibbs Jr, Raymond W. 2017. *Metaphor Wars*. Cambridge, UK: Cambridge University Press.

Gibbs Jr, Raymond W. 2006. *Embodiment and Cognitive Science*. Cambridge, UK: Cambridge University Press.

Gogoll, Jan, and Julian F Müller. 2017. "Autonomous Cars: In Favor of a Mandatory Ethics Setting." *Science and Engineering Ethics* 23 (3). Springer: 681–700.

Goldberg, Bruce. 2016. "Mechanism and Meaning." In *Investigating Psychology*, 62–80. Routledge.

Goldsmith, John A, and Geoffrey J Huck. 2013. *Ideology and Linguistic Theory: Noam Chomsky and the Deep Structure Debates*. Routledge.

Haidt, Jonathan, and Craig Joseph. 2007. "The Moral Mind: How Five Sets of Innate Intuitions Guide the Development of Many Culture-Specific Virtues, and Perhaps Even Modules." *The Innate Mind* 3. Citeseer: 367–91.

Haraway, Donna. 1988. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14 (3): 575. doi:10.2307/3178066.

Haraway, Donna J. 2016. *Staying with the Trouble: Making Kin in the Chthulucene*. Duke University Press. https://www.dukeupress.edu/staying-with-the-trouble.

Harman, Graham. 2016. *Immaterialism: Objects and Social Theory*. Cambridge, UK: Polity. http://ca.wiley.com/WileyCDA/WileyTitle/productCd-1509500960.html.

———. 2018a. *Object-Oriented Ontology: A New Theory of Everything*. New York, NY: Penguin Random House.

———. 2018b. *Object-Oriented Ontology*. Pelican.

Hoven, Jeroen van den. 2013. "Architecture and Value-Sensitive Design." In *Ethics, Design and Planning of the Built Environment*, edited by Claudia Basta and Stefano Moroni, 224. Springer Science & Business Media. https://books.google.ca/books?id=VVM_AAAAQBAJ&dq=moral+value+such+as+freedom,+equality,+trust,+autonomy+or+privacy+justice+%5Bthat%5D+is+facilitated+or+constrained+by+technology&source=gbs_navlinks_s.

———. 2017. "The Design Turn in Applied Ethics." In *Designing in Ethics*, edited by Jeroen van den Hoven, Seumas Miller, and Thomas Pogge, 11–31. Cambridge, UK: Cambridge University Press. doi:10.1017/9780511844317.

Hoven, Jeroen van den, Gert Jan Lokhorst, and Ibo van de Poel. 2012. "Engineering and the Problem of Moral Overload." *Science and Engineering Ethics* 18 (1): 143–55. doi:10.1007/s11948-011-9277-z.

Hoven, Jeroen van den, and Noëmi Manders-Huits. 2009. "Value-Sensitive Design." In *In A Companion to the Philosophy of Technology*, edited by Jan Kyrre Berg Olsen, Stig Andur Pedersen, and Vincent F. Hendricks. Oxford, UK: Wiley-Blackwell. doi:10.1002/9781444310795.ch86.

Hoven, Jeroen van den, and John Weckert. 2008. *Information Technology and Moral Philosophy*. Edited by Jeroen van den Hoven and John Weckert. Cambridge University Press. http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521855495.

Hughes, James. 2014. "Compassionate AI and Selfless Robots: A Buddhist Approach." In *Robot Ethics: The Ethical and Social Implications of Robotics*, edited by Patrick Lin, Keith

Abney, and George A. Bekey, 69–84. MIT Press.

Huldtgren, Alina. 2014. "Design for Values in ICT." In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, edited by Jeroen van den Hoven, Pieter E Vermaas, and Ibo van de Poel, 1–24. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-6994-6_35-1.

Johnson, Mark. 1993. *Moral Imagination: Implications of Cognitive Science for Ethics*. Chicago, IL: University of Chicago Press.

———. 2013. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. University of Chicago Press.

———. 2014. *Morality for Humans: Ethical Understanding from the Perspective of Cognitive Science*. Chicago, IL: The University of Chicago Press. http://press.uchicago.edu/ucp/books/book/chicago/M/bo17322899.html.

Kahn, Charles H. 1987. "Plato's Theory of Desire." *The Review of Metaphysics*. JSTOR, 77–103.

Kant, Immanuel. 2015. *The Critique of Practical Reason*. Edited by Andrews Reath and Mary Gregor. Cambridge, UK: Cambridge University Press.

Kemeny, Margaret E. 2009. "Psychobiological Responses to Social Threat: Evolution of a Psychological Model in Psychoneuroimmunology." Elsevier.

Koehn, Daryl. 2012. *Rethinking Feminist Ethics: Care, Trust and Empathy*. London, England: Routledge.

Krakowiak, K Maja, and Mina Tsay-Vogel. 2015. "The Dual Role of Morally Ambiguous Characters: Examining the Effect of Morality Salience on Narrative Responses." *Human Communication Research* 41 (3). Oxford University Press Oxford, UK: 390–411.

Kramp, Mary Kay. 2003. "Exploring Life and Experience through Narrative Inquiry." In *Foundations for Research*, 119–38. Routledge.

Lakoff, George, and Mark Johnson. 2003. *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

Latour, Bruno. 2017. *Facing Gaia: Eight Lectures on the New Climatic Regime*. Polity.

Llewellyn, Rebecca, Chrystal Jaye, Richard Egan, Wayne Cunningham, Jessica Young, and Peter Radue. 2016. "Employing Imaginative Rationality: Using Metaphor When Discussing Death." *Medical Humanities*. Institute of Medical Ethics, medhum-2016.

MacIntyre, Alasdair C. 1988. *Whose Justice? Which Rationality?* London, England: Duckworth London.

Meillassoux, Quentin. 2009. *After Finitude: An Essay on the Necessity of Contingency*. Edited by Ray Brassier. Bloomsbury Publishing. https://www.bloomsbury.com/uk/after-finitude-9781441173836.

Mordini, Emilio. 2007. "The Narrative Dimension of Nanotechnology." *Nanotechnology Perceptions* 3 (March): 15–24. http://www.cssc.eu/public/The narrative dimension of nanotechnology.pdf.

Morton, Timothy. 2016. *Dark Ecology: For a Logic of Future Coexistence*. New York, NY: Columbia University Press.

———. 2018. *Being Ecological*. Boston, MA: MIT Press. https://mitpress.mit.edu/books/being-ecological.

Müller, Vincent C, and Nick Bostrom. 2016. "Future Progress in Artificial Intelligence: A Survey of Expert Opinion BT - Fundamental Issues of Artificial Intelligence." In , edited by Vincent C Müller, 555–72. Cham: Springer International Publishing. doi:10.1007/978-3-

319-26485-1_33.

Narvaez, Darcia. 2016. *Embodied Morality: Protectionism, Engagement and Imagination*. Springer. https://books.google.ca/books?id=1WBBDAAAQBAJ&dq=embodied+morality&lr=&source=gbs_navlinks_s.

Newell, Alan F., P. Gregor, M. Morgan, G. Pullin, and C. Macaulay. 2011. "User-Sensitive Inclusive Design." *Universal Access in the Information Society* 10 (3): 235–43. doi:10.1007/s10209-010-0203-y.

Nussbaum, Martha C. 2001. *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. 2nd ed. Cambridge, UK: Cambridge University Press. http://www.cambridge.org/gb/academic/subjects/philosophy/ethics/fragility-goodness-luck-and-ethics-greek-tragedy-and-philosophy-2nd-edition#ZSWphmvlMH11Ollq.99.

Oxley, Julinna. 2011. *The Moral Dimensions of Empathy: Limits and Applications in Ethical Theory and Practice*. London, England: Springer.

Peterson, Jordan B. 1999. *Maps of Meaning: The Architecture of Belief*. New York, NY: Routledge. https://www.routledge.com/Maps-of-Meaning-The-Architecture-of-Belief/Peterson/p/book/9780415922227.

Pinch, Trevor, and Wiebe E. Bijker. 1987. "The Social Construction of Facts and Artifacts." In *The Social Construction of Technological Systems : New Directions in the Sociology and History of Technology*, edited by Wiebe E. Bijker, Thomas Parke. Hughes, and Trevor Pinch, 405. MIT Press. https://books.google.ca/books?id=B_Tas3u48f8C&printsec=frontcover&dq=The+Social+Construction+of+Technological+Systems&hl=en&sa=X&ved=0ahUKEwjvtbqrkfHXAhUMMd8KHdXnDGEQ6AEIKDAA#v=onepage&q=The Social Construction of Technological Systems&f=false.

Putnam, Hilary. 1981. *Reason, Truth and History*. Vol. 3. Cambridge University Press.

Roache, Rebecca. 2008. "Ethics, Speculation, and Values." *NanoEthics* 2 (3): 317–27. doi:10.1007/s11569-008-0050-y.

Rolf, Matthias, and Nigel Crook. 2016. "What If: Robots Create Novel Goals? Ethics Based on Social Value Systems." In *CEUR Workshop Proceedings*, 20–25.

Rorty, Richard. 1989. *Contingency, Irony, and Solidarity*. Cambridge, UK: Cambridge University Press. http://www.cambridge.org/gb/academic/subjects/philosophy/philosophy-social-science/contingency-irony-and-solidarity?format=PB&isbn=9780521367813#JtPRZ0kE9KUjmqvG.97.

Rundell, John. 2016. "Creativity and Judgement: Kant on Reason and Imagination." In *Rethinking Imagination*, 87–117. Routledge.

Sarbin, Theodore R. 2001. "Embodiment and the Narrative Structure of Emotional Life." *Narrative Inquiry* 11 (1). John Benjamins Publishing Company: 217–25.

Segerstrale, Ullica. 2016. "The Westermarck Thesis as a Thinking Tool for Sociobiology." In *Evolution, Human Behaviour and Morality*, 97–118. Routledge.

Stocker, Michael. 1976a. "The Schizophrenia of Modern Ethical Theories." *Journal of Philosophy* 73 (14): 453–66. doi:10.2307/2025782.

———. 1976b. "The Schizophrenia of Modern Ethical Theories." *Journal of Philosophy* 73 (14): 453–66. doi:10.2307/2025782.

Taub, Sarah. 2012. "Iconicity and Metaphor." *Sign Language–An International Handbook. Berlin: Walter de Gruyter*, 388–412.

Tiège, Alexis De, Yves Van de Peer, Johan Braeckman, and Koen B Tanghe. 2018. "The Sociobiology of Genes: The Gene's Eye View as a Unifying Behavioural-Ecological Framework for Biological Evolution." *History and Philosophy of the Life Sciences* 40 (1). Springer: 6.

Tilmouth, Christopher. 2007. *Passion's Triumph over Reason: A History of the Moral Imagination from Spenser to Rochester*. Oxford, UK: Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199212378.001.0001/acprof-9780199212378.

Timmermans, Job, Yinghuan Zhao, and Jeroen van den Hoven. 2011. "Ethics and Nanopharmacy: Value Sensitive Design of New Drugs." *NanoEthics* 5 (3): 269–83. doi:10.1007/s11569-011-0135-x.

Turner, Mark. 1994. *Reading Minds: The Study of English in the Age of Cognitive Science*. Princeton, New Jersey: Princeton University Press.

———. 2006. *The Artful Mind: Cognitive Science and the Riddle of Human Creativity*. Oxford University Press. https://global.oup.com/academic/product/the-artful-mind-9780195306361?cc=ca&lang=en&.

———. 2010. "Conceptual Integration." In *The Oxford Handbook of Cognitive Linguistics*, edited by Dirk Geeraerts and Hubert Cuyckens. Oxford Handbooks Online. doi:10.1093/oxfordhb/9780199738632.013.0015.

Turner, Mark, and Gilles Fauconnier. 2000. "Metaphor, Metonymy, and Binding." *Metaphor and Metonymy at the Crossroads: A Cognitive Perspective* 356. Berlin, New York: Mouton de Gruyter: 356.

Umbrello, Steven. 2018a. "Atomically Precise Manufacturing and Responsible Innovation: A Value Sensitive Design Approach to Explorative Nanophilosophy." doi:10.13140/RG.2.2.30486.24649/1.

———. 2018b. "The Moral Psychology of Value Sensitive Design: The Methodological Issues of Moral Intuitions for Responsible Innovation." *Journal of Responsible Innovation* 5 (2). Taylor & Francis: 186–200. doi:10.1080/23299460.2018.1457401.

Umbrello, Steven, and Seth D. Baum. 2018. "Evaluating Future Nanotechnology: The Net Societal Impacts of Atomically Precise Manufacturing." *Futures* 100 (June): 63–73. doi:10.1016/j.futures.2018.04.007.

Umbrello, Steven, and Angelo Frank De Bellis. 2018. "A Value-Sensitive Design Approach to Intelligent Agents." In *Artificial Intelligence Safety and Security*, edited by Roman V. Yampolskiy, 395–410. CRC Press. doi:10.13140/RG.2.2.17162.77762.

Unger, Roberto Mangabeira. 1976. *Knowledge and Politics*. New York, NY: Simon and Schuster.

Varela, Francisco J, Evan Thompson, and Eleanor Rosch. 2017. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT press.

Warnier, Martijn, Francien Dechesne, and Frances Brazier. 2014. "Design for the Value of Privacy." In *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains*, edited by Jeroen van den Hoven, Pieter E Vermaas, and Ibo van de Poel, 1–14. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-6994-6_17-1.

Wood, Allen W. 1999. *Kant's Ethical Thought*. *Modern European Philosophy*. Cambridge: Cambridge University Press. doi:DOI: 10.1017/CBO9781139173254.

Wynsberghe, Aimee van. 2012. "Designing Robots With Care: Creating an Ethical Framework

for the Future Design and Implementation of Care Robots." University of Twente. doi:10.3990/1.9789036533911.

———. 2013. "A Method for Integrating Ethics into the Design of Robots." *Industrial Robot: An International Journal* 40 (5): 433–40. doi:10.1108/IR-12-2012-451.

———. 2016. "Service Robots, Care Ethics, and Design." *Ethics and Information Technology* 18 (4). Springer Netherlands: 311–21. doi:10.1007/s10676-016-9409-x.

Wynsberghe, Aimee van, and Scott Robbins. 2014. "Ethicist as Designer: A Pragmatic Approach to Ethics in the Lab." *Science and Engineering Ethics* 20 (4): 947–61. doi:10.1007/s11948-013-9498-4.