

Enactivism, second-person engagement and personal responsibility

van Grunsven, Janna

DOI

[10.1007/s11097-017-9500-8](https://doi.org/10.1007/s11097-017-9500-8)

Publication date

2018

Document Version

Accepted author manuscript

Published in

Phenomenology and the Cognitive Sciences

Citation (APA)

van Grunsven, J. (2018). Enactivism, second-person engagement and personal responsibility. *Phenomenology and the Cognitive Sciences*, 17(1). <https://doi.org/10.1007/s11097-017-9500-8>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Author J.B. van Grunsven

Enactivism, Second-Person Engagement and Personal Responsibility

Abstract

Over the course of the past few decades *4E approaches* that theorize cognition and agency as *embodied, embedded, extended*, and/or *enactive* have garnered growing support from figures working in philosophy of mind and cognitive science (Cf. Chemero 2009; Dreyfus 2005; Gallagher 2005; Haugeland 1998; Hurley 1998; Noë 2004; Thompson 2007; Varela et al. 1991). Correspondingly, there has been a rising interest in the wider conceptual and practical implications of 4E views. Several proposals have for instance been made regarding 4E's bearing on ethical theory (Cf. Colombetti & Torrance, 2009 & Cash, 2013). In this paper I contribute to this trend by critically examining the enactive contribution made by Giovanna Colombetti and Steven Torrance (2009) and by laying the foundations for an alternative enactive approach. Building off recent enactive approaches to social interaction, Colombetti and Torrance maintain that many of our actions and intentions “and in particular the ethical significance of what we do and mean” are “emergent from the interactions in which we participate” (2009, 523). Taking this seriously, they argue, entails a radical shift away from moral theory's traditional emphasis on individual or personal responsibility. I challenge their suggestion that accepting a broadly enactive 4E approach to cognition and agency entails the kind of wholesale shift they propose. To make my case I start by revisiting some of the general theoretical commitments characteristic of enactivism, including some relevant insights that can be gathered from Vasudevi Reddy's broadly enactive approach to developmental psychology. After that I examine both the arguments internal to Colombetti and Torrance's proposal and, in an effort to sketch the beginnings of an alternative view, I draw some connections between enactivism, the ethics of care and P.F. Strawson's work on personal responsibility. I believe that a consideration of the commonalities but also the differences between these views helps advance the important conversation concerning the link between enactivism and questions of personal responsibility in ethical theory that Colombetti and Torrance have undeniably helped jumpstart.

Introduction

Our understanding of ourselves and others as responsible agents deeply permeates our social interactions; it is reflected in common practices such as blaming and praising, accusing and forgiving, punishing and acquitting, as well as in a wide range of inter-personal affective states and attitudes such as guilt, resentment, and indignation. The specific form of these attitudes and practices and who is recognized as their appropriate target may be culturally and historically variable, but *that* they play a fundamental role in how our intersubjective lives

are textured seems difficult to deny. We might say, following P.F. Strawson, that our responsibility-practices are a pervasive natural fact of human life without which it is “doubtful whether we should have anything that we could find intelligible as a system of human relationships” (2008, 36).

The specific nature of responsibility, though, has been cashed out in a variety of different ways. Of decisive influence is often one’s (implicit or explicit) theory of human mindedness. The standard view in philosophy of mind and cognitive science states that human mindedness is marked by its representational character; we make sense of the world around us as knowers and we insert ourselves into the world as agents on the basis of discrete information-gathering and action-motivating representational mental states—paradigmatically *beliefs* and *desires*—presumed to be causally efficacious, located within an individual agent’s mind-brain, and available for an agent’s reflective deliberative choice-making activities. On the basis of this picture it seems natural to think about our practices of praising and blaming people and of holding them accountable for their comportment in the world as grounded in the connection between the relevant mental states held by an individual and the action that these mental states are thought to bring about. By examining this connection we can establish whether an individual agent possesses the right kind of ownership of their actions to be viewed as an autonomous agent who counts as a clear locus for moral praise and blameworthiness.¹

The representationalist picture of mindedness has been challenged by various theories in the field of embodied, embedded, extended, and enactive cognition [or 4E]. Though there are important differences between the various inflections of 4E, there is wide shared agreement that

- (1) **Human agents** are, in Andy Clark’s words, “natural born cyborgs” in the sense that the cognitive and affective processes through which they sustain their meaningful relationships to the world are not intracranial but are often, if not always, distributed across their brain, body and wider material and social environment (Clark 2003).
- (2) **Actions** are not necessarily the product of deliberated upon desires and intentions; rather, much of our agential lives unfold at a “pre-reflective” level; as agents we are often motivated by a perceptual grasp of what a given situation “affords” (Gibson 1979).
- (3) **Experience** is not simply the result of brute givens impinging on an agent’s sensory apparatus. Rather, it is an agential achievement; an interpretative act that correlates with an agent’s skills and reflects “the

established dynamics of interaction between perceiver and world (Noë, 2004, 215).

- (4) **Autonomy** is not what our Kantian tradition suggests, that is, it is not the ability to self-sufficiently legislate our own actions by taking up a detached rational standpoint that severs us from our habitual, pre-reflective ties to the world. Rather, living agents maintain their autonomy precisely in virtue of a constitutive dependency on their environment. As such, autonomy must be reconceived relationally (Cf. Cash 2010; Colombetti & Torrance 2009; Thompson 2007).²

My aim here is not to rehash or add to the empirical research and philosophical arguments invoked to support these commitments, nor will I discuss the various ways in which different strands in the 4E field have cashed them out. Instead I am concerned with how the 4E image of human agency, specifically as theorized along enactive lines, bears on our understanding of ethical life and in particular on the notion of personal responsibility.

The four commitments of 4E that I just identified seem to motivate questions about the nature and possibility of responsibility that echo worries expressed in traditional debates on freedom, responsibility and determinism. The exact meaning of the thesis of determinism is not without ambiguity.³ Broadly construed, however, it expresses the thought that all our current behavior is in some sense determined by external and antecedent conditions, where these conditions can be causal laws of nature, events from the past, desires, habits, inclinations and the like.⁴ Some worry that if determinism is true, then the justification for our responsibility-practices (our practices of praising, blaming, forgiving and resenting people for their actions) is undermined. This worry seems to carry over to 4E, since 4E theorists believe that human agents are natural born cyborgs who perform many of their day-to-day actions pre-reflectively and whose agential autonomy is a relational achievement that is both constrained and enabled by environmental factors that are at least in part outside their control.⁵ Indeed, though they don't express it *as* a worry, Giovanna Colombetti and Steve Torrance seem to suggest as much in their essay "Emotion and Ethics: an Inter-(en)active Approach" (2009).⁶ Building off enactive approaches to social cognition, Colombetti and Torrance aim to offer a phenomenological examination of the "myriad shared, complicit, disputed, resolved, dissolved, rebutted, etc., significances which emerge in a constantly shifting, more or less shadowy way, in any interactional situation" – an examination which reveals that "the ethical significance of what we do and mean" is "emergent from the interactions in which we participate" (2009, 520, 518 & 523). They emphasize that our interactions often gain a kind of autonomy of their own—what they call

interaction-autonomy—that enables and constrains the autonomy of individual agents and they furthermore argue that a proper acknowledgment of this demands a “shift” in how we think about individual or personal responsibility – a shift that “constrains us to defocus ... from questions of individual responsibility, exculpation, blame and praise, and encourages us to focus on the ethical qualities of the interaction itself.” To defocus from questions of individual responsibility, praise, and blame, is, according to Colombetti and Torrance, to take up “a more considered view” (2009, 522-3).⁷

My main aim here is to challenge Colombetti and Torrance’s suggestion that accepting a broadly enactive 4E approach to cognition and agency warrants a shift away from the notion of individual responsibility. To make my case I start by revisiting some of the general theoretical commitments characteristic of enactivism, including relevant insights that can be gathered from Vasudevi Reddy’s broadly enactive approach to developmental psychology. After that I examine the arguments internal to Colombetti and Torrance’s proposal and, in an effort to lay the foundations for an alternative enactive stance, I will look at some external ethical theoretical resources. Though Colombetti and Torrance helpfully situate their (inter)enactive approach to ethics within the landscape of contemporary normative ethics and metaethics, they leave out two postures, namely the ethics of care and P.F. Strawson’s second-person approach to moral responsibility, which can be fruitfully brought into dialogue with enactivism. What makes a dialogue between Colombetti and Torrance’s enactivist view, the ethics of care and Strawson’s stance particularly productive is that they share several relevant points of commonality. Most centrally, all three proposals foreground a picture of human agents understood in their affective second-person engagements with one another. At the same time, the conclusions that care ethicists and Strawson draw from this starting point when it comes to the meaning of individual responsible agency are quite different from those drawn by Colombetti and Torrance. My hope is that a consideration of the commonalities and the differences between these views helps advance the important conversation concerning the link between enactivism and ethical theory that Colombetti and Torrance have undeniably helped jumpstart. For one thing I believe it helps show that although Colombetti and Torrance’s proposal grows out of a phenomenological preoccupation with second-person interaction and the “myriad ... significances which emerge in a constantly shifting ... way, in any interactional situation”

their proposal's conclusion regarding the meaning of personal responsibility actually undermines this initial phenomenological second-person preoccupation (2009, 520).

Enactivism and the autonomous sense-making lives of living beings

Colombetti and Torrance's primary aim is to "explore some ways in which the study of social interaction, when developed on enactive lines, can illuminate the study of ethics" (2009, 516). In this section I will give a general overview of the enactive approach to cognition and agency. In the next section I discuss what it means to understand social interaction and cognition from an enactive standpoint. This allows me to then properly evaluate the specific ethical upshots Colombetti and Torrance distill from this enactive picture, especially concerning the justifiability of the notion of personal responsibility.

Enactivism is first and foremost a theory on the nature and emergence of mindedness in the natural world. Arguably, to be a minded being is to enjoy a perspective onto a world in which one reliably responds to things that are present to one as relevant. Drawing on contemporary developments in biology, enactivists argue that this key feature of mindedness can already be identified in a preliminary sense in the simplest forms of life. At the cellular level we can begin to note what Evan Thompson calls "a basic *formal organization* of life" marked by a recursive logic of *self-production*; living beings actively constitute themselves as unified centers of activity by producing their "own self-maintaining processes, including an active ... boundary that demarcates inside from outside and actively regulates interaction with the environment" (2007, 64). This, Thompson adds, makes a living organism "an individual in a new and precise sense that begins to be worthy of the term *self*" (2007, 75). What emerges in the very same gesture with life's self-production is a notion of the physical environment as a value-laden world of affordances: "an organism's world is primarily a context of significance in relation to that organism's particular manner of realizing and preserving its precarious identity" (Froese & Di Paolo 2009, 444). A living being, then, isn't simply governed by a world of heteronomous external forces acting upon it. Rather, it confronts its world on its own terms, occupying a self-regulating perspective onto an environment that is meaningfully organized in correlation with its self-maintaining activities and its history of embodied dynamic interactions with its environment.

Enactivists employ the term *autonomy* to refer to the self-constituting lives of living beings. It is important not to equate the enactive notion of autonomy with self-sufficiency. In order to retain their integrity as unified bounded centers of activity, living beings must continually incorporate and respond to features of their environment. In Francisco Varela's terms, "the living system must distinguish itself from its environment, while *at the same time* maintaining its coupling," where coupling, in its broadest sense, "refers to the necessary and permanent embeddedness and dependency of the self on the environment" (1991, 85 & 103). The autonomy of a living being is thus relationally achieved and inextricably tied to precarious dependency on, and exposure and perceptual responsiveness to environmental affordances.

Enactivists hold that it is precisely because the autonomy of living beings is precarious, because living beings are porous bodily beings—exposed to and dependent on what is other—that they are in their very being cognitive, or, in the language of enactivism, *sense-making*, world-enacting creatures (Cf. Jonas 1996). Understood as sense-making, cognition is theorized as an affectively motivated (inter)active rather than an observational-representational affair:

Exchanges with the world are inherently significant for the cognizer and this is a definitional property of a cognizing system; the creation and appreciation of meaning or *sense-making* in short. ... Sense-making is an inherently active concept. Organisms simply do not passively receive information from their environments, which they then translate into internal representations whose significant value is to be added later. Natural cognitive systems are simply not in the business of accessing their world in order to build accurate pictures of it. They actively participate in the generation of meaning in what matters to them; they enact a world (De Jaegher & Di Paolo, 2007, 488).

It is important to add that enactivism doesn't deny that we can, for wholly legitimate explanatory purposes, disregard the organism qua precarious autonomous system and instead investigate it in terms of, say, physico-chemical changes and their local effects. The point is that when we make this move, we will see very different relations:

For any system it is always possible to adopt a heteronomy or external-control perspective, and this can be useful for many purposes. Nevertheless, this stance does not illuminate – and indeed can obscure – certain observable patterns of behavior, namely, patterns arising from the system's internal dynamics rather than external parameters. ... A heteronomy perspective does not provide an adequate framework to investigate and understand this phenomenon; an autonomy perspective is needed (Thompson, 2007, 50).

According to enactivism, then, when it comes to explaining the cognitive agential lives of living beings, an external heteronomous perspective will always be limited as it skips over the primary property of the system that makes its behavior precisely cognitive and agential, namely its precarious autonomy in virtue of which it occupies a perspective onto its environment with which it interacts meaningfully.⁸ I emphasize this point about levels of explanation here because it bears on Colombetti and Torrance's approach to responsible agency. As I discuss below in *Is individual blameworthiness misguided? Colombetti and Torrance's proposal*, I think Colombetti and Torrance fail to attribute an appropriate role to the autonomy perspective of the individual agent in explaining the nature of inter-individual interaction and its ethical significance. Instead they focus almost exclusively on the impact of interaction-autonomy, sidelining the active sense-making perspectives of the individual agents caught up in inter-individual interaction. Before getting to their proposal and my specific worries about it, let's first turn to the enactive image of social interaction that they take themselves to inherit and that grows out of the enactive theory of cognition as sense-making.

Enactivism, social cognition and the priority of second-person engagement

What we have seen so far is that 1) living beings are in the business of maintaining their precarious autonomy, where 2) the ongoing activities of autonomous self-constitution both *enact* and *depend* on a perceptual-cognitive responsiveness to environmental relevance or "affordances." Hence, 3) enactivism understands cognition and perception not as a representational observational affair that is modally distinct from action. Rather, to speak with John Haugeland "[t]he primary instance [of cognition] is rather interaction, which is simultaneously perceptive and active, richly integrated in real time" (1998, 221).

Enactivists approach the emergence of human autonomous selfhood by foregrounding that "in our human case," our interactive "perceptual and motor attunement to the world ... is primarily to an environment of ... *the intentional actions of others*" (Thompson, 2007, 80, my italics). The proposal that human agents first and foremost enact, sustain, and scaffold their precarious autonomous lives via a perceptual and motor attunement to the intentional actions of others, reiterates the anti-representationalist posture of enactivism. After all, according to the standard representationalist approach to social

cognition, intentions are located inside the mind-brain of a given individual, hidden from view and unavailable for perception to take hold of (Cf. Ratcliffe 2007). The minded life of another isn't the sort of thing one perceives. Rather, social cognition is theorized as the sophisticated capacity to attribute the appropriate hidden mental states to physical behavior that is by itself psychologically bare and thought to serve as "a prerequisite for normal social interaction" (Frith & Happé, cited in Ratcliffe 2007, 5). It is worth noting the image of social life that naturally grows out of a representationalist picture of social cognition. Consider the following characterization by Daniel Dennett:

First you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in many – but not all – instances yield a decision about what the agent ought to do; that is what you predict the agent will do (1987, 17, cited in Ratcliffe 2007).

Social life, on this view, is theorized by beginning with the individual agent who, on the basis of *observation*, arrives at a point at which other people's behavior can be efficaciously *explained* and *predicted*. Now, of course figuring out the intentions, beliefs and desires of others and explaining and predicting their behavior can be terribly important in the regulation of social life. However, enactivists have widely questioned whether this third-person detached stance of observation and prediction is as foundational and pervasive as representationalist theories of social cognition have taken it to be. Shaun Gallagher, for instance, argues that much of our social life, both as adults and as developing social agents, is negotiated through second-person interactions that are facilitated by a direct perceptual grasp of the expressive bodily behavior of others, affording a range of more or less appropriate possibilities for action and response (Gallagher, 2007, 354). As I will discuss shortly, feminist thinkers from the care ethics tradition and P.F. Strawson have been equally suspicious of the move to characterize our relations to other minded beings as beginning with a detached stance of observation and prediction and they too attribute priority to second-person interactions and relations of dyadic involvement.

From an enactive standpoint, we can speak of the emergence of social cognition through dyadic second-person interaction in a two-fold sense:

- 1) Dyadic interactions are integral to the development of a wide range of *skills* that enable us to reliably identify and respond to the intentions, desires, and needs of others (more on this developmental point in a moment). And
- 2) Coordinated interactions enact social *meanings* that are not reducible to the intentions of individual agents.

In cashing out this second point, Enactivists Ezequiel Di Paolo and Hanne de Jaegher (2007) have introduced the notion of *participatory sense-making*, which Colombetti and Torrance build off in their account. Di Paolo and De Jaegher characterize participatory sense-making as follows:

[W]hat we call *participatory sense-making* ... [is] the coordination of intentional activity in interaction, whereby individual sense-making processes are affected and new domains of social sense-making can be generated that were not available to each individual on her own (2007, 497).

[Participatory sense-making] is not reducible to individual actions or intentions but installs a relational domain with its own properties that constrains and modulates individual behavior (2007, 494).

We can invoke countless examples to bring out the idea that coordinated dyadic interactions enact a relational domain of meaning that enables and constrains individual action. Think of the bickering couple that finds it impossible to break out of their established dynamic despite their individual intentions for doing so; or of a choreographed dance, performed by two trained dancers moving in unison; or of a class-room well into the semester, where the history of interactions between a teacher and her students has established a social environment of perceived opportunities for dialogical exchange. Often, then, when agents interact, the situation or domain of interaction will acquire “a life of its own,” or, in Colombetti and Torrance’s words, it will acquire its own “interaction autonomy,” where “interaction autonomy ... can be described as the way in which the interactional process that takes place between two or more agents has its own, more or less complex, dynamics, which typically has its own independent momentum” (2009, 518).

Though Colombetti and Torrance add that dyadic social interactions are characterized by a reciprocal interplay *between* agent-autonomy and interaction-autonomy, we will see (and I have already flagged) that they have a tendency to overemphasize the force of interaction-autonomy at the expense of agent-autonomy. De Jaegher and Di Paolo, by contrast, stress that while in our day to day interactions with others the possibilities for action afforded by

the situation are relationally shaped, participatory sense-making only works—*it only counts as participatory sense-making*—if both agents are at the same time continually and mutually responsive to the other’s agent-autonomy. Even the most fluid dance can only succeed if I perceive and respond to how my partner’s “autonomy demands frequent readjustment of my individual sense-making. ... I must alter my actions contextually in order to reencounter the other and in the process, sometimes, be encountered myself when her sense-making unexpectedly modulates my own” (De Jaegher & Di Paolo 2007, 504). De Jaegher and Di Paolo add that if this responsiveness were to break down the character of the relationship changes: “If the autonomy of one of the interactors were destroyed, the process would reduce to the cognitive engagement of the remaining agent with his non-social world. The ‘other’ would simply become a tool, an object, or a problem for his individual cognition” (2007, 492).

There is robust developmental support for the enactive proposal that our understanding of the minded lives of others is not in the first instance “a problem for [an observer’s] individual cognition,” but emergent from dyadic second-person interactions. Developmental psychologist Vasu Reddy argues that infants too young to have the *concept* of, say, an intention or a belief, “are not ‘just perceiving’ others’ intentional actions, but are jumping right in to imitate, respond, anticipate, and invite, adjust to and, before the end of the year, even deliberately disrupt them (Reddy, 2008, 175). The distress typically exhibited by infants when dyadic engagement is interrupted by unexpected or unwanted moments of breakdown and asynchrony is indicative of the significance of “proto-dialogical” turn-taking for neonates and developing infants (Cf. Reddy, 2008, 74).⁹ Yet mismatches and interactive breakdown are also developmentally crucial, especially when they are reliably followed by attempts on the caretaker’s behalf to invite interactional repair. When our history of interactions with others has enabled a playful or trusting relation to the ever-imminent reality of dialogical breakdown, a basis for being open to new and surprising ways of encountering the other gets established (Reddy, 2008, 83-4). Following Reddy, some enactivists argue that what is acquired through interaction and its constitutive moments of breakdown and repair is “a pre-reflective knowledge of how to deal with others – how to share pleasure, elicit attention, avoid rejection, re-establish contact. The infant acquires specific interactive schemes ... that are needed for keeping up the respective interaction” (Fuchs & De Jaegher, 2009, 481).

If this enactive developmental approach to social cognition is right, then other persons are firstly perceived as subjects who afford engagement and address and who in turn can engage and address me. Reddy hints at a link between our engaged experience of others as beings who afford address and the ethical experience of others as beings who *obligate* us: “Not only is the experience of the other person more immediate and more powerful in direct engagement, but it calls out from you a different way of being, an immediate responsiveness, a feeling in response, and an obligation to ‘answer’ the person’s acts” (Reddy, 2008, 27). In addition to being *developmentally* and, as I am starting to suggest, *ethically* prior, Reddy also explicates how our engaged second-person grasp of the other is *epistemically* richer: “it is only if we are actively involved with persons that we can perceive them as they are” (Reddy, 2008, 29). This is because, she adds, there “is a circularity ... in terms of mind knowledge: what we know of minds must depend on our engagements with them, but these engagements must themselves depend on what we know of them. ... The more you engage with other minds, the more there must be to engage with” (Reddy, 2008, 31-2).¹⁰ If this conception of our experience of other persons is right, then by taking up a third-person stance of observation, explanation and prediction towards an other we will fix in generalities or formulas something that by definition will always in part escape and transcend such formulas. We will miss not the propositional content of a person’s beliefs and desires but the manner in which she inhabits and negotiates the world as a precarious autonomous being who *continually* realizes herself as a person through her comportment with the world and with others.¹¹ A powerful illustration of this thought and of the kind of responsibility to the other that this entails can be found in the ethics of care of Eva Kittay to which I now briefly turn.

Enactive social cognition and the ethics of care – a brief interlude

In her essay “The Personal is the Philosophical is the Political” Kittay challenges the viewpoint that 1) a person’s ability to place moral obligations on us is dependent on her possession of a range of cognitive capacities, where 2) the sophistication of these capacities is directly proportional to the obligations a person is owed. This picture, defended by Jeffrey McMahan among others, has a direct bearing on how we think about the ethical significance of the lives of individuals whose cognitive capacities are severely limited – individuals such as Kittay’s cognitively disabled daughter Sesha. McMahan maintains that, if Kittay were to

remove herself from her personal attachments to her daughter, attachments tethered to a history of second-person engagements, then a detached listing of and comparison between Sesha's cognitive capacities and the capacities of, say, "a rat or a pig" would reveal to Kittay that her "daughter ... has no grounds to claim justice, and it is less bad to kill her than to kill 'one of us' (Kittay, 2009, 610 & 608). In a passage that is worth citing at length, Kittay draws attention to the reductive dehumanizing effect of McMahan's proposed method for establishing an individual's moral worth, a method that requires a stepping back from our affective ties of engagement to others and a taking up of a detached third-person attitude:

I would return to my daughter, Sesha, and find myself trying to analyze the features that differentiated her from the nonhuman animals with whom she was being compared, features that would make her worthy of [moral] personhood. As I did so, I would simultaneously shrink away in disgust from such reflections. ... [T]o respond to the challenge to articulate the differences between a human animal with significantly curtailed cognitive capacities and a relatively intelligent nonhuman animal means that one first has to see the former as the latter. That is the moment of revulsion. Relating with that stance to my daughter as my daughter is an impossibility. ... Of course, a part of the experience I am describing involves the paradox of trying to study another subject, and in so doing turn a subject into an object. I was studying my daughter and my relationship to her, and such study does seem to require an objectification that is at odds with the relationship of two subjects. (Kittay 2009, 612-3).

Echoing a point we saw earlier with Reddy, Kittay argues throughout her essay that seeing her daughter *as* her daughter and *qua* subject and preserving the "fulgent sweetness" of her "being" requires an engaged, affective, interactive attitude. In fact, not only does Kittay make the case that there is something paradoxical about attempting to perceive another subject *qua subject* by taking up a detached third-person stance of description and explanation. She also argues that an attached participant stance towards others is in fact essential for getting right those very capacities that McMahan himself prioritizes in establishing the moral worth of any given being.¹² McMahan, for instance, claims that "the profoundly cognitively impaired are incapable ... of deep personal and social relations" (cited in Kittay, 2009, 616) Kittay is able to show that this claim "is seriously misinformed" precisely by drawing on her close-personal interactions with severely cognitively disabled individuals:

My daughter now lives in a group home with five other people who are all considered severely mentally retarded, and have been so since birth. Two of her housemates lost their fathers within the period of a month. One, a young woman diagnosed with Rett's syndrome, would be found sitting with tears streaming down her face after she was told that her father was extremely ill and would die. In the case of the other, a young man who invariably greets me with a huge smile, I was witness to the howling, wailing

grief minutes after his mother and sister informed him of the death of his father. They most likely left not knowing what he had understood, and only learned of his response when they later spoke to the staff. We are speaking here of the capacity to understand the very abstract concept of death, the death of a beloved person. So much for cavalier claims that the severely retarded cannot form profound attachments (2009, 617).¹³

According to Kittay, any *epistemically responsible* judgment that we arrive at as moral agents regarding the lives of others, requires an understanding of the other that grows out of involved second-person interactions. Referencing “the close personal attentiveness” that nobel-winning scientist Barbara McClintock and primatologist Jane Goodall “devoted 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 who will, as a result, “often fail to get a glimpse into the lives of these [beings]” (Kittay, 2009, 619-20). Kittay’s conception of epistemic responsibility is intimately connected to a readiness to acknowledge the limits of what we can know about the capacities of another, especially from the disengaged armchair. Calling for a stance of *epistemic humility*, Kittay puts the point as follows:

Now [exactly] what cognitive capacities Sesha possesses I do not know, nor do others. And it is hubris to presume to know. 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. ... It is because I see Sesha close up, because I have a deep and intimate relationship with her, that I am able to see what is hidden from those who are not privileged enough to see her when she opens up to another (2009, 619).

To restate this point in Reddy’s words: “what we know of minds must depend on our engagements with them, but these engagements must themselves depend on what we know of them.” That is, how we *see* others is determined by what we *do* with them, it is shaped by our history of interactions with them, by the skills we draw on and by our readiness to be responsive to the precarious autonomous perspective they occupy. Relatedly, Sesha’s capacities and intentions qua agent cannot be isolated from the possibilities for action and self-constitution that her environment affords her, where the latter is in part constituted by the inter-individual interactions available to her and the relational domain of meaning that those interactions enact. Thus, like enactivists, care ethicists such as Kittay defend a relational conception of human autonomous selfhood. In a passage that reveals some of the

striking commonalities between enactivism and the ethics of care that I have been gesturing towards, care ethicist Virginia Held's writes:

The ethics of care ... characteristically sees persons as relational and interdependent, morally and epistemologically. Every person starts out as a child, dependent on those providing us care, and we remain interdependent with others in thoroughly fundamental ways throughout our lives. This is not to say that we cannot become autonomous; feminists have done much interesting work developing an alternative conception of autonomy in place of the liberal individualist one. ... the autonomy sought within the ethics of care is a capacity to reshape and cultivate new relations. The ethics of care ... sees many of our responsibilities as not freely entered into but presented to us by the accidents of our embeddedness in familial and social and historical contexts. It often calls on us to *take* responsibility (Held, 2006, 13-4).

Indeed, Held sees "the compelling moral salience of attending to and meeting the needs of the particular others for whom we take responsibility," as "the central focus of the ethics of care," where, to follow Kittay's point, the ability to take responsibility for the complex lives of particular others is intimately bound up with the perceptual capacities we develop through our close-personal dyadic engagements with them (2006, 10).

Is individual blameworthiness misguided? Colombetti and Torrance's proposal

We just saw that the ethics of care, with its underlying relational conception of autonomous selfhood, doesn't dismiss the notion of individual responsibility altogether. Rather, it develops out of this picture a specific way of understanding individual responsibility, where responsibility amounts first and foremost to the exercise of capacities and activities through which we attempt to be appropriately responsive to the claims that the precarious lives of particular others make on us. The view that the countless everyday concrete ways in which we interact with particular others, care for others, respond to others, have an emphatic ethical significance is also endorsed by Colombetti and Torrance. They emphasize that:

[S]ocial interaction itself ... turns out to be not just an interaction between agents who are essentially ethically neutral, where ethical considerations occasionally come in. Rather, it may be that the negotiative dance of participatory sense-making is inevitably ethical in nature: that what we participate in is, to its very bones, an ethical communal sense-making or value-making (2009, 523).

However, contra the care ethicists I have just discussed, who see a direct link between our relationally negotiated lives and the importance of taking responsibility for the particular

others whose lives are bound up with ours, Colombetti and Torrance challenge the legitimacy of individual responsibility-taking and responsibility-ascription. Let's take a look at the manner in which they arrive at this stance.¹⁴

Colombetti and Torrance begin by accusing standard approaches to normative ethics—consequentialism, deontology, and virtue ethics (as I mentioned they do not consider the ethics of care)—of being “too exclusively individuocentric in nature, too focused on the alone-in-a-crowd single agent” (2009, 517).¹⁵ What this emphasis leaves out, they add, is “any exploration of the deep ethical ramifications of the participatory, collective dynamics of human inter-relations” (2009, 517). Now, Colombetti and Torrance admittedly grant, albeit in passing, that “the appraisal of individuals” is “obviously important for ethics” (2009, 517). They furthermore add, I think rightfully so, that such individual appraisals “should take place *within the context of* a crucial consideration of the processes of interaction between agents” (2009, 15). However, in considering the “deep ethical ramifications” of these “processes of interaction” Colombetti and Torrance don't actually offer a proposal that captures how appraisals of the individual fit in with considerations of interaction-autonomy. On the contrary, they suggest that when it comes to the evaluation of agents-in-interaction, responsibility-ascriptions and stances of praise and blame are at bottom out of place. Colombetti and Torrance arrive at this conclusion while building off a concrete case that captures two distinct encounters between Audrey, a woman with advanced dementia, and two of her caretakers. The first encounter is between Audrey and an experienced care assistant. Let's take a look at the specifics:

It was mealtime in the home, and Audrey was wheeled into the lounge. Audrey ... can say only a few phrases (like ‘Go to bed’, ‘I love you’ and a few swear words) ... Audrey also has very bright blue, large, expressive, beautiful eyes. The [first experienced] care assistant ... placed the bowl of splodge on her lap, and started to put spoonfuls into Audrey's mouth—all the while looking out of the window. Audrey's eyes tracked left, right, left, right—looking for a face to hook herself into. But there wasn't one, just more faceless spoonfuls of food arriving out of nowhere into her mouth. When she had finished, the care worker stood up, stood over Audrey and wiped her mouth—still with no eye contact, and Audrey's eyes scanning her face (Perrin & May, 2000, 51. Cited in Colombetti & Torrance, 2009, 521).

The character of the second encounter is decidedly different:

Later that afternoon, when a new shift was on duty, a punk walked into the day room. She had bright blonde dyed hair, ... lots of earrings, ... Doc Martens. ... She pulled the low stool up next to Audrey, and sat astride it so that her face was right up close to

Audrey's face. Audrey beamed as her eyes hooked in with the punk's eyes. Both women were twinkling at each other. Without any words, the punk slowly and gently fed Audrey sips of tea, eyes still locked together and smiling, beaming, deeply communicating. The punk stroked Audrey's face every so often, and Audrey nuzzled in for more. The punk was new, she hadn't any training and she didn't know that she was keeping Audrey in the world for a little longer (Perrin & May, 2000, 51. Cited in Colombetti & Torrance, 2009, 521).

A care ethical analysis of these two different encounters would likely foreground that in a dyadic relationship characterized by profound asymmetry, the first nurse fails to be responsive to the care and attention that Audrey requires and invites through her expressive bodily comportment. Whatever perceptual capacities the nurse may have developed through her years of working in the field of care, at the moment described above she seems to perceive Audrey as a mouth that affords feeding, a task that affords completion, rather than as a dependent person whose ability to flourish is directly bound up with the attitudes and opportunities for interaction that she makes available to her. Whereas the experienced nurse remains experientially blind to the morally relevant needs and demands that Audrey places on her, ignoring or failing to feel what Held calls "the moral force of the responsibility to respond to the needs of the dependent," the inexperienced punk can be described as being acutely responsive to Audrey as a precarious relational being (Held, 10). And by accepting and returning Audrey's expressive invitations for connection, she directly contributes to Audrey's ability to "stay in the world for a little longer."

Initially, it seems Colombetti and Torrance are opting for a similar reading of the case. They begin by urging that the "ethical character" of the first situation is marked by a third-person structure: "the care assistant treats Audrey as a 'her'" (or perhaps even an 'it') "rather than a 'you' (or a 'thou')." There is a disconnect between the participants: a failed search for connection on the part of Audrey, and a *refusal* to connect on the part of the carer" (Colombetti & Torrance, 2009, 521, my italic). It quickly becomes clear, however, that Colombetti and Torrance deem it misguided to view the experienced nurse's "refusal to connect" as something for which she carries personal responsibility. Instead, what Colombetti and Torrance take to be the salient take-away point of both the "third-person" encounter and its second-person counterpart is that our ethical evaluations should be primarily directed at "the quality of the interaction[s]," where a consideration of how the agents contribute to those interactions individually seems to be a secondary matter (2009, 521). Focusing on interaction processes, they specify, should inhibit us from viewing "the

first carer as having not considered carefully enough the way to achieve the best results for the patient, or as following bad rules, or as having a defective character” (2009, 522). Such traditional normative ethical readings of the situation, they add “would ‘blame’ [the experienced nurse] for poor management of the care situation and for poor treatment of the ‘blameless’ patient. She ‘fails’ to provide the kind of affective stimulation that Audrey needs” (2009, 522). Colombetti and Torrance take this to be a misguided evaluation of the situation. Such an emphasis on personal responsibility is seen as reductive, as failing to give due credit to the irreducibly inter-personal character of human interaction:

A more considered view would see the blocked, frustrated, character of the first episode as emerging from the way the two participants *mutually engage* with each other, and also as coloured, perhaps, by the way a history of previous episodes has shaped the two participants’ responses in the present. To take this less individuocentric approach would be *to distance oneself from making judgments of individual responsibility* (2009, 522, my italics).

Let me start by agreeing with Colombetti and Torrance that it is an important upshot of an enactive approach to agency and interaction that we should pay more attention to the inter-personal relational domains within which actions unfold and acquire their meaning. What is more, this can have the ethically important potential to relieve individual agents from an overly liberalist notion of individual moral responsibility, according to which your actions are never anything other than strictly yours, as they are the culminations of your individual intentions and desires. Nevertheless, there are a couple of things about Colombetti and Torrance’s stance that I find theoretically questionable as well as morally troublesome:

- (1) Note how they characterize not just the second interaction but also the first encounter as a form of *mutual engagement* or participatory sense-making. However, as we saw earlier with De Jaegher and Di Paolo, participatory sense-making only counts *as* participatory sense-making to the extent that both agents continually track and are responsive to the other as a *you* rather than as “a tool an object, or a problem for [one’s] individual cognition” (2007, 492). By granting that, in the first case, the nurse treats Audrey precisely not as a you but as a ‘her’ (or, as I suggested, as a task-to-be-completed) Colombetti and Torrance seem to offer two conflicting readings, where the same encounter both fails and succeeds to meet the necessary requirements to count as a form of mutual engagement.

- (2) Having labeled both encounters as instances of mutual engagement, Colombetti and Torrance conclude that both encounters require an evaluative stance that focuses first and foremost on the interaction itself, where doing so will allegedly motivate us to adopt an evaluative stance of non-judgment (“to take this less individuocentric approach would be *to distance oneself from making judgments of individual responsibility*”). The idea seems to be that in noticing that interactions are never reducible to the individual intentions of each participating agent, we also ought to realize that we are never quite warranted to place blame on individual agents-in-interaction. But this all-or-nothing stance seems to undermine Colombetti and Torrance’s own phenomenological concern with the “myriad shared, complicit, disputed, resolved, dissolved, rebutted, etc., significances which emerge in a constantly shifting, more or less shadowy way, in any interactional situation” (2009, 520). One would expect that a theory that foregrounds this will be, to speak with Virginia Held, “sensitive to contextual nuance and particular narratives,” where responsibility-ascription presumably gains traction in some of those narratives and loses its legitimacy in others (2006, 10). Colombetti and Torrance’s proposal ultimately ends up “making the abstract and universal claims of” some of the “more familiar moral theories” that they themselves criticize (Held 2006, 10). That is, while they point out the important difference between second and third-person encounters, and while they furthermore explicitly set out to capture how human interaction and “human ‘feelings of connectedness’ involve a complex interplay of various levels of empathy and other-grasping,” they simultaneously wipe out these important distinctions and gradations, treating all inter-personal encounters as inviting the same kind of evaluative stance of non-judgment.
- (3) As we have seen, the proposal that we should “distance” ourselves “from making judgments of individual responsibility” follows from Colombetti and Torrance’s suggestion that each interaction is constituted by the contributions made by all parties involved. From an ethical standpoint, this is a deeply problematic move. By essentially flattening out the distinction between second-person engagements and third-person encounters and what it means to evaluate them properly, Colombetti and Torrance’s proposal opens the door to a characterization of not just vulnerable Alzheimer patients but also of rape-victims and other victims targeted in acts of

violence or limited in their agency through oppressive social environments as playing their equal part in their own degradation.¹⁶

(4) Perhaps my reading of the encounter between Audrey and the first nurse is clouded by the way in which standard ethical theory has shaped my evaluations, but I find it nearly irresistible to interpret the situation as one in which the nurse is failing both at a social and ethical level – that is in her role as a nurse and in her direct human response to Audrey as a living breathing dependent autonomous person. A nurse who systematically dissociates herself from her role and the obligations it entails should be held responsible for continuing to occupy that role knowing she can no longer bring herself to fulfill. And I think we would not feel surprised if we learned that the nurse, reflecting back on the long day, could feel a sense of moral remorse in how she treated Audrey. Of course it could turn out that *excusing conditions* apply to this particular situation and in that case we may want to withhold from blaming the nurse for her “refusal to connect” with Audrey. But excusing conditions are *excusing* for a reason, namely because they bring out features of a situation, or facts about an agent, that makes the ascription of blame lose traction where it is normally wholly appropriate (perhaps we find out that the nurse has been forced to work in spite of an overpowering burn-out). What Colombetti and Torrance are effectively suggesting, however, is that the excusing stance is applicable tout court; that to notice that individual actions and intentions are constrained and shaped by the relational domains within which they unfold *is* to realize that a defocusing away from praise and blame is always “the more considered view.” Shortly I will turn to Strawson’s work to raise an objection to this conclusion, which, I believe, is in an important sense antithetical to Colombetti and Torrance’s own commitment to the priority of second-person engagement.

(5) As we have been seeing, Colombetti and Torrance seem to believe that a prioritization of interaction-autonomy over agent-autonomy always delivers the “more considered” evaluation of face-to-face inter-individual encounters. But doesn’t this depend, by enactivism’s own lights, on what it is we are aiming to illuminate about a given encounter? It is true, of course, that enactivism understands agent-autonomy in a manner that simultaneously requires us to look at the environment in which a given agent is embedded. Thus, an investigation of agent-

autonomy inadvertently bleeds into an investigation of interaction-autonomy. But, as we saw in *Enactivism and the autonomous sense-making lives of living beings*, it is integral to the enactive framework that a proper understanding of cognition and agency begins with a conception of living beings as autonomous selves who “confront their environment on their own terms;” how living agents respond to their environment is a reflection of how they perceive their environment, where perception is in turn an agential achievement internally linked to an agent’s skills and concerns (Thompson 2007, 70). But if this is true, it isn’t clear to me how we arrive at “a more considered view” of instances of inter-individual interaction by downplaying the extent to which these interactions are shaped by the ways in which the autonomous perspective of each interactor informs the affordances they respond to.¹⁷ Indeed, it seems plausible that by refraining from ascriptions of personal responsibility in contexts of participatory sense-making, we will obscure rather than “illuminate ... certain observable patterns of behavior” characteristic of typical adult forms of coordinated interaction (Thompson 2007). Or, to re-cite Strawson, it might be that if we remove our understanding of ourselves and our perception of others as responsible agents from our everyday ongoing interactions with one another it is “doubtful whether we should have anything that we could find intelligible as a system of human relationships” (2008, 36).¹⁸

Now, perhaps an approach to responsibility that focuses on the ongoing everyday ways in which we take responsibility for the (often precarious dependent) lives of others, doesn’t seem fully satisfactory to Colombetti and Torrance because they are concerned with the *justification* of the idea of individual responsibility. “Sure,” they might wager, “we may feel obligated to take responsibility for the lives of particular others, but is this sense of obligation justified? Or does a 4E conception of human agency show that while we may *feel* and *take* responsibility continuously we actually *aren’t* responsible as individual agents? After all, being a human agent, 4E style, means always being situated, it means always being shaped by others through ongoing close-coupled interaction processes and as such we lack the kind of autonomy that seems to be a necessary enabling condition for genuinely *being* responsible.

I want to conclude this paper by arguing that enactivists need not maintain that a 4E image of agency undermines the justification for our responsibility-practices. In fact, as I have already flagged, the manner in which Colombetti and Torrance call into question the

justificatory ground for our responsibility-practices causes them to partially undermine their own commitments to a second-person phenomenologically oriented approach to human interaction, or participatory sense-making, and its ethical significance. I will make my case through a brief consideration of points made by P.F. Strawson in his seminal essay “Freedom and Resentment.”

Strawson on responsibility, determinism, and the second-person standpoint

Strawson’s view of the nature and locus of responsibility grows out of his engagement with the thesis of determinism. As I discussed in the introduction, the thesis of determinism states that an agent’s current behavior is in some sense determined by external and antecedent conditions, where these can be causal laws of nature, desires, events from the past, etc. (as I also sketched in the introduction, there is relevant overlap between the way in which the thesis of determinism and the 4E thesis might seem to threaten the justification of our responsibility practices). One posture towards the nature and possibility of responsibility generated by deterministic considerations is that of the “pessimist,” as Strawson calls him.¹⁹ Strawson’s pessimist believes that some form of unconstrained autonomous agency is a necessary condition for our responsibility-practices and that autonomy thus construed is incompatible with the validity of determinism, such that, *if* determinism is true, *then* our responsibility-practices lose their enabling condition and are thus unjustified. Colombetti and Torrance seem tempted by a version of this thought. After all, they suggest that the truth of 4E with its conception of autonomy as situated and profoundly shaped and enabled by the relational domains within which human actions unfold makes the notion of personal responsibility essentially misguided or, at best, reductive; it is “the more considered view” to defocus from questions of individual praise and blame and instead focus on the quality of interactions themselves.

In making this point, the particular sub-species of pessimism that Colombetti and Torrance come close to embracing is *hard-determinism*, where the hard-determinist believes in the validity of the thesis of determinism and hence denies that human beings are free responsible agents. One of the problems with hard-determinist pessimism is its counter-intuitiveness. To view all human actions as mere causal happenings and to declare our practices of blaming, praising, resenting, forgiving and punishing null and void, seems to go

against our pervasive everyday pre-theoretical self-conception as it maintains that we are not the kinds of beings we take ourselves and others to be. Now of course Colombetti and Torrance aren't quite like the hard-determinist pessimist. After all, as thinkers who inherit the phenomenological tradition and who are explicitly concerned with the "myriad shared, complicit, disputed, resolved, dissolved, rebutted, etc., significances which emerge in a constantly shifting, more or less shadowy way, in any interactional situation" they seem committed to attribute precisely the kind of priority to our everyday pre-theoretical self-conception that the hard-determinist delegitimizes (Colombetti and Torrance 2009, 520). And yet, as I hope to indicate using Strawson's analysis, by simultaneously suggesting that the truth of 4E challenges the meaning of our responsibility practices Colombetti and Torrance inadvertently undermine the priority of this pre-theoretical experiential perspective.

The view that our responsibility practices could lose their justificatory ground in virtue of the truth of a theoretical discovery external to those practices is a view that, according to Strawson, "over-intellectualizes the facts" needed to ground our responsibility practices. Strawson argues that a different approach to our responsibility-practices begins to emerge when we remind ourselves of something "it is easy to forget when we are engaged in philosophy, especially in our cool, contemporary style," namely "what it is actually like to be involved in ordinary interpersonal relationship, ranging from the most intimate to the most casual" (2008, 23). Much like enactivists working on social interaction, Strawson thus begins by calling attention to the various manifestations our everyday pre-theoretical inter-personal relationships and by noting that our affective second-person responses or to one another, as Strawson terms them, our "reactive attitudes," play a constitutive role in the negotiation of these different relationships. (2008, 23) Think, for instance, of the resentment you might feel when a waiter's unresponsiveness to your attention-catching attempts seems to stem from an intentional effort to ignore you, motivated, perhaps, by seemingly racist sensibilities. According to Strawson, reacting to this situation with resentment just is a way of targeting the waiter as someone who has intentionally violated the normative expectations for interpersonal regard that we have in our ongoing everyday engagements with one another; it is a pre-theoretical way of identifying the waiter as a moral agent whose action is appropriately susceptible to judgments of blame. We can, of course, be wrong about relevant specifics of the situation. Imagine, for instance, that we come to learn that the waiter's behavior stemmed not from malicious racist motives, but from innocuous inexperience (he

simply doesn't yet possess the quick room-scanning abilities of an experienced waiter). Uncovering such information should have a direct impact on our reactive attitudes. Resentment will no longer feel appropriate and neither will the ascription of blame that the experience of resentment precipitates.

The significance of this point and what it indicates about the locus of our responsibility practices is fleshed out in detail in Strawson's *argument from excusing conditions*, where excusing conditions refer to circumstances under which reactive attitudes "do or do not seem natural or reasonable or appropriate" (2008, 23). Excusing conditions come in a variety of forms. We can speak, for instance, of *injury*-excusing conditions when we find out the agent acted in some sense involuntarily ("he was pushed" "he was forced to"), unknowingly ("He didn't know," "He didn't realize"), or under exceptional circumstances that reasonably trump his accountability ("He had to do it," "It was the only way") (Strawson 2008, 23). What is specific to injury-excusing conditions is that though they excuse the agent for the harm caused by the particular action under consideration, they tend not to affect our conception of the agent as morally responsible tout court (we will still resent and be disposed to blame our waiter for any genuine intentional ill will he may go on to express moments later). *Agency*-excusing conditions, by contrast, are those in which a wide range of our reactive attitudes no longer feel appropriate in response to the agent herself. In such cases, Strawson writes, "the circumstances ... present the agent as psychologically abnormal – or as morally undeveloped. The agent was himself; but he is warped or deranged, neurotic or just a child. When we see someone in such a light as this, all our reactive attitudes tend to be profoundly modified" (2008, 24). The modification Strawson speaks of entails a transition from a participant attitude—which comes with the full range of expectations, demands, and reactive feelings to which we are prone in our typical adult relationships and interactions—to an attitude Strawson terms as "objective," where to inhabit the objective attitude is to view a person from a standpoint of explanation, prediction and manipulation, say as "an object of social policy' as a subject to be managed, or handled, or cured, or trained" (2008, 25).²⁰

Note how, much like enactivists who challenge the commonplace that social cognition is first and foremost grounded in a third-person stance of observation, explanation and prediction, Strawson maintains that a third-person attitude towards other persons indicates a modification of, and thus presupposes, the more pervasive orientation we have

toward one another when we occupy a participant attitude. Strawson adds that our susceptibility to the second-person reactive attitudes in response to another's "good or ill will or indifference or lack of concern" expressed towards us tends to be interconnected with two important analogues. Firstly, most if not all of us who are susceptible to second-person reactive attitudes will also experience feelings of guilt, shame, remorse etc. when we realize we ourselves have flouted the kind of regard another warrants from us. And secondly, we will also be capable of vicariously experiencing resentment on behalf of others when their claims to regard have been refused or otherwise violated. Thus, when a vulnerable Alzheimer's patient is treated by her care assistant without regard for her desire for attention and connection, we might feel a sense of moral indignation (Strawson's term for vicarious resentment) on behalf of the patient and we might expect a sense of remorse or retroactive shame on behalf of the nurse. Much like the second-person reactive attitudes, their self-reflexive and vicarious analogues are generally modulated or inhibited upon the discovery of injury and agency-excusing conditions (the discovery of deplorable work-conditions and an impending burn-out make a caretaker's callous behavior explainable in a way that modulates the responsibility we are, under normal conditions, prone to ascribe to her).

As we have been seeing, the discovery of relevant excusing conditions entails a bracketing—sometimes momentarily and sometimes permanently—of our involved attitude and of our view of the other as a full-fledged participating member in inter-personal moral interactions with their corresponding expectations of mutual responsiveness. Instead we shift to the objective stance, from which we go on to explain or predict a person or a bit of their behavior. The crucial question Strawson asks is whether the validity of the thesis of determinism, or of any theoretical truth, such as that of 4E, would require us to permanently take up the objective stance and warrant a *total* repudiation of the reactive attitudes? Would it require us to detach from "that complicated web of attitudes and feelings which form an essential part of the moral life as we know it, and which are quite opposite to objectivity of attitude ... would, or should, it mean the end of gratitude, resentment, and forgiveness; of all reciprocated adult loves; of all the essentially personal antagonisms?" (Strawson 2008, 24-5). To rephrase the question in terms of Colombetti and Torrance's example and the proposal they distilled from it: would accepting the 4E thesis entail that, no matter how different the interactions between Audrey and her two caretakers are and no matter how different our vicarious reactive attitudes to these two situations may be, that we are doing more justice to

the truth of the situation when we neither praise the punk for her behavior nor blame the nurse? Would it entail that we should adopt a universal evaluative stance of non-judgment, as Colombetti and Torrance seem to believe?

Strawson rejects this conclusion by making the following point: because the agency-excusing conditions that compel us to bracket our reactive attitudes are by definition conditions that capture a *deviation* from our typical modes of social interaction, a permanent bracketing of the reactive attitudes cannot be the upshot of the thesis of determinism or any thesis for that matter. After all, “it cannot be a consequence of any thesis which is not itself self-contradictory that abnormality is the universal condition” (2008, 28). To expand on the claim that the truth of determinism neither “could” nor “should” propel “us always to look on everyone exclusively in th[e objective] way,” Strawson furthermore advances the broadly Wittgensteinian view that the reactive attitudes are simply a part of our natural history (2008, 26). They are, to put it in enactive terms, part and parcel of the developmentally and epistemically primary ways in which we come to have an understanding of what it means to be an intentional human agent at all. Although Strawson grants that this doesn’t make it *logically* inconceivable that the theoretical validity of the truth of determinism could lead to the evaporation of our reactive attitudes, he proposes that this radical revision of human life as we know it “is, for us as we are, practically inconceivable” (2008, 26).²¹ That is, Strawson wagers that:

The human commitment to participation in ordinary inter-personal relationships is, I think, too thoroughgoing and deeply rooted for us to take seriously the thought that a general theoretical conviction might so change our world that, in it, there were no longer such things as inter-personal relationships as we normally understand them; *and being involved in inter-personal relationships as we normally understand them precisely is being exposed to the range of reactive attitudes and feelings that is in question*. ... A sustained objectivity of attitude, and the human isolation which that would entail, does not seem to be something of which human beings would be capable, even if some general truth were a theoretical ground for it (2008, 26).²²

As it happens, enactivism qua theory endorses and tries to further illuminate precisely this “human commitment to participation in ordinary inter-personal relationships.” As such one would expect Colombetti and Torrance, who inherit and contribute to the enactive second-person approach to social cognition, to wholeheartedly endorse this Strawsonian point. And yet, what a consideration of Strawson’s project has helped bring out is that their proposal regarding the evaluation of agents-in-interaction essentially

entails the universalization of those excusing conditions that modify much of the participant attitudes that are part and parcel of our intersubjective lives as we know it. In other words, and this is my concern, there seems to be an odd oscillation in Colombetti and Torrance's account between the objective and the engaged perspectives and the priority attributed to them in the capturing of our interactional lives with one another. While Colombetti and Torrance place primary emphasis on second-person experience, they simultaneously seem to suggest that this experience insofar as it gives us a sense of the other as someone for whom I can take responsibility—and a sense of myself as someone who is called upon to take responsibility for the other—is at bottom misguided. Ultimately, their suggestion that the “more considered” view of personal responsibility warrants that we “distance” ourselves tout court “from making judgments of individual responsibility” entails a turning away from the phenomenology of second-person involvement and a privileging of a third-person approach to personal responsibility in interaction.

In treating the notion of individual responsibility as misguided or limited at best Colombetti and Torrance seem to presuppose that ascriptions of individual responsibility can only mean something on the basis of precisely the methodologically individualist representationalist conception of agency that they themselves emphatically criticize.²³ In taking this stance they are overshooting their target. What I have tried to indicate is that there are other options available that share several important starting points with the enactive approach to social interaction embraced by Colombetti and Torrance; both Strawsonian and care ethical approaches view responsibility as a stance or role that we as relational subjects can succeed or fail to inhabit with respect to the lives of others with whom we are bound up. To take this thought in an enactive direction: typically developed human agents possess a refined know-how of what it is to invite, thwart, sustain, interrupt, and promote second-person interactions with others. Correspondingly, we possess an intimate grasp of how the others with whom we interact (and how we ourselves) can succeed or fail to maintain their autonomy within the context of a given interaction. Even if much of what we do is shaped by our environment; if our intentions aren't ‘our own’ in the traditional sense; if getting others right is a messy and endless task because we ourselves and the other are constantly shaped in relation, still, to interact with others, to engage in the “negotiative dance of

participatory sense-making” is to have an experiential grasp of the other as a being who makes claims on us, who obligates us, and who we can ‘block from view’ by disengaging. To be engaged in genuine interaction is to see the other, at least during the interaction itself, as someone to whom I am in some sense obligated to respond, and thus to know myself as someone who can take responsibility.

As I said before, I grant that Colombetti and Torrance’s emphasis on the interpersonal relational domains within which actions unfold and acquire their meaning can have the ethically important potential to relieve individual agents from an overly liberalist notion of individual moral responsibility, according to which your actions are never anything other than strictly yours, as they are the culminations of your individual intentions and desires. But while the developments in embodied embedded cognition that Colombetti and Torrance inherit and contribute to may put pressure on this broadly liberalist way of thinking about the locus of individual responsibility they do not commit us give up on its meaning altogether. On the contrary, an enactive approach to social interaction understood as participatory sense-making and the idea of mutual responsiveness to the other’s precarious autonomy that this entails, can, especially when combined with Reddy’s developmental account, help further our understanding of how our second-person engaged experience of the other is constitutive of what it means to be a morally responsible agent. By providing empirical evidence for the view that, starting in early infancy, typically developed human agents perceive other persons as precarious autonomous sense-making beings who afford engagement and whose claims on us we can fail and succeed to be responsive to, enactivism can help expand on the Strawsonian and care-ethical view that an involved participant orientation towards others is primary not just in a developmental and an epistemological, but also in an ethical sense. And by offering careful phenomenological analyses of the “myriad shared, complicit, disputed, resolved, dissolved, rebutted, etc., significances which emerge in a constantly shifting, more or less shadowy way, in any interactional situation” enactivism can help nuance and provide further support for the view that it is “only by attending to this range of” intersubjectively negotiated significances and second-person attitudes that we can “recover ... a sense of what we mean, i.e. of all we mean, when speaking the language of morals, we speak of desert, responsibility, guilt, condemnation” (Strawson, 2008, 24).

Bibliography

Ayer, A.J. 1954. "Freedom and Necessity," in *Philosophical Essays*, St. Martin's Press, New York, 1954

Benson, Paul. 2000. "Self-Worth and the Social Character of Responsibility," in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Ed. Catriona Mackenzie & Natalie Stoljar). Oxford University Press

Cash, Mason. 2010. "Extended Cognition, Personal Responsibility, and Relational Autonomy," in *Phenomenology and the Cognitive Sciences*. 9:645-671

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. MIT Press

Clark, Andy. 2003. *Natural Born Cyborgs: Minds, Technologies and the Future of Human Intelligence*. Oxford University Press

Colombetti, G., & Torrance, S. (2009). Emotion and ethics: An inter-(en)active approach. *Phenomenology and the Cognitive Sciences*, 8(4), 505–526.

Dreyfus, H. L. (2005). "Overcoming the Myth of the Mental: How Philosophers can Profit from the Phenomenology of Everyday Expertise." *APA Pacific Division Presidential Address*, 1–24.

Dreyfus, H.L. 2009. "The Myth of the Pervasiveness of the Mental," unpublished keynote talk at *The Phenomenological Body – Its Spaces and Limits*, graduate student conference at the New School for Social Research (March 2009)

Di Paolo, E., Rohde, M., & De Jaegher, H. (2010). Horizons for the enactive mind: Values, social interaction, and play. In J. Stewart, O. Gapenne, & E. A. Di Paolo (Eds.), *Enaction Toward a New Paradigm for Cognitive Science* (pp. 33–88). Cambridge, MA: MIT Press.

De Jaegher H, and Di Paolo E (2007). "Participatory Sense-Making: An enactive approach to social cognition." *Phenomenology and the Cognitive Sciences*, 6(4), 485-507.

Froese, T., & Di Paolo, E. A. (2009). Sociality and the life–mind continuity thesis. *Phenomenology and the Cognitive Sciences*, 8(4), 439–463.

Fuchs, T., & Jaegher, H. D. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8(4)

- Gallagher, Shaun. 2005. *How the Body Shapes the Mind*. Oxford University Press
- Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2(3-4), 353–365.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin.
- Haugeland, John. 1998. "Mind Embodied and Embedded," in *Having Thought: Essays in the Metaphysics of Mind*. Harvard University Press.
- Held, Virginia. 2006. *The Ethics of Care: Personal, Political, and Global*. Oxford University Press.
- Hume, David. 200). *A Treatise of Human Nature*. (D. F. Norton & M. J. Norton, Eds.) Oxford University Press.
- Hurley Susan. 1998. *Consciousness in Action*. Harvard University Press.
- Jonas, Hans. 1996. "Evolution and Freedom: on the Continuity among Life-Forms." In *Mortality and Morality: A Search for the Good After Auschwitz*. Northwestern University Press.
- Kittay, Eva. 2009. "The Personal is Philosophical is Political. A Philosopher and Mother of a Cognitively Disabled Person sends Notes from the Battle Ground." *Metaphilosophy* 40 (3-4):606-627
- Noë, Alva. 2004. *Action in Perception*. Cambridge, Mass: MIT Press
- Noë, Alva. 2010. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. Hill and Wang. New York, NY
- Ratcliffe, M. (2007). *Rethinking commonsense psychology : a critique of folk psychology, theory of mind and simulation*. Basingstoke New York: Palgrave Macmillan.
- Strawson, P.F. 2008. "Freedom and Resentment." In M. McKenna & P. Russell (Eds.), *Free Will and Reactive Attitudes*. Ashgate Publishing Company.
- Reddy, Vasudevi. 2008. *How Infants Know Minds*. Harvard University Press.
- Thompson, Evan. 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard (Belknap)
- Wiggins, David. 2003. "Towards a Reasonable Libertarianism," reprinted in *Free Will* (ed. Gary Watson). Oxford University Press,
- Varela, Francisco. 1991. "Organism: A Meshwork of Selfless Selves," in Alfred I. Tauber (ed.) *Organism and the Origins of Self*, Boston Studies in the Philosophy of Science, Springer Science
- Varela, F., Thompson, E., & Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

¹ There are various proposals on how we should cash out the nature of this connection. Two helpful discussions can be found in Paul Benson (2000) and Mason Cash (2010).

² This list of commitments is not meant to be exhaustive. I have chosen to foreground these aspects of 4E because they most clearly apply to the discussion below, concerning the nature and possibility of responsible agency.

³ Strawson indicates as much when he says to be of “the party of those who do not know what the thesis of determinism is.” (1) See also, A.J. Ayer, 1954 (esp. 117-8) & Wiggins’ (2003).

⁴ Cf. A.J. Ayer, (1954)

⁵ In a paper delivered at a 2009 conference at The New School for Social Research, Hubert Dreyfus struggled with this very issue when he tried to make room for responsibility within his specific conception of unreflective action (or “absorbed skillful coping”).

⁶ I say “indirectly” because Colombetti and Torrance do not explicitly engage with the traditional freedom and determinism debate. As we will see, though, some of the conclusions they draw echo conclusions drawn in that debate.

⁷ I use the terms individual and personal responsibility interchangeably throughout the paper.

⁸ Varela gives the following example: “a bacterium swimming in a sucrose gradient” can be “analyzed in terms of the local effects of sucrose on membrane permeability, medium viscosity, hydromechanics of flagellar beat, and so on” (Varela, 1991, 85). However, “the sucrose gradient” and the bacterium’s “flagellar beat are interesting to analyze only because the entire bacterium points to such items as relevant ... Remove the bacterium as a unit, and all correlations between gradients and hydrodynamic properties become environmental chemical laws, evident to us as observers but devoid of any special significance” (Varela, 1991, 85).

⁹ Echoing two points we just saw with De Jaegher and Di Paolo, Reddy emphasizes that (1) “dialogue create[s] and transform[s] the individual and the realm of meaning,” and (2) it “demonstrate[s] the recognition of another being. You would not have a dialogue with someone unless you took for granted, or at least hoped, that he or she was a minded being, capable – at some level – of understanding and responding” (Reddy, 2008, 66). Building off Colwyn Trevarthen’s proposal, Reddy argues that there are convincing reasons to attribute at least four structural features of adult dialogical interaction to dyadic non-verbal infant-caretaker “proto-communication,” namely “*communicative acts* (such as expressions, words, gestures), *self-synchrony* (the ability to produce organized and coherent actions), *interactional and affective synchrony* (the ability to relate your own actions and emotions to the other’s actions and emotions), [and] *turn-taking* (an ability to take turns in acts)” (2008, 71). Numerous other developmental psychologists agree that by the time that they are 2 months-old, human infants engage in proto-dialogues with their caretakers (Cf. Mary Catherine Bateson; Lynne Murray; Daniel Stern; Colwyn Trevarthen; Ed Tronick; as discussed in Reddy, 2008, 68-89).

¹⁰ Likewise, not only is there more of the other that we come to see in the process of engagement, but we ourselves are affected by the interaction process; we are “constantly being re-shaped as an entity in relation and ... gradually building up awareness of ... [ourselves] in these relations” (Reddy, 2008 148-9).

¹¹ I do not want to deny that this can sometimes be for wholly legitimate purposes (think of the Surgeon operating on a patient or the psychoanalyst analyzing, explaining and predicting the behavior of his patient).

¹² Kittay herself rejects the capacities-based approach to moral worth advocated by McMahan and defends a relational approach instead (see 2009, especially 623-5).

¹³ As Kittay notes in a footnote: “it is not unreasonable, in the case of this young man, that he held back his grief to spare his mother and sister.”

¹⁴ Colombetti and Torrance’s enactive contribution to ethical theory is self-reflectively programmatic and tentative: “What we are presenting here must be seen as very much a first outline sketch” (516). In offering this sketch Colombetti and Torrance take up a maximally broad perspective: they engage with debates in *metaethics* by touching on non-cognitivism and by presenting “the enactive view” as providing “a fresh perspective on the traditional ‘reason versus emotion’ dialectic in ethics.” They furthermore gesture towards connections between enactivism and *moral psychology* by gesturing at links between enactivism and Spinoza’s ethics of autonomous self-emancipation, as well as conceptions of moral accountability. Finally, they suggest enactivism represents an approach to moral evaluation that differs from mainstream utilitarian, deontological, and virtue ethical theories

in *normative ethics*. Though this broad approach may further Colombetti and Torrance's programmatic goal of indicating a variety of ways in which enactivism offers insights that are currently underdeveloped in ethics it also has a clear downside. This is because this maximally broad perspective requires Colombetti and Torrance to skip over some of the important details needed to gauge how novel their enactive approach genuinely is (For instance, virtue ethics already challenges the "reason emotion dichotomy" that Colombetti and Torrance speak of, and one could add furthermore that this dichotomy trades on a strawman image of Kantian ethical theory. Furthermore, as I point out in this paper, Colombetti and Torrance leave undiscussed the striking commitments shared by enactivists and thinkers from the ethics of care. Someone who does discuss the link between feminist ethics of care and the field of embodied cognition is Mason Cash (2010).

¹⁵ Colombetti and Torrance initially present their "ethical account centering on participatory sense-making and interaction ... as supplementing other accounts, as making good some deficiencies and silences in those accounts, rather than as supplanting them" (517). They suggest that three of the most prominent normative ethical theories currently on offer each get something right about ethical life and that "a more reasonable view is to see each approach as offering a distinctive and important contribution to an overall picture." It seems highly questionable to me that the commitments of a virtue ethicist, an act-utilitarian and a Kantian can be unified into one coherent "overall picture," since they operate with decisively different conceptions of human nature, the springs of human action, and what gives those actions moral worth. What is more, as Colombetti and Torrance advance they seem to veer in a more critical direction that rejects what they see as the starting-point of Kantian, Utilitarian, and Virtue ethics: namely the image of the "alone-in-a-crowd single agent."

¹⁶ I want to thank Hanne Jacobs for drawing my attention to this problematic upshot.

¹⁷ Indeed, whereas Colombetti and Torrance urge us to defocus from responsibility-ascriptions at the level of individual agents, Di Paolo et al (2010) insist that "because an enactive approach places great importance on the autonomy of the individuals involved [in interaction], this approach to social cognition, while focusing on the interaction process, paradoxically also gives social agents an autonomy and role that has not been thematized before: that of participation in contrast to mere observation" (72).

¹⁸ Although I don't want to overintellectualize the facts and characterize infants, people with severe cognitive disabilities and advanced Alzheimer's disease as robustly responsible agents, it might not be implausible to understand the activity of turn-taking characteristic of even pre-verbal second person dialogical interaction as a precursor of a genuine taking and attributing of responsibility. After all, turn-taking entails a targeting of the other and a relating to oneself as an intentional agent who can succeed or fail to play their part in the promotion of interaction.

¹⁹ The thesis of determinism has generated roughly three different postures regarding the nature and justifiability of our responsibility, referred to by Strawson as that of the skeptic, the optimist and the pessimist respectively. Since a version of the pessimist's posture is operative in Colombetti and Torrance's proposal I will, for my current purposes, restrict myself to a brief discussion of the pessimist stance and Strawson's critical engagement with it. For a 4E proposal that opts roughly for the pessimist's counter-view, namely the optimist's approach, see Mason Cash (2010). Cash argues that taking responsibility (and not avoiding it altogether on the basis of a theoretical insight that our agency might be at its heart "morally impaired") is something agents will continue to do because "A defense of 'I cannot help it, my genes and society made me that way' or 'my environment made me do it' abdicates the responsibility for attempting to improve and to earn the right to be trusted, a right that a society otherwise accords to all members who have shown themselves capable of living up to the responsibility that this trust assumes." (2010, 652) Membership in a community of trusted inter-actors is thus at stake. Freedom and responsibility are "accorded" on the basis of an "earn[ed] ... right to be trusted," which we achieve by claiming ownership where our communal norms tell us it is appropriate to take ownership: "recognition of this potential for improvement, and thus for taking control over factors that otherwise might be external influences on one's actions is crucial. It can enable us to admit that the potential for systems wider than individual agents to produce actions does not automatically absolve individuals of all responsibility. A forward-looking justification for practices of punishment would support sometimes applying relevant sanctions even to people with a history or environment that might undermine their ability to act responsibly, in order to increase their ability to do so" (2010, 652-3). Similarly, Strawson's optimist argues that a deterministic conception of human conduct is perfectly compatible with our practices of holding people accountable for their actions; of praising and blaming them for what they do. My beliefs, desires, and intentions may be fully determined by prior conditions or external constraints that preclude the possibility for me to act from a place of unconstrained freedom, but as long as my conduct was caused by *my* desires, and not by means of coercion or compulsion, we have a sufficient ground for justifying our responsibility-practices. Specifically, the optimist's idea is that holding people responsible, praising and blaming them for their voluntary non-

coerced actions, can shape their motivational states in highly desirable ways. This ‘forward looking’ approach, with its emphasis on the after-the-fact influence of incentives like praise and blame on an agent’s motivational states, justifies responsibility practices for being useful tools in regulating, manipulating or training human behavior. The pessimist objects, and Strawson agrees, that this emphasis on behavior-manipulation cannot capture our sense that our responsibility-practices track moral desert and that they “are expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes. Our practices do not merely exploit our natures, *they express them*” (Strawson, 36)

²⁰ Of course a fair range of reactive attitudes will still be wholly appropriate for persons who warrant agency-excusing conditions. Our discussion so far has precisely emphasized the epistemological and ethical priority of the second-person engaged attitude with its ability to contribute to the flourishing of persons like Sessa and Audrey. At first glance, Strawson’s way of juxtaposing the objective and the reactive attitude may seem to entail that people like Sessa and Audrey, who are certainly not appropriate targets of the kind of resentment characteristic of typical adult human relationships, are rightfully viewed through the lens of the objective attitude. Strawson touches on this issue when he acknowledges that “the simple opposition of objective attitudes on the one hand and the various contrasted attitudes which I have opposed to them must seem as grossly crude as it is central.” Nuancing his view, he adds, “parents and others concerned with the care and upbringing of young children cannot have to their charges either kind of attitude in a pure or unqualified form. They are dealing with creatures who are potentially and increasingly capable of holding, and being objects of, *the full range of human and moral attitudes*, but are not yet truly capable of either. The treatment of such creatures must therefore represent a kind of compromise, constantly shifting in one direction, between objectivity of attitude and developed human attitudes” (2008, 32, my italics). The point, then, is not that Audrey and Sessa warrant a wholly objective stance *as persons*, but that a certain way of understanding their behavior, namely as behavior for which they are responsible, loses traction. Our engagements with Audrey and Sessa and “the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships ... cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other” (Strawson 2008, 25). And to have one’s reactive attitudes curtailed in this way, Strawson wagers, *just is* to place the agent in question outside the community of responsible agents.

²¹ Alva Noë argues for a similar point in *Out of Our Heads* (2010) when he writes: “Our commitment to other minds is ... not really a theoretical commitment at all. We don’t come to learn that others think and feel as we do, in the way that we come to learn, say, that you can’t trust advertising. Our commitment to the consciousness of others is ... a presupposition of the life we lead together. I cannot both trust and love you and also wonder whether in fact you are alive with thought and feeling.” (33)

²² Of course a committed Strawsonian pessimist could respond (and perhaps, although this strikes me as unlikely, Colombetti and Torrance would want to respond) that the natural or psychological impossibility for us to give up the reactive attitudes does not rationally justify them and that “the real question is not a question about what we actually do, or why we do it. It is not even a question about what we would *in fact* do if a certain theoretical conviction gained general acceptance. It is a question about what it would be *rational* to do if determinism were true, a question about the rational justification of ordinary inter-personal attitudes in general” (2008, 27). However, as Strawson points out, the pessimist’s concern with rational justification is located at the wrong level. *If* it turned out we had a choice to suspend our reactive attitudes and the inter-personal relations negotiated through them on the basis of a rational theoretical insight, Strawson argues, this choice would derive its rationality from “an assessment of the gains and losses to human life, its enrichment or impoverishment; and the truth or falsity of a general thesis of determinism would not bear on the rationality of *this* choice” (2008, 28).

²³ Thanks to Alice Crary for urging me to make this more explicit.