Electronic Theses and Dissertations

2021

Difference and Demand: Toward a Levinasian Psychopathology

Edward Anthony Lenzo

DIFFERENCE AND DEMAND:

TOWARD A LEVINASIAN PSYCHOPATHOLOGY

by

Edward Lenzo

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Philosophy

The University of Memphis

May 2021

For Others

Acknowledgments

This work would not have been possible if not for a relation with a plurality of others: mentors, colleagues, friends, and family.

I want to begin by thanking my committee. Shaun Gallagher's work has inspired me since I encountered it as an M.A. student at Colorado State University; it has been a privilege to conduct my work under his tutelage. Tom Nenon, who has an incredible skill for making the abstruse intuitable, has been instrumental in my navigation of Husserl's voluminous work. It was Kas Saghafi who introduced me to the thought of Emmanuel Levinas, and I am grateful to him for calling the totality of my prior understanding into question. Without Joel Krueger's assistance, my explication of Levinas' thought and its application to issues in contemporary phenomenological psychopathology would have been unacceptably unclear.

I thank the faculty, staff, and graduate student body of the Department of Philosophy at the University of Memphis. The Department has cultivated a rigorously intellectual atmosphere that simultaneously does away with much of the competitiveness and aggression too often present in academia. Instead, the Department has nurtured a friendly, collegial environment where criticism is always in service of responsible, original thought. I owe a special debt of gratitude to Cathy Wilhelm and Connie Diffee: they have always treated me with love and acceptance, and without their expertise and hard work the Department simply would not function. I also thank the Department of Sociology, Philosophy and Anthropology at the University of Exeter, and in particular Joel Krueger and Giovanna Colombetti, who welcomed me for a semester of research abroad and provided a number of opportunities to share my work and to learn from others.

I thank my dear friends, who have also invariably been mentors to me. I strive to emulate Philip Turetzky, who is to me an intellectual exemplar; the breadth and depth of his

Abstract

  This dissertation is an attempt to approach psychopathology from a broadly Levinasian perspective. I do so by situating concerns and insights raised by Emmanuel Levinas relative to phenomenological and "4e" (embodied, embedded, extended, enactive) approaches to psychology, cognitive science, and psychopathology. There has been some recent take-up of Levinas by these approaches, but it has remained somewhat superficial or otherwise inaccurate. It is my position that Levinas' concepts of expression as a fundamentally social, ethical, and asymmetric relation, embodiment as sensitive exposure to and enjoyment of the world, and subjectivity as the substitution of one's own existence for the other, can deepen our understanding of intersubjectivity, ethical experience, and pathology phenomena.

  In *Chapter 1*, I demonstrate that phenomenology relies on similarity in its articulation of interpersonal understanding, and that this becomes problematic in the context of psychopathology, where difference becomes thematic. In *Chapter 2*, I articulate Levinas' conception of expression, which presupposes difference rather than similarity and accomplishes an interpersonal relation that is fundamentally ethical. In *Chapter 3*, I show how such a conception can contribute to understanding by considering the ways that expression can direct one's attention to that which was not previously known. I do so by considering the theory of natural pedagogy and dynamical systems approaches in psychopathology. In *Chapter 4*, I present some underlying assumptions of 4e approaches to psychopathology, specifically focusing on enactivism, and consider enactivist accounts of ethics and ethical experience. It is my contention that these accounts fail to make sense of the possibility of overcoming self-interest, and accordingly of various forms of ethical experience. However, bolstering these accounts with a Levinasian concept of expression produces a powerful understanding of the production of understanding across difference. In *Chapter 5*, I apply Levinasian insights to interpreting specific

pathology phenomena associated with depression, narcissism, and schizophrenia. Finally, I

address the applicability of concerns raised in *Chapter 1* to my own account.

Table of Contents

Introduction

  This dissertation is an attempt to approach psychopathology from the perspective of the ethical metaphysics of Emmanuel Levinas. My goal is to put Levinasian concepts into communication with problems and insights in psychopathology, particularly from the perspectives of phenomenology and enactivism. My thesis is that these two areas of work can contribute to one another. Specifically, Levinas' concept of expression can help us to better understand intersubjectivity and provides a novel approach to the problem of incomprehensibility and the nature of so-called "self-disturbances" in psychopathology; on the other hand, enactivism provides us with an account of how it is that we arrive at mutual sense, particularly across seemingly radical difference. Phenomenological and enactivist psychopathology heavily draws on classical phenomenologists such as Husserl, Heidegger, and Merleau-Ponty. It is only recently that some have begun to incorporate certain of Levinas' insights, but this remains piecemeal and somewhat misrepresentative of the fundamental dimensions of his work.

  I stress the importance of the ethical dimension of human intersubjectivity for making sense of other persons (and thus for any theory of interpersonal sense-making, too), and especially the asymmetry of the ethical relation. I support these Levinasian insights by appealing to (and reinterpreting) a naturalistic theory of learning (the theory of natural pedagogy), and go on to apply Levinas to the approaches in philosophy of cognitive science and psychopathology known as interaction theory and participatory sense-making. I then demonstrate certain limitations in current enactivist accounts of ethics and offer Levinasian correctives. Finally, I draw out the implications of Levinas' concepts of embodiment, sensibility, and expression for

our understanding of the self in relation to self-disturbances in pathology phenomena, e.g., schizophrenia.

Psychopathology is a field that attempts to understand the other person across differences that have, at times, seemed incomprehensible. Phenomenology strives to make positive sense of that which it studies, that is, it aims to say what some given phenomenon *is*, not merely what it is not in relation to a network of other, better known phenomena. Phenomenological psychopathology, then, attempts to make positive sense of the other across difference. Emmanuel Levinas, who embraced the spirit if not the letter of Husserlian phenomenology (Derrida, 1978, p. 107; Levinas, 1981, p. 183) took more seriously the question of the subject's relation with an irreducible other than do the phenomenologists who act as touchstones in psychopathology.

Beginning with Husserl, phenomenology is first an egology, an endeavor on the part of the philosopher who, reflecting on their own experience, purports to discover general structures of consciousness. On such an approach, intersubjectivity is the question of the presence of others within experience, i.e., the appearance of others to or for consciousness. Enactivism, inspired by a diverse host of philosophical and scientific thought (including the phenomenology of Husserl, but also that of Heidegger, Merleau-Ponty, and other phenomenologists, as well as biology, neuroscience, dynamical systems theory, and cognitive science), first gained attention in conjunction with the theory of autopoiesis, according to which all organisms sensitively navigate their environments in terms of what matters for sustaining their own existence. All organismic life is therefore cognitive, in the sense of *sense-making*, and also affective, since sense-making refers to the way the environment affects the organism in terms of organismic values. For the enactivists, sense-making begins with the organism, though enactivism has now turned to seriously consider interactions between organisms.

There are three major questions of intersubjectivity within the domains specified above. First, how is it that I know or understand the other person as a person? Put in phenomenological terminology, how does the other person appear to me in experience as precisely an other person, distinct from myself and my own consciousness? Second, how is it that I know or understand the particular states of the other person, that is, how do I know or understand that the other person is undergoing a certain experience, e.g., happiness, anger *about* X, and so on? Third, how do relationships with others condition or structure my cognitive-affective processes in general? That is, how is my understanding of not only the other but also of myself, the world, and worldly things conditioned by others?

These questions have primarily been considered in terms of constitution and epistemology, appearance and understanding. Phenomenology and enactivism also remind us – as do other traditions within cognitive science, at times – that such things extend beyond categories of mind traditionally conceived as cognitive, e.g., propositional thought, reasoning, categorical identification of phenomena, to also include affective forms of sense-making: we understand the world, in part, by how the world affects us in terms of emotions, moods, feelings, and even organismic desires and needs. We understand in terms of what matters to us. It is my claim that the answers to all three of these questions depend not only on appearance and understanding – not just on my own cognitive/affective activity – but also on an ethical relationship with others, making front and center not what matters to me but rather what matters to others, even despite my own activity and self-interest.

It is not as if phenomenology and enactivism have not broached the topic of ethics. Indeed, ethics was a major concern for Husserl and other phenomenologists (De Warren, 2017; Donohoe, 2004). Though, for Husserl, ethics was seen as emerging from the general principles of

knowledge and rationality, and referred primarily to norms and values. Enactivism, too, has shared this emphasis. The bulk of enactivist thought on the topic of ethics has been contributing to an understanding of the generation of values and norms, not only for an individual but also at a societal level, and the ways that such things feedback on the cognitive/affective processes out of which they arise. (Participatory sense-making, developed by De Jaegher & Di Paolo, 2007, is a good example of this, as are Colombetti & Torrance, 2009, and Varela, 1999. They all show how normativity arises through social relations (or the stronger claim, that meaning itself is normative), and correspondingly constrain the further production of meaning.)

Levinas, however, has a different idea of ethics. Ethics describes a primary relationship with another person characterized by an asymmetric responsibility. It is primary because it describes the social relation *par excellence*; it describes contact with another person even before ever coming to understand her. The other person, simply by expressing herself to me, makes a demand that I always must respond to in some way or other. This demand is a responsibility independent of whether the other will reciprocate any response I issue. Levinas goes as far as to argue that we are responsible even for the other's responsibility. This extreme, asymmetric responsibility is at the core of ethics for Levinas, and is prior to and, more importantly, *better than* any value or norm that I myself participate in; not only is it a source of values and norms, but it is only through an asymmetric responsibility that I can put my own concerns aside and even act *against* norms (which, it must be noted, can themselves produce and sustain violence) *for* the other.

Not only does Levinas' formulation of ethics help us to understand ethical experience in general, and especially phenomena such as extreme altruism, but it also helps us to understand *understanding* others, in the more common cognitive/affective sense. It is due to the other

mattering to us independently of ourselves that we are forced to reckon the world and bestow upon our perceptions an intersubjective signification. Indeed, it is the other who first introduces the possibility of a significance beyond myself. Although this is somewhat in line with other major phenomenological thinkers, Levinas' insight is that this process already bears an ethical significance and furthermore that responsibility for the other is the ultimate justification of any objective claims; the other can always disrupt my understandings and call me to answer for them.

A Levinasian approach finds as an ally the theory of natural pedagogy, according to which we learn general kinds by being addressed by others (Gergely & Csibra, 2009). The intersubjective relationship, then, is not only an ethical relationship but, due to the exigence of that relationship and its motivation of significance, it is also an asymmetric teacher/student relationship: I learn from the other as she directs my attention to the world. This phenomenon is well supported by the philosophical-scientific literature on infant learning and expression. Of especial interest are the concepts of primary and secondary intersubjectivity, which I reinterpret in making the argument for the primacy of ethics.

This account of asymmetric responsibility co-implicates a novel theory of embodiment. Within phenomenological and enactive approaches to cognition and psychopathology, the body is primarily understood by reference to certain couples of opposing terms, e.g., object body/lived body, body image/body schema. Irreducible to these formulations is the body as *sensitive vulnerability*, as the very condition of our being, beyond but conditioning any thought we may have of the body or anything else. It is as a sensitive being, vulnerable to the world writ large, that I can first come to encounter anything at all beyond myself, including the other who approaches me as a body. In responding to the call of the other, I come to account for myself – account for my bodily actions and freedom, and reckon my own place in the world relative to the

5

other; this accounting is the formation of a further ethical, thematic self, in addition to the basic embodied self – or what I call an embodied *ego* – already presupposed. For this reason, I use Levinas to better understand the nature of self and self-disruption in psychopathology phenomena, including depression, narcissism, and schizophrenia.

My reader must keep in mind a number of themes that, while not always explicit, undergird this work. The first is the relationship between comprehension and understanding, on the one hand, and incomprehension or "ununderstandability" on the other. Incomprehensibility has long been considered a problem in psychopathology, since at least as early as Karl Jaspers, but tracing back its roots to antiquity. In this dissertation, the project is not to deny incomprehensibility nor to overcome it by some special method; rather, my claim is that understanding *always* emerges out of prior incomprehension: first we communicate, then we understand (the other, the world, ourselves). The question is of the relationship between incomprehension and comprehension, thus the special interest taken in learning and sense-making.

A second theme to keep in mind is the relationship between theory and practice in psychology broadly conceived (including, e.g., psychiatry and psychopathology; at times I use the word perhaps *too* broadly to also include cognitive science). I suspect that theory and practice cannot be sharply distinguished, though of course that does not mean neither pole can be emphasized for various purposes. Some psychiatrists, such as Giovanni Stanghellini (2019), argue that psychological practice must be based on theory (though, this is fortunately not a conclusion he seems to adhere to at all times). On the other hand, a Levinasian position that roots itself in a face-to-face encounter with another person must emphasize the role that practice plays in the very formulation – and accountability – of theory. Here, the most important thing to bear

6

in mind is that my claims throughout ought to always have some import for at least one but, often, both sides of this dichotomy.

A third is the question of naturalism. In this dissertation, I ultimately provide certain naturalistic defenses and implications of Levinas' thought. Yet, Levinas is often seen as resistant to naturalization. Within psychopathology research today, the project of "naturalizing phenomenology" looms large, and one should wonder, then, how Levinas fits into this project. Though I do not treat the question explicitly in the body of this work, I can share some initial thoughts on the matter. First, there are many ideas of what constitutes a "naturalistic" philosophy. Quite broadly, I take naturalism to refer to an understanding of some phenomenon, X, in terms of a system of concepts that ultimately refer to 'natural' phenomena as determined by natural science. For instance, a naturalistic account of desire would ultimately found the reality of that state on whatever things or forces our physicists (and perhaps chemists and psychologists) claim exist. Second, bearing in mind this sense of naturalism, Levinas bears an interesting relationship to the project of naturalism. As readers will see in *Chapter 2* below, Levinas' ideas of embodiment, enjoyment, and even hedonism seem amenable to such a project. On the other hand, Levinas explicitly maintains that one's relationship to an other cannot be totalized under any system of concepts, and thus it would follow that it cannot be totalized under a system of natural concepts. In a perhaps deflationary reading, this is no more radical than maintaining that ethics cannot be reduced to physics or moral psychology: it is just not a natural scientific domain. On the other hand, it is because ethics *demands* justice that we are left to reckon with a perhaps *uncomfortable naturalism*. Because we always relate not only to one other but to a plurality of others, Levinas maintains that we must "compare incomparables" (Levinas, 1981, p. 158), that is, divide our attention and resources among plural others. Therefore, ethics demands

exactly what it prohibits: the application of general concepts to the other. Insofar as "the Other's material needs are my spiritual needs" (Cohen, 1986, p. 24), I must apply material, i.e., natural, concepts to understanding these needs and for executing a response. But this is an uncomfortable naturalism since in doing justice to plurality through the application of general concepts I unavoidably do violence to the non-totalizable Other.

Before moving to chapter outlines, I want to make a terminological specification already suggested above. Herein, I distinguish between morality and ethics. Morality refers to values and principles for orienting action according to those values. That is, morality deals with *morals* construed as specific principles, lessons, or standards. To be moral refers to one's adherence to morals. Ethics deals with the source of meaning of values. Here, since Levinas is the major figure, ethics refers to a responsible relationship with others, out of which we (ought) derive morality. Enactivist ethics has hitherto, accordingly, been primarily an enactivist account of *morality*: an account of the processes of production of specific values, principles, and norms. This distinction is not meant in any disparaging way – morality is wed to ethics, and both are utmost concerns; rather, the distinction is meant only for conceptual clarity.

In *Chapter 1*, I focus on the foundational text of phenomenological psychopathology, *General Psychopathology* by Karl Jaspers, and the take-up of that project by transcendental phenomenology, with an emphasis on Edmund Husserl, Edith Stein, and to a lesser extent Martin Heidegger. I demonstrate that Jaspers' concept of empathy more closely resembles that of Husserl and Stein than it does that of his contemporary, Theodor Lipps, which has become central to contemporary simulation-theory accounts of interpersonal understanding. Empathy, for Jaspers, Husserl, and Stein, is the primary way that we understand others, and depends on perceiving bodily movements that resemble my own, as well as familiarity with the kinds of

experience that, in my own case, accompany such movements. The movements of the other, just as our own, take on significance as *expression*, i.e., psycho-physical unities of those movements and associated mental acts. Where empathy fails, according to Jaspers, understanding becomes impossible. In the case of certain pathology phenomena, such as thought insertion, this renders the target phenomenon incomprehensible. In such cases, the only recourse is to a deficit method, by which we understand only what the phenomenon *is not*. I critique the deficit method on epistemological and ethical grounds, as well as the transcendental phenomenological modification thereof. In the latter case, a special problem is introduced that continues to shape the rest of the dissertation in the background: the discovery of essential structures of experience that phenomenology claims to accomplish depends on certain interpersonal relations, and such relations cannot be presupposed to hold in relevant psychopathology cases.

In *Chapter 2*, I explicate Levinas' concept of expression. Rather than psycho-physical unity, expression is primarily conceived by Levinas as communication: the other expresses herself insofar as she communicates with me in such a way as to demand a response. In demanding a response, the expression of the other is inherently disruptive of consciousness, rather than being understood as an act constituted for consciousness. In order to explicate expression, I begin with Levinas' understanding of embodiment as vulnerable sensibility, which becomes important for subsequent chapters, especially *Chapter 5*. The crucial point is that our relation to the other, according to Levinas, is not a perceptual but an ethical one. It is this ethical relation that contemporary phenomenological and enactivist approaches to intersubjectivity have yet to seriously address (though there have been a number of accounts of what I call *morality*). Formulating the interpersonal relationship in terms of ethics also allows us to approach the incomprehensibility problem from a new direction: on this account, all understanding is founded

on incomprehension; we first communicate, without presupposing any understanding whatsoever, and it is only in responding to the call of the other that we come to understand. *Chapters 3* and *4* explore the ways that we arrive at understanding from incomprehension.

In *Chapter 3*, I begin by revisiting Jaspers' incomprehensibility thesis and exploring its implications for a Levinasian approach. Importantly, a number of thinkers maintain that such an approach would be "without method," "unsettled," or would otherwise endorse some version of the incomprehensibility thesis. This need not be the case. I go on to focus on the theory of learning implicated by Levinas' position, according to which learning begins with the other directing our attention. This is complemented by the theory of natural pedagogy, which relies on address as motivating general learning, as well as dynamical systems theory (which becomes especially interesting in the domain of psychopathology). I then return to the incomprehensibility thesis and explicate the forms of understanding available on a Levinasian approach, as well as their implications for clinical and theoretical psychopathology, i.e.., that theory is founded on and accountable to practice, that interpersonal understanding is particular and limited, that incomprehensibility in the problematic sense results from unjustified generalization, and that attentional modulation and affect can fruitfully guide research (à la "praecox feeling").

In *Chapter 4*, I turn to consider interaction theory and participatory sense-making, two accounts of intersubjectivity within enactivism that explore the idea of social cognition. According to interaction theory, social understanding often just *is* social interaction, and according to participatory sense-making, we make sense of the world – and ourselves and each other – together. There are undoubtedly true insights here. I argue, however, that these accounts overlook asymmetry, and in doing so are unable to adequately account for understanding contrary to common sense and self-interest. This discussion leads into a consideration of

enactivist ethics, represented here by the work of Francisco Varela, Giovanna Colombetti and Steve Torrance. Levinas can help us to make adequate sense of ethical experience that enactivist ethics has yet to account for. On the other hand, enactivist ethics and accounts of social cognition can lend to Levinas a compelling story for transitioning from ethical asymmetry to *doxic symmetry*, i.e., the social establishment of mutual sense. I thus see the Levinasian account as complementary with enactivism, at least in this regard.

In *Chapter 5*, I consider the implications of Levinas' conception of embodiment (introduced in *Chapter 2*) for psychopathology. In recent years, many pathology phenomena have been characterized in terms of self-disturbance, e.g., schizophrenia has been conceived of in terms of a loss of one's sense of agency or ownership, or an unstable first-personal perspective, or a failure to differentiate oneself from others. I apply Levinas here to distinguish between an embodied, sensitive ego and a social, responsive I, or personal self. This allows me to consider a number of possibilities regarding self-disturbance, linking each to some form of pathology or other, but concluding that responsibility for the needs and responsibility of others requires that we rethink the scope of psychiatry beyond the limits of either the self or dyadic relationships. On the way, I address two major concerns, which originate in my treatment of phenomenology in the first chapter: (1) that a Levinasian position collapses into a different form of the deficit model (represented here by the Seattle School of psychology), and (2) that the approach here commits the same kind of problematic structuralism levelled at phenomenology.

Chapter 1

Phenomenological Psychopathology:
Familiarity and Alterity in Jaspers, Husserl, and Contemporary Thought

Introduction

Phenomenological psychopathology is largely considered to have originated as a field with the publication of Karl Jaspers' *General Psychopathology*, though today it is just as influenced by classical phenomenologists such as Edmund Husserl and Martin Heidegger. In this chapter, I begin with Jaspers' *General Psychopathology*, explicating the conception of empathy at play in that work. This is the crucial concept because it is empathy that Jaspers identifies as understanding "in the true psychological sense" (Jaspers, 1997, p. 304). An act of empathy, or psychological understanding, requires two conditions: perception of bodily movement and personal familiarity with the associated mental experience. I demonstrate that this conception of empathy is congruent with Husserl's own and not, as is sometimes presumed, with the concept as formulated by the simulation theory of social cognition. Where empathy fails, Jaspers maintains that our only recourse is to a deficit method, according to which we articulate only what some (pathology) phenomenon *is not*. I argue against the deficit method on epistemological and ethical grounds. Next, I show how transcendental phenomenology can move past a Jasperian, or "strict," deficit method, in virtue of positing universal structures of consciousness. This, however, is also problematic: the applicability of such structures to a person other than oneself presupposes certain things – broadly, resemblance of essential possibilities and inclusion in communal practices – that simply cannot be presupposed (and are in fact often lacking or actually undermined) in the relevant cases of pathology phenomena. In summary, psychopathology thematizes divergences from ordinary consciousness in such a way as to undermine the applicability of these methods.

Jaspers' General Psychopathology

In his foundational *General Psychopathology*, Karl Jaspers attempted not only to collect and analyze some of the best work in the field at the time, but also to establish the fundamental concepts of psychopathology and to systematize its methods. With this work, he is often credited with founding the field. He conceives of 'empathy' as central to genuine psychological understanding and, accordingly, to psychopathology. Empathy is a specific form of perception that is both similar to but importantly distinct from ordinary forms of (object-) perception. The details of this mode of perception are, however, underdetermined by the text. I here explicate what can be determined from the text, then employ two other, more developed theories of empathy – those of Theodor Lipps, on the one hand, and phenomenologists Edmund Husserl and Edith Stein, on the other – that continue to be influential from Jaspers time to our own, in order to disambiguate the concept and flesh out its remaining details

Empathy

Jaspers distinguishes between (i) subjective experience, (ii) objective performances, (iii) somatic accompaniments and (iv) meaningful objective phenomena. Respectively, these refer to (i) first-personal experience, to be described by phenomenology, (ii) psychic performances, which are conceived roughly as mental or intentional acts,[1] such as apperception or memory, (iii) bodily events that are in themselves psychologically meaningless and merely accompany psychic performances and (iv) expressive behavior, i.e., actions that may only be understood through considering the agent's creative productions and relation to a meaningful world. These

---

[1] According to Husserl, Franz Brentano discovered the intentional structure of the "mental," though he did not fully realize the significance of this discovery nor develop a systematic method of intentional analysis. Jaspers, however, was familiar with Husserl's (early) work, and can be understood as more explicitly appreciating these points. Jaspers did not, however, typically engage in the intentional method all the way down to explicitly uncovering the modes of the appearance of the other for the observing subject, e.g., the psychologist. That is, he was interested primarily in the constitution of noema – what characterizes the observed phenomena in question *as object* – and did not typically push the question of noesis.

phenomena are interrelated, and in order for understanding to be truly psychological (according to Jaspers), it cannot satisfy itself with mere description of individual states or causal explanations of mental and physical acts, but must grasp meaningful phenomena. The way in which a psychologist can grasp meaningful phenomena requires an act of empathy: "There has to be an act of empathy, of understanding" (Jaspers, 1997, p. 55).

Jaspers does not explicitly define empathy. However, there are some consistent uses and determinations throughout the text. As one of the earliest mentions of empathy within the work, he writes "where we understand how certain thoughts rise from moods, wishes, or fears, we are understanding the connections in the true psychological sense, that is by empathy (we understand the speaker)" (Jaspers, 1997, p. 304). Beginning here, we can elucidate at least three aspects of empathy that remain consistent throughout the text. First, empathy is, for Jaspers, psychological understanding, i.e., understanding "in the true psychological sense."[2] Empathy is consistently taken to be identical to or necessary for psychological understanding throughout the text (as can be seen, for example, at Jaspers, 1997, p. 311, quoted below). This idea of necessity is consistent with a number of theories, prevalent at Jaspers' time as well as (in modified forms) our own, that designate the fundamental act by which persons come to understand one another in a broad sense as "empathy" (such as those theories of Lipps and Husserls, considered below).

Second, empathy involves an understanding not of psychological states or processes, i.e., "thoughts," considered in isolation, but rather in the context of other states, such as "moods, wishes, or fears." Jaspers is interested, in part, in establishing the fundamental concepts of psychology and psychopathology, and this involves a determination of the being of the phenomena in question. For Jaspers, this determination is thoroughly phenomenological, in the

---

[2] He continues: "Empathic understanding […] always leads directly into the psychic connection itself" (Jaspers, 1997, p. 304).

sense that it must accurately describe the phenomena *as lived by* the subjects that undergo them. Drawing on Wilhelm Dilthey as well as Husserl, Jaspers takes the experiential life of a given person to be interpretable only as a whole; an individual experience can be adequately understood only in its interconnections with others. Thus, though a description of particular kinds of "thoughts" is desired, such a description can only be adequate to the subject's lived experience if it factors in the context out of which those thoughts arise. The closely related third point is that empathy grounds a form of understanding that is directed to understanding *the person*, rather than some abstract psychological entity divorced not only from its psychological context but also from the context of a specific, individual, actual human personality and life.

Reinforcing these points, and allowing us to determine the conditions that make empathy possible, Jaspers writes: "we can have no psychological understanding without empathy into the content (symbols, forms, images, ideas) and without seeing the expression and sharing the experienced phenomena" (Jaspers, 1997, p. 311). Empathy thus delivers insight into specific content, which Jaspers takes to include symbols, forms, images and ideas. As a grasp of a psychological context, empathy necessarily involves a certain structural awareness of the experiential life of the other, but this is inseparable from an understanding of the specific contents that concretely realize that structure. In our empathetic relationship to an individual person, we understand how, for example, a specific anger arises from a specific fear, how anger arises from fear *for that person*, etc. (Jaspers, 1997, p. 303). These kinds of understanding form the basis of psychological theory, which then strives to abstract from any individual and understand relationships between psychological kinds in general. This inseparability between mental form and content – that every psychological act has its content and vice versa, and that adequate description must consider them both in relation to one another as well as to other acts

and content – is a product of Jaspers' take-up of the fundamental insights of Dilthey and

Husserl's development thereof.

Furthermore, expression and the sharing of experienced phenomena (or simply,

experience-sharing) are conditions of empathy. According to Jaspers, "we never can perceive the

psychic experiences of others in any direct fashion, as with physical phenomena […] we can

only make some kind of representation of them" (Jaspers, 1997, p. 55).[3] What is given directly in

experience of the other is never their mental life, but only their expressive behavior. Ideally,

expression is given "first-hand" in a personal conversation or perceptual observation of the

person in question, though written word is also a form of expression and thus can also ground

psychological understanding. Expression is not mere behavior, i.e., movement, however; the

perception of bodily movements considered in isolation would be insufficient for empathy, as

these movements could not of themselves reveal the psychic context that empathetic

understanding requires. Following Dilthey, Jaspers conceives of expression as a psychophysical

unity. Behavior is grasped directly by perception, while the psychical correlate thereof is

understood by a kind of reference to ourselves. He writes:

> Psychology traditionally takes on the task of bringing into consciousness material of which
>
> we are unaware. Evidence for such insight has always rested on the fact that, circumstances
>
> being favourable, other people could notice the same things, provided they had undergone the
>
> same experiences. (Jaspers, 1997, p. 306)

---

[3] It might be objected here that empathy, as "representational," is accordingly not a kind of perception. I believe that we are better off interpreting it as perception, and the reasons why will become clear when we consider Husserl and Stein's account of perception. To briefly rebut the objection, it is sufficient to note that perception *always* involves something not given but rather "represented." There remains an important difference between empathy and other forms of perception, however, in that the psychological aspects of the other person could never be given directly ("primordially," in Husserl's terms), whereas the absent but represented sides of spatio-temporal objects, in principle at least, always could.

Having had similar experiences as the person to be understood, a fact presumably to be verified through self-description and analogy to one's own expressive behavior, allows one to contextualize the behavior directly perceived and thereby indirectly understand the psychical experience of which the behavior is an expression.[4] The sharing of experience between the person to be understood and the would-be empathizer is thus constitutive of the meaningfulness of expression – it is that through which behavior can be psychologically understood – and thus experience-sharing limits psychological understanding. "We do not understand a person merely by isolating out a number of such contents, but the degree to which the psychologist is at home with them will limit and condition his psychological understanding" (Jaspers, 1997, p. 308). Since experience-sharing limits empathy, Jaspers also takes culture to be relevant to psychological understanding (Jaspers, 1997, p. 307). Those familiar experiences with which one is "at home" depend, to considerable degree, on one's culture.

Some examples. First, consider a case in which someone's loved one has passed away. You see them talking, smiling, and joking around as usual – as if nothing were wrong. Had you yourself never lost a loved one, or perhaps even had you never lost a loved one *and also behaved similarly* then, you may not understand at all why this person seems so content, if not outright happy. You might assume that they didn't much care for the deceased, or perhaps are in complete denial. Yet, had you undergone a similar experience, or had someone else explained it to you, or perhaps if you simply know this person very well (e.g., you know their priorities, how they've dealt with other hardships, their general outlook, etc.), you might realize that this behavior is a strategy for putting their friends at ease and for trying to counteract their own grief. Second, consider the context of an academic conference. Someone outside of academia is

---

[4] Husserl elaborates the indication of this similarity or sharing, as well as the mode of its possible verification.

unlikely to view a raised arm with a single finger extended as a 'follow-up' question or comment in distinction from a raised hand which, in that context, signifies a new question or comment. In general, outside of a context in which a particular movement is accepted as an '*expression of …*,' that same movement can take on an utterly distinct meaning or, perhaps, even none at all. These examples illustrate, I hope, that the theory of expression with which Jaspers is operating is a broadly structuralist one, in the sense that the meaningfulness of the behavior in question depends on that behavior's relations to other behaviors and psychical states, both belonging to the subject in question as well as those that the interpreter (e.g., the psychologist) has experienced for themselves.

To recap: For Jaspers, psychological understanding is grounded on empathy, which in turn depends on the perception of expressive behavior and the sharing of experience. Like object-perception, it depends on certain qualities being given to sensation, i.e., the physical aspect of the expression considered as bodily movements. Unlike object-perception, however, aspects of what empathy discloses are in principle un-perceivable and must instead be "represented." That is, I can never *see* the psychological experience of the other person, though I could in principle see, for example, the backside of the mug I am holding, if only I were to turn it around. It is in this way that my experience of myself and my experience of others is fundamentally distinct, despite only being able to make sense of the latter by reference to the former.

Psychopathology

According to Jaspers, the benefit of psychological understanding for psychopathology is to bring certain experiences to awareness, that is, to allow the psychopathologist to represent relevant psychological phenomena through empathy. In many cases, he claims, we can

understand psychopathology phenomena as increases or diminutions of non-pathological phenomena (such as can be represented by empathy), or as otherwise normal experiences brought about by abnormal causes (theorized by empirical psychology) [or motivational relations.] In the former case, we can have a legitimate form of understanding since the difference between the experience of the other and my own is only a matter of degree. In the latter, we couple understanding with explanation: I can understand the kind of state that the other is in, and I can understand the context out of which it emerges, but I must explain the connection between these in terms of an aberrant causality; the phenomena could be described as a familiar state arising from unfamiliar causes, or familiar causes resulting in a different state – though still a familiar one – than what we would ordinarily predict.

In certain cases, however, understanding is not possible at all (though causal explanation might still be applicable). For example, in the case of "made phenomena," present for some persons diagnosed with schizophrenia, e.g., thought insertion and delusions of bodily control, the relevant psychological phenomena are "beyond empathy" and are *in principle ununderstandable*. He writes:

> The most profound distinction in psychic life seems to be that between what is meaningful and *allows empathy* and what in its particular way is *ununderstandable*, 'mad' in the literal sense, schizophrenic psychic life. (Jaspers, 1997, p. 577)

In the latter case, "we find changes of the most general kind for which we have no empathy but which in some way we try to make comprehensible from an external point of view" (Jaspers, 1920, p. 88; translation mine).

If empathy is the foundation of psychological understanding, then we can understand this ununderstandability, or incomprehensibility, in terms of the failure of empathy. Recalling that

19

empathy involves a grasp of the psychological context out of which the target phenomenon arises, Christoph Hoerl (2013) argues that empathy is undermined in the case of "made phenomena" because such phenomena do not follow an established psycho-generative structure: they appear as if from nowhere (to borrow an expression from Martin and Pacherie's "Out of Nowhere" (2013)). Since this structure is part of the proper object of empathy (precisely: grasping this structure is a part of grasping the target phenomenon, which is incomprehensible considered in isolation), if it is not possible to grasp this structure, then the target phenomenon likewise cannot be understood. It is unclear to me, however, whether it is truly the case that made phenomena cannot be understood in terms of their generation and placement within a psychological context. Indeed, many delusions and hallucinations build on the contents of experience and develop over time, and there is substantial literature on the ways that schizophrenic experience, for instance, emerges out of certain experiences, such as childhood trauma, the loss of a loved one, severe depression or anxiety, substance abuse, difficultly adjusting to a foreign culture, and so on, and the ways that the sense of some pathological state develops from meaningful experiences of the subject generally. Rather than creation ex nihilo, such phenomena only appear *as if* ex nihilo, in the sense that they often, at least initially, appear without expectation and, according to popular approaches, with either an altered or absent sense of ownership or agency. They can be understood in context, just perhaps not at first, or not in terms of an experientially unified subject of experience. This latter feature has earned schizophrenia its characterization as "dissociative."

But there are other reasons why such phenomena may be, in principle, incomprehensible on a Jasperian approach. Jaspers claims that there are "phenomenological elements" that are inaccessible to the non-pathological subject, or even to anyone but the individual who undergoes

such experiences. He goes as far as to claim, plausibly in certain respects, that schizophrenic persons typically do not understand the specific character of one another's experiences any better than do non-pathological observers, though there can be a kind of co-development of symptomology such that contents of experience adhere to a general structure; for example, he discusses the case of a family who co-create a common narrative through which to interpret their delusions, and he suggests even that one's situation within narrative practice can even induce certain forms of schizophrenia like symptoms (e.g., the parents in his example seem to suffer from schizophrenia, and their narrative practices of sense-making, report, and explanation shape the way that their children perceive the world, such that they would not receive a diagnosis of schizophrenia on their own or even if simply removed from the narrative context, but nevertheless experience a kind of pseudo-schizophrenia due to their environment). In such cases, involving fundamentally inaccessible phenomenological elements of the other's experience, we can interpret incomprehensibility as the lack of the experience-sharing necessary for an act of empathy; the observer simply does not know what the other is going through, since they have not experienced the relevant kind of phenomena themselves.

Furthermore, empathy requires the perception of expressive behavior, and behavior counts as expressive when it resembles my own behavior, that is, when it is behavior that I understand as expressive of certain experiential facts. If the experiential fact is unfamiliar to me, that is, if it is an inaccessible "phenomenological element," then it is at least plausible that I may fail to understand any behavior that expresses this unfamiliar state. Behavior that expresses ununderstandable psychical phenomena will either be familiar or not; if the latter, it will be ununderstandable, if the former it will be misleading.[5]

---

[5] Some of Laing's early work, especially *The Divided Self* (1965) makes the claim that schizophrenic persons, for example, purposefully mislead; their behavior is a kind of "act" put on to cope with their fundamental "ontological insecurity." I

We can thus understand Jaspers' incomprehensibility claim, i.e., the claim that some

psychical phenomena are inaccessible to empathy and thus cannot by psychologically

understood, in terms of an undermining of the proper object of empathy (specifically an

undermining of the psychological context that is a constitutional part of the phenomena to be

understood), or in terms of unfamiliar experience or expression. If one does not participate in the

forms of experience and bodily movement that pertain to the other, or if one at least has not

cultivated an understanding of these forms by experiencing certain descriptions that put one in

such a state as to be able to understand – perhaps an "as-if" state, so to speak – then

psychological understanding will not be possible.

Deficit Modeling

The only recourse to the failure of empathy in the face of radical incomprehensibility,

according to Jaspers, is to negatively circumscribe the target phenomena, understanding them

indirectly "by saying what they are not" (Jaspers, 1997, p. 577). Following Louis A. Sass (2014),

I call this approach "deficit modeling." Deficit modeling represents the limits of psychological

understanding, and is, for a Jasperian approach, the psychologist's only available method in the

face of difference, i.e., in the absence of the similarities that found psychological understanding.

It takes a number of forms in the literature: the target phenomenon may be described as

dissimilar from familiar phenomena in a specific way characterized by some kind of lack or

absence relative to ordinary experience, or as resulting from some deficient or malfunctioning

psychological process (e.g., the literature on "cognitive deficits" or "deficits in reason" as

making sense of pathology phenomena), or as reflecting present but "underdeveloped"

mechanisms, or simply as *not* some other familiar phenomenon. It is this latter that Jaspers seems

do not endorse this position, but there is a kernel of truth to it: *we all* "dress up" our actions, or "act" in the theatric
sense, from time to time and for various purposes, however innocent or otherwise they may be.

to explicitly suggest, but he makes use of these other forms throughout his work, and many contemporary phenomenological approaches to psychopathology do the same.

It must be noted that mechanistic approaches, at least for Jaspers, can offer a form of *explanation* of the target phenomenon, but not *understanding*: they present the phenomenon in terms of an external causality that, while underlying psychological phenomena as a necessary, naturalistic referent, is in itself psychologically meaningless and inapplicable to understanding the kinds of connections exemplified by psychological relations (this is one form of circumscription from "an external point of view," referenced above). At the level of understanding, deficit models cannot offer a positive understanding of phenomena: they never say what the phenomenon *is* or even what it is *like* (a typical goal for the phenomenologist), but only what it is not.[6]

This absence of positivity is not only epistemologically limited (i.e., the deficit model provides only a negative circumscription), but also risks internalizing deficits to the person in question. It risks characterizing persons themselves as deficient. While not necessarily incorrect in all cases, the Jasperian system cannot legitimate such a move across difference: the person taken to be deficient in some psychologically meaningful way is still a form of understanding and thus presupposes empathy; where empathy fails, such a characterization cannot be made. Despite being illicit, this tactic is a common one. Ironically, the most obvious example deals with deficits of empathy, and it is here that we can most clearly see the limitations and dangers of the method.

Numerous scientific and philosophical theorists, including Jaspers (see, as a small sample, Zahavi (2014), Fuchs and Koch (2014), Shah et. al. (2019)), attribute a failure of

---

[6] This has implications for causal explanation. If we cannot determine the sense of the phenomenon, cannot say what it *is*, this will impact our ability to discover "its" causes and effects.

empathy to the subject of certain forms of psychopathological experience. For Jaspers, such a deficit of empathy is the primary indication of the presence of pathology in general. Accordingly, he applies such a deficit to a variety of pathological conditions. In certain cases, the deficits are posited in terms of causal mechanism, and in others in terms of diminutions of familiar phenomena. In both cases, however, there is still enough similarity between the psychologist and their patient that some positive understanding can be grasped by empathy. This approach, when the deficit is precisely of the subject's capacity to empathize, indicates an asymmetry: the psychologist can understand (empathize with) the person who cannot understand (empathize with) the psychologist, and does so by attributing some deficiency to that person. A *pure* deficit method – one according to which *all we can do* is negatively circumscribe the phenomena in question – is typically reserved for certain experiences associated with schizophrenia, such as "made phenomena." In this case, nothing positive can be said of the other and even the attribution of a deficient capacity for empathy must be interpreted not as psychological understanding but as a kind of explanation.

In contemporary work, autistic and schizophrenic persons, for example, are often described as lacking (or otherwise experiencing difficulty with) ordinary empathetic perception. Such a description is a particular instantiation of the deficit approach in which the deficit in question is precisely empathy, seen by Jaspers to undergird psychological understanding in general, and this deficit is taken to belong to the subject in question. One consequence of this view, from a Jasperian perspective, would be that schizophrenic or, for instance, autistic individuals, would have a diminished aptitude for psychological understanding. There is some tension here, since it would seem that such a person would – unlike the struggling psychologist – be familiar with certain kinds of experience that would be necessary for empathy: would not the

autistic person be in a better position, familiar as she may be with the relevant phenomena, to empathize with and thus to understand autism than a non-pathological psychologist?[7] In any case, I am not sure there is support for such a claim, that autistic or schizophrenic persons are less capable of psychological understanding, and if this is not the case then we must question either Jaspers' view of empathy as necessary for psychological understanding, or the attribution of such a deficit to the other. I think there are independent reasons for rejecting both.

A number of explanations are available for the appearance of deficient empathy. Recall that empathy, for Jaspers, is grounded on the perception of expressive behavior as well as a sharing of relevant experience. There are at least two ways that we can express an unfounding of empathy in cases of schizophrenia and autism. The first involves a kind of perceptual distortion. Martin and Pacherie (2013), among others, articulate that a common schizophrenic difficulty in object recognition can be understood in terms of shifting sensational boundaries, e.g., certain features of the visual field are disunified, not smoothly integrated, and do not maintain stable, coherent relations as to give rise to any definite object (at least not in the "normal" way). Increasingly, autism is thought to involve an attentional component, such that some aspects of the environment are attended to to a much lesser degree than is normal, such as facial features and movements.[8] Empathy relies on a certain perceptual territorialization, according to which the expressive movements of the other are perceived as unified spatiotemporal, psycho-physical acts. Yet, if either unification or attention are disrupted, they may not appear this way. A disruption to the perception of expressive behavior would unfound the capacity for empathy, which

---

[7] Jaspers' treatment of the family that co-narrates a shared delusion may indicate that he does at least endorse the claim that the pathological person is in no better position to understand the non-pathological, despite apparent familiarity.

[8] See, for a sample of the autism eye-tracking literature, Dapretto et. al. (2006), Falck-Ytter et. al. (2013), Jones & Kin (2013), Klin et. al. (2002), Riby & Hancock (2009), Sterling et. al. (2008), and Takarae et. al. (2004).

presupposes such perception. Importantly, such an explanation does not posit a disruption of a sense of self as central to deficient empathetic capabilities but does entail an altered perceptual relationship with one's environment.

In Jasperian fashion, we began with a kind of intuition that revealed a certain incomprehensibility. The incomprehensibility is made sense of (negatively) by a deficit model, that is, 'the other is like me except without [empathy].' An explanation is then given to make sense of this deficit. But, the appearance of deficient empathy could be explained in entirely different terms. I offer two relational approaches that attribute this appearance to certain relations between individuals, rather than internalizing them to individuals.

Empathy is made possible by perception of expressive behavior and shared experience. If two persons do not share *both* expressive practices and relevant experience, then they will be unable to understand each other, at least relative to those expressions and experiences that they do not share. In psychopathology cases in which we are inclined to posit some "inaccessible phenomenological element," then empathy is undermined not in an individual but between the two. The psychologist does not empathize with their patient and vice versa. If empathy has to do with matching – of both experience and expressive behavior – then it is reciprocal: if I can empathize with you then, at least in principle, you could empathize with me. The psychologist, as normal – that is, as ordinarily being able to empathize with most people within his culture, and even with most patients insofar as their experiences diverge from his own only by degree or in terms of abnormal causal relations – attributes this deficient empathy to the abnormal person, the person whom he can only conceive, if he reads his Jaspers, by positing such a deficit. 'The other is like me, like the normal person, but without […],' and this makes sense of the failure of empathy that actually bottoms out in the relationship between these two, in their difference that

cannot, for the Jasperian, be understood positively, by indexing it to the person thus negatively circumscribed.[9] This is not to say that the existence of difference is denied, but only that persons cannot be understood by empathy across difference and, as I take to be equivalent, that the difference itself cannot be understood on a Jasperian account as what it is but only as what it is not. But the key point is that the difference in question is relational: it is not something that inheres in one individual but is rather a discrepancy between the person to be understood and the would-be understander. In the face of this discrepancy, to understand the other by deficit violates the ununderstandability claim, and more importantly is unfounded: one only knows that oneself and the other are dissimilar, and it is just as likely that the psychologist, and not the patient, is, in some meaningful way, deficient. It is not clear, however, that all difference, whether it can be conceived with any specificity, must be conceived as deficit.

Furthermore, the perception of expressive behavior does not only depend on one's individual capabilities but is inherently intersubjective and normative: without a shared context to grant movements with significance, bodily motion is merely mechanical. That is, expressive movements gain their significance precisely *as expression*, and thus as meaningful phenomena, by their relation to a context in which those movements count as expressions. This presupposes inclusion in or familiarity with shared practices. I remind the reader of the role of culture in empathy, and of the examples of the death of a loved one and the conference question. Persons with mental illness are typically excluded, at least within the American context, from various practices and communities, and sometimes are placed within isolated communities in the form of institutions such as the asylum, the ward, the psychiatric hospital. Outside of a shared context, it is no wonder why they may seem unable to empathize, but the point must also be made of the

---

[9] This indexing might even be seen *as* the circumscription.

27

psychologist. If the psychologist and patient do not share practices, which can be the result of institutional forces as well as novel behaviors produced to navigate a world experienced "abnormally," then empathy fails, and each figure will be unable to empathize with the other: what becomes motivation for a deficit model for the psychologist becomes the deficit for the patient. Additionally, the enterprises of psychology and psychopathology generate theory-laden significances all their own, familiarity with which cannot be reasonably expected of a non-specialist, but which are taken with the psychologist to the clinical meeting. (Furthermore, as suggested above, the attribution of deficient empathetic capabilities excludes one from the community of psychologists: psychological understanding depends on it).

It may be true that certain forms of experience or life can genuinely involve an *appearance* of deficient empathetic capabilities, but the deficiency cannot be understood in individualistic terms; it is inherently relational, undergirded by experience sharing and inclusion and exclusion in communal practices. To the extent that such a deficiency is taken to be problematic (as the term "deficit" often suggests), intervention must accordingly take place at the level of communities and social practices, and not (at least primarily or exclusively) at that of the individual.

It is admitted by Jaspers that the deficit method does not produce understanding, but only a negative circumscription, and that the method is only negative, i.e., it attempts to say what some phenomenon *is not*, but never what it *is*, or *is like*. Psychopathologists do not often content themselves with such an approach. According to Sass (2014, p. 126),

> There are many psychiatrists whose adoption (often un-thinking) of a pure-deficit model and neurobiological reductionism places them in a similar camp [to those who 'place schizophrenia beyond the pale of the comprehensible']. But many subsequent

psychopathologists who have thought about these issues (Henriksen in press; Sass 2003, 2004a) are disinclined to accept the *radical* nature of Jaspers' distinction [between understandable and ununderstandable], or his view about schizophrenia in particular. Indeed, very few contemporary thinkers of cognitive science, psychiatry, psychology, etc., especially of philosophical bend, would satisfy themselves by thinking of their work as 'mere, psychologically meaningless explanation' that never tells us what the phenomena in question are but only what they are not. The unacceptability of a strictly negative epistemology is even more intense for the phenomenological thinkers, who strive to produce a positive description of their target phenomena; like Hume, phenomenologists broadly agree that if you cannot show me the intuition that warrants a representation, then that representation is literally non-sense.

We can see this refusal to adhere to the negative, a refusal which concretely transgresses the legitimate bounds of deficit modeling while nevertheless being founded in it, at play in much contemporary work. For example, Matthew Ratcliffe, Thomas Fuchs, Joel Krueger, and Jean-Remy Martin and Elizabeth Pacherie do not content themselves with purely negative descriptions: they attempt to understand what some phenomenon *is*. The dominant method begins by positing a deficit, then tracing it through a conceptual system to grasp the positive difference such a deficit would make for the remainder of mental life. For example, Ratcliffe (2013; 2015) argues that the lack of a sense of possibility and change in major depression results in, or partly constitutes, a corporealization of the body. But Jaspers cannot ground this method: where we understand in the true psychological sense, we understand by empathy; where we do not understand by empathy, we can only negatively circumscribe, saying what something *is not*. What, then, can ground the positive signification demanded by contemporary phenomenological psychopathology? What is this conceptual system through which I can seemingly understand the

29

other across difference? I take my hint to this question from the fact that contemporary psychopathologists of various camps, but especially those operating within a phenomenological approach, often take empathy as either the fundamental relation to the other or at least as methodologically primary: we must interact with and observe the other, take their testimony seriously, interpret their behavior, and attempt to understand the phenomenon, as what it *is*, in terms of the familiar.

Transcendental Phenomenology

In contemporary philosophy of cognitive science and psychopathology, the term "empathy" is conceived differently by at least two major approaches: simulation theory and phenomenology. These conceptions of empathy are distinct and, ultimately, incompatible, though they are sometimes conflated. Jaspers is taken to be a significant, even crucial figure by both camps. I argue that Jaspers' approach to empathy is best understood in its relation to the work of Edmund Husserl and Edith Stein on the topic. Though Jaspers is traditionally understood as establishing "phenomenological" psychopathology, in the sense that it is concerned with the lived experience of the subject to be understood, the ways in which his views relate to canonical "Phenomenologists" is debated. Empathy is a crucial notion through which we can articulate one such relationship but, more importantly, transcendental phenomenology – specifically considered from the perspective of empathy and psychopathology, and in its claim to discover necessary structures of experience – allows us to augment Jaspers' proposed "deficit method" and better understand and justify many contemporary strategies in the field.

However, before developing these connections between Jaspers and transcendental phenomenology (*viz.* Husserl and Stein), I should first consider the alternative. I return to the former task in the following section. Contemporary simulation theory approaches to

psychopathology (see Gallese et. al., 1996; Gallese & Goldman, 1998; Goldman, 1995; Gordon, 1995a-c & 1996; Heal, 1995a, 1995b, 1996, 1998, & 2005) make use of 'empathy' and specifically relate the concept to Jaspers. Simulationist approaches trace their roots at least as far back as a contemporary of Jaspers, Theodor Lipps (1851-1914), whose theory of empathy was the prevailing one at the time that *General Psychopathology* was first published (1913). Indeed, Edith Stein's seminal *Zum Problem der Einfühlung* [On the Problem of Empathy] did not appear until three years later, in 1916, and though Husserl references empathy in several senses and for a variety of purposes throughout his work, his first fairly close account was not published until, arguably, 1931, with the *Méditations cartésiennes* [Cartesian Meditations]. These considerations make it plausible to assume that Jaspers' notion is an heir to Lipps', but this is a mistake.

Theodor Lipps, and Simulationism

Theodor Lipps conceived of empathy [Einfühlung :: *feeling-into*] as a kind of projection of oneself onto (or into) the other. Lipps' account is, like Jaspers', rooted in expression. According to Dan Zahavi, Lipps "emphasizes the role of expression and argues that gestures and expressions manifest our emotional states, and that the relation between the expression and what is expressed is special and unique, and quite different from, say, the way smoke represents fire" (Zahavi, 2010, p. 288). Furthermore,

> Lipps argues that when I see a foreign gesture of expression, I have a tendency to reproduce it, and that this tendency also evokes the feeling *normally* associated with the expression. It is this feeling which is then attributed to the other through projection. It is projected into or onto the other's perceived gesture, thereby allowing for a form of interpersonal understanding. (Zahavi, 2010, p. 288; italics mine)

In this regard, Lipps' approach to empathy is continuous with contemporary simulationist theory and, insofar as empathy essentially references expressive behavior as well as one's own experiential life, it might also seem closely aligned with Jaspers' notion.

However, we can elucidate several discrepancies between the two formulations, as presented here. First, Lipps' conception of empathy specifically focuses on emotions and emotional content; Jaspers' is broader in this respect, extending to "thought," "moods," etc. Second, Jaspers' account of empathy involves a grasp of the psychological context out which the phenomena to be understood emerges *for the subject empathized with*. For Lipps, however, I "instinctively" (Zahavi, 2010, p. 288) reproduce the other's bodily movement, this generates a feeling in me, and I project my feeling onto the other. In this case, the other's expression is considered only as a bodily movement (an "objective performance," perhaps) that has a causal effect on me, and the only psychological framework involved in the process is my own. Not only might my own psychological context differ from someone else's, even granting that, at a given time, I and the other experience the "same kind" of phenomena (that is, my anger and someone else's anger might differ radically in terms of psychological context while nevertheless both counting as 'anger'), I nevertheless needn't grasp or otherwise understand even my own psychological context: I simply live a psychological state or process, generated within a framework that may remain obscure to me, and impose that individual state or process onto the other. Lipps' empathy is thus severed from the psychological context of the other in at least one way relevant to Jaspers' conception. Third, relatedly, the target of Lipps' conception of empathy is a particular emotional state of the other: what I understand is their current emotion. For Jaspers, however, what we understand by empathy is the other person, considered as a whole. Fourth, empathy on Lipps' conception amounts to what might instead be called "sympathy"

(*same-feeling*): I actually undergo the emotion that I project onto the other. There is thus a kind

of matching between the experience of the empathizer and the person empathized with. For

Jaspers, however, there is no such requirement: the empathizer need only have had, at some point

in time, a similar experience.[10] It is certainly not the case, for Jaspers, that empathy is based on a

motor response that reproduces in me, at the very time of the act, the feeling that is to be

(subsequently) projected onto the other.[11]

Phenomenological Empathy: Husserl and Stein

Jaspers' notion of empathy is importantly distinct from Lipps' and, for many (typically,

all) of the same reasons, contemporary simulationist formulations. On the other hand, it is

closely related to Edmund Husserl and Edith Stein's account, which agrees with Jaspers' on

those points just presented. Historically speaking, this similarity can only be loosely attributed to

any direct influence that Husserl and Stein may have had on Jaspers.[12] More plausibly, the

connection is owing to the fact that all three figures were influenced by the work of Wilhelm

Dilthey and Max Weber, the latter of which Husserl influenced in turn (Muse, 1981).[13] In any

---

[10] Despite the quote (Jaspers, 1997, p. 306) used to establish this point on page 4 above, it is sometimes unclear whether one has to have actually undergone a similar experience themselves, or rather if they need only have been familiarized with it in a looser sense. Jaspers asserts that the practicing psychologist must study literature and poetry, and that doing so builds the kind of familiarity with experiences in a way useful for developing their craft. An analysis of the kind of familiarity gained by such study would be necessary for further disambiguating the point, but such an analysis is not forthcoming in Jaspers. I thus stand by the earlier quote and suppose, literally and naïvely, that one must have "had undergone the same experiences" (Jaspers, 1997, p. 306).

[11] In his "Phenomenological Approach to Psychopathology" (1968, originally 1912), Jaspers seemingly defined empathy similarly to Lipps: "Subjective symptoms cannot be perceived by the sense-organs, but have to be grasped by transferring oneself, so to say, into the other individual's psyche; that is, by empathy." Yet, I do not believe this definition remains consistent even within this article, and it is no longer operative by the time on *General Psychopathology*, only a year later. Indeed, prior to this definition, he counts Lipps as a phenomenological predecessor to Husserl, and so it is arguable that at this time already Jaspers' conception of empathy is somewhat confused between Lipps and Husserl, though informed by both.

[12] Jaspers does indeed cite Husserl as the source of his "phenomenological" method, but, again, neither Husserl nor Stein had developed a nuanced theory of empathy at the time that *General Psychopathology* was published.

[13] Indeed, Weber's use of *Verstehen* to refer to a kind of interpretative or communal sense making, often called both "understanding" as well as "empathy" can help explain the close connection between the two concepts in Jaspers.

case, the work of Husserl and Stein can be seen as a development of the concept of empathy that is consistent with Jaspers and has implications for his view and method. Not only that, but many approaches in contemporary psychopathology and philosophy of psychology and cognitive science cite both Jaspers and Husserl (or some figure who, broadly, develops the latter's approach, e.g., Heidegger) as major influences.

The phenomenological notion of empathy is best worked out by Edith Stein and Edmund Husserl.[14] The former devotes the entirety of her *Zum Problem der Einfühlung* [On the Problem of Empathy] to developing the concept in terms of perception, individuation, and expression. She also provides a discussion of other senses of 'empathy,' including that of Theodor Lipps and Max Scheler, both of whom Stein and Husserl alike are critical of. Husserl is largely in agreement with Stein's account of empathy, though they arguably (and depending on where you look in Husserl) disagree at several points, including the various forms or "levels" of perceiving foreign consciousness, what each level requires, and the proper objects of certain forms of empathetic perception. For my purposes, the theory of empathy I will make use of is drawn primarily from Stein's *On the Problem of Empathy* and Husserl's *Cartesian Meditations* (as well as selections from the 1923/24 lectures, published in the Husserl, 2019 (*First Philosophy*), which preceded the *Meditations* in which they are included in a modified form). Husserl's work in *The Crisis of European Sciences and Transcendental Phenomenology* presents a somewhat different

---

[14] Dan Zahavi carefully presents a number of phenomenological approaches to empathy, including Husserl and Stein's, in multiple places. For some examples, see his articles "Empathy, Embodiment, and Interpersonal Understanding: From Lipps to Schutz" (2010) and "Beyond Empathy: Phenomenological Approaches to Intersubjectivity" (2001), as well as his book *Self and Other: Exploring Subjectivity, Empathy, and Shame* (2014). He correctly notes that there is no *one* approach to empathy shared by the phenomenologists, though all accounts – he thinks – share certain features. For my purposes here, Husserl and Stein develop the most relevant approach. Heidegger and Merleau-Ponty displace the significance of empathy, Levinas presents a critique of empathy that I think is underappreciated by Zahavi and the literature in general (this comes into play in the following chapter), and Scheler, Sartre, and Schutz (others that Zahavi focuses on) will not be discussed in any detail here. While Zahavi's work represents, amongst other things, an invaluable resource for anyone attempting to understand phenomenological notions of empathy, and while I am personally indebted to his work in this way, I have some different interpretations than does Zahavi and, in any case, prefer to anchor the current discussion directly in the texts and my own interpretation to the greatest extent possible.

picture than these earlier works and becomes quite important in view Heidegger's influence on

Husserl, and so I will explicate certain insights from *The Crisis* where appropriate. Since Stein

and Husserl are largely in agreement regarding empathy, at least until somewhere between the

Meditations and *The Crisis*, I will treat their accounts as unitary, except where noted.

Whereas Jaspers is primarily interested in grounding our psychological understanding of

others, the Phenomenologists prioritize intersubjectivity more broadly, asking how it is that we

can be and actually are related to others, and how we can understand (or fail to understand)

others at all. The answers to these questions have important implications for phenomenology in

general, particularly for Husserl. Attempting to reformulate the aims and methods of psychology

through Franz Brentano's "discovery" of intentionality, Husserl maintains that "it would of

course be pointless to treat the positive science of *intentional psychology* and *transcendental

phenomenology* separately" (1931, p. 147).[15] He notes that psychology and transcendental

phenomenology have the same content: the psychological ego and the transcendental ego are one

and the same. Thus, the insights provided by transcendental phenomenology should have direct

relevance to the significance and practice of psychology. Indeed, it is almost trivial to suppose

that an account of how we can understand others *at all* will be relevant to the question of a

specifically psychological understanding of others.

According to Stein, empathy "is a kind of act of perceiving *sui generis*," "the experience

of foreign consciousness in general" (Stein, 1989, p. 11). By *sui generis,* Stein and Husserl mean

that such a perception is a form of intentionality irreducible to other acts, such as memory or

imagination. However, Stein and Husserl do appeal to the nature of memory and imagination to

elucidate the nature of empathy. In remembering joy, for instance, Stein writes:

---

[15] This claim and related also appear in Husserl's *Crisis* (1970) and Encyclopedia Britannica article(s) *Phenomenology* (1997)

the "I" as the subject of the act of remembering, in this act of representation, can look back at

the past joy. Then the past joy is the intentional object of the "I," its subject being with and in

the "I" of the past. Thus the present "I" and the past "I" face each other as subject and object.

(Stein, 1989, p. 8)

Within the structure of the act of memory, there is a distinction between the "I" that remembers

and the "I" being remembered, but these are both experienced with a "consciousness of

sameness," that is, as both being *me*. This is not true in the case of empathy: "The subject of the

empathized experience, however, is not the subject empathizing, but another" (Stein, 1989, p.

10). This distinction between self and other is constitutive of empathic intentionality: were the

"I" empathized-with experienced as a duplication of myself – as a modified but nonetheless same

'me' – then empathy would no longer be an experience of foreign consciousness; rather, it would

be a matter of projecting ourselves into the perceptually present body of the other and in such a

case we would never be able to experience anyone else irreducible to ourselves. (This forms a

foundational part of the phenomenological critique of Lipps).

At its most basic level, empathy is an experience of the other person as a subjective

perspective on the world – as a site of sensation and animation – similar to but distinct from me.

The empathetic act itself is achieved by associative and identifying syntheses, as well as what

Husserl calls "bodily coupling" and "transfer of sense." We experience our own expressive

movements as psychophysical unities, such that bodily movement and mental life, in the form of

intentions, effort, moods, etc., are inseparably associated. That is, I am never originally given my

own body as a *mere* body, either in its movements or as an entity. My movements are not given

as mere mechanical motions but, instead, as actions; my body is no mere physical object like any

other but a necessary aspect of lived experience and a point of orientation for and reception of all

my sensible experience. When I swing my arm, the movements are inseparable from my sensations, intentions, thoughts, concerns, what I know, what I want, etc. For example, I swing a hammer because I want to build, or to destroy, because I know the hammer is the tool for the job, because I can lift the hammer and see the nail, the wood, and doing so involves a sensation of my body as moving, a feeling of the heft of the hammer, the contraction of muscle, etc. not as objective phenomena observed like any other but rather as given to me directly. The swing becomes constituted as a psychophysical phenomenon: it consists of visible movements but also intentional states that, in my own case, I live through. When I witness the expressive behavior of someone else – when I see them swing a hammer – I directly perceive only the physical aspects of the act, the visible movements. However, this is an act like mine – a psychophysical act – and their body, as performing such an act, is perceived as a body like mine. As like mine, our bodies are "coupled" – they are distinct instances of the same *kind*, and as being of the same kind, there occurs a "transfer of sense" at the level of perception, informing the *what* that is perceived: the expressive behavior of the other, *as expressive behavior* (like mine), is perceived as a psychophysical unity, though I can only perceive one aspect thereof. This already grounds the sense of the other as a source of animation, i.e., of action rather than mere movement.[16]

What separates empathy from other forms of perception, and also from projection in Lipps' sense (which involves, first, a kind of sympathy), is that the mental correlates of the expressive behavior of someone else can never be given to me in the same way as those of my

---

[16] In *Cartesian Meditations* (1931, pp. 117-120), Husserl also argues for the constitution of the sense of the other as a locus of sensation and animation in the following, somewhat convoluted, way: We know from our own experience that, though I can conceive of my own action *as if* I occupied a different space than I currently do, I cannot *actually* occupy two distinct spaces simultaneously. When I actually see the expressive movement of a body *like mine* somewhere other than in the space that my body occupies, then it must be concluded that it is not *my* body but the body of an other. This body is something seen, and specifically something seen that is like my own body in form and action, and so it is known as my own body, but different – as an "alter ego" – as a locus of sensation and animation, like me, but distinct from me. It is this perception that grounds the category of "man," and indeed Husserl refers to the body of the other as the first body, and the other as the first man.

own; mine are presented to me as lived through, but those of the other are *appresented* as belonging constitutively to their expressions. According to Husserl, the experience of the other in empathy is "direct," in that the other is given to me as actually present "in the flesh." However, empathy never gives the experiential life of the other *primordially*, in the sense of being given in the same way as my own. As noted, this is not a failure of empathy: the non-primordiality of the givenness of the other is what distinguishes them from me, and thus what makes empathy an experience *of the other* and not just a projection of myself. It is an essential feature of empathy.

We can here note and resolve an apparent tension between Jaspers and the phenomenologists. According to Jaspers, empathy is representational and indirect, whereas for the phenomenologists it is direct but non-primordial.[17] Despite their terminology, we can see that "direct" and "representation" have different meanings for these thinkers: by representational, Jaspers means not presented to me, whereas for Husserl the term means not given *at all*. Jaspers would consider appresentation a form of representation, while Husserl would not. By "direct," Jaspers means that something is given to me immediately, whereas for Husserl "direct" means given as here, "in the flesh." Both agree that empathy is a *mediate* mode of awareness, but for Husserl it is nevertheless direct. We can thus correspond Jaspers' "direct" with Husserl's "primordial," and Husserl's appresentation is, in Jaspers' terminology, a form of representation. (Thus, the way that empathy is "direct" for Husserl, namely that empathy appresents the states of the other, would qualify empathy as "representational" for Jaspers.) For both, our experience of the other is distinct from our experience of ourselves: it is non-primordial, we do not experience their psychological states as we do our own; it is mediate, it requires a relation between the

---

[17] The act itself is primordial – I live through the *act* of empathy; it is the content, i.e., the experiential life of the other, that is presented to me non-primordially.

behavior of the other and our own psychological states bridged by perception and our own expressive acts.

Furthermore, the non-primordiality of the content of empathy grounds our sense of the world as *the objective world*. Were it not for empathy, we would be in the situation of what may be called being-alone, solipsistic being specified only partly by the "reduction to the sphere of ownness" articulated in *Cartesian Meditations*.[18] Without an awareness of others, our experience of the world is founded on what is given by sensation, and objects take the form of that which exceeds any particular intentionality of my own. My mug, for instance, exceeds my viewing it from this angle… or this one… or this one… With my experience of the other, however (i.e., empathy), the object becomes understood as exceeding *all* my actual or possible perspectives on it. It is something available for some other ego, i.e., for a foreign consciousness irreducible to me.

It is crucial that this "reduction" is an abstraction from lived experience; an all-too-common mis-interpretation of the *Cartesian Meditations* takes Husserl to claim that we are at first, or at least fundamentally, before becoming explicitly aware of particular concrete others, in the situation of being-alone, and that the world must then somehow be constructed from resources entirely *my own*. Instead, the world that we live in is already suffused by the other, by objects and by practices that already presuppose intersubjectivity. The "reduction" is meant to begin with this world and then, by "bracketing," abstract to that which is only given to me primordially, that is, originally in sensation, and to show that doing so would not negate the possibility of the experience of the other ordinarily taken for granted. The goal is not show that

---

[18] Being-alone can only ever be partly explicated because, in my understanding, the situation is itself incoherent. Even within the sphere of ownness, according to Husserl, we experience others; we simply bracket (or "parenthesize") the constitutional effects of intentionalities directed at other egos for our most fundamental strata of experience – nature as given by sensation – while we nevertheless must acknowledge such intentionalities as part of what is given to us.

either the world or the other, nor even their appearances as constituted for me, are contained in or constructed by myself, but rather to demonstrate that phenomenological self-reflection, and the insights derived from this practice, are not committed to solipsism.[19]

To sum up the phenomenological notion of empathy here presented, and particularly in relation to Jaspers: it is consciousness of the other, conceived of as a mediate form of perception, based on the perception of expressive behavior, bodily coupling (i.e., perception as of a body *like mine*, in its form and action), and a transfer of sense from the unified nexus of my psycho-physical reality to the body of the other. That is, I see the movements of the other as one aspect of a psycho-physical unity, just as are my own movements; I perceive their expressions as the same type of thing as mine, i.e., expressions of a mental life, similar to but distinct from my own. This all agrees with Jaspers' notion, and so I now opt to use them interchangeably, except where some difference becomes salient in which case it will be noted and considered.

Transcendental Structures and Positivity

My purpose has been to read Jaspers as broadly consistent with phenomenology as developed by Husserl and Stein, specifically regarding the notion of empathy crucial to the former's proposed method of psychology and psychopathology. Where Jaspers and these phenomenologists seemingly diverge is that the former is concerned with an empirical, scientific method, while the phenomenologists first and foremost are concerned with transcendental philosophy (Wiggins & Schwartz, 2013, p. 25).[20] We saw above that a shortcoming of Jaspers'

---

[19] Robert W. Jordan provides a nuanced discussion on this topic in his "On Not Living in the Primordial World," presented at a number of conferences and reserved for an anthology that was never released. In this work, he also illustrates that the aforementioned interpretation is based on flaws in the work, but that "Husserl himself corrected the worst of them in manuscripts written only five to six months after the book manuscript was given to French translators. These were published in Husserliana XV (1973) some fourteen years ago."

[20] Remember, however, that Husserl takes psychology and phenomenology to be closely related, as sharing subject matter, and so believes the insights of transcendental phenomenology will be able to be taken up by psychology. The

approach is that, where empathy fails, we are limited to a purely negative understanding of the phenomena that interest us.[21] Transcendental Phenomenology, however, offers us a way to augment an empathy based approach so as to provide deficit models with a positive significance or meaning. This is due to Transcendental Phenomenology's discovery of 'necessary' or 'invariant structures of subjectivity.'

Transcendental Phenomenology claims to discover necessary structures of subjectivity. For instance, all experience is structured by temporality, and empirical intuition is always given in accordance with a horizon of possibilities. While these are both necessary structures of subjectivity, they are not of the same level: the former founds the latter, as the open-ended future undergirds possibility. These structures are necessary *essentially*. That is, they are conceived through a kind of eidetic reduction and variation that aims to uncover eide, or essences, by abstracting from individual concrete phenomena to a kind of ideal type. As an oversimplification, the method involves an imaginative variation of features of some phenomenon, in an attempt to understand what could be different (what could vary) while the phenomenon nevertheless remains of the same type (for some given type of interest), and, crucially, what could not vary without a fundamental change in kind. That is, the method attempts to discover essential invariants. For example, were the notes of a melody to occur not in succession but simultaneously, they would no longer count as a melody but instead as a chord. It is essential to a melody, then, that notes follow one another over time rather than occurring all at once. It is not essential, however, that the notes follow one another in a specific rhythm, or in a certain order, or that the melody be produced on one instrument rather than another. These latter

---

two fields are not utterly distinct, however Husserl awards priority to the former whereas Jaspers focuses nearly, if not entirely, on the latter.

[21] Though, that is not to say that we can have no *explanations*, viz. objective causality.

features are inessential to melody, whereas the former – that the notes unfold sequentially – is essential. The method of eidetic reduction and variation allows Husserl to make the claim that consciousness is *essentially*, i.e., *necessarily or invariably* temporal, structured by a horizon of possibilities, and so on.

While Husserl does offer other derivations of these structures, the method of eidetic reduction and (imaginative) variation offers us the clearest understanding of the sense in which these structures are meant to be "necessary," and, importantly for my purposes here, these methods are frequently taken up in contemporary phenomenological psychopathology, especially when it comes to apparent cases of radical difference. First, the sense in which these structures are necessary is that they are necessary for the phenomenon in question to count as the type of thing it is. For something to count as consciousness, for example, it must be temporally structured. For something to count as empirical intuition, there must be a reference to a horizon of possibilities. Husserl does allow, however, negative reference: a particular intuition may *lack* a horizon of possibilities, but at that point it counts as a different or abnormal type of phenomenon, something to be understood precisely by reference to this lack relative to normal experience. There is thus reference here to not only a kind of essential, ideal type and identity, but also to similarity in normative and non-normative senses: the reflecting phenomenologist performs a reduction and variation *on his own experiences* (that is, she is self-reflecting), discovers a universal type, and imports this to the sense of the phenomenon at a constitutional level: the phenomenon *is* thus-and-so, universally, based on a specific form of self-reflection. One's own experience is thus taken as the norm against which to test variations. But, because of the method, the findings are taken to exceed one's own subjectivity: they identify a universal type that *any* experience, no matter who's, must conform to if it is to count as the same type of

thing that I discover for myself. There is a related and (at least seemingly) richer sense in which these structures are necessary. Husserl develops a system of strata which are ordered by grounding relations (e.g., temporality grounds or structures possibility).[22] Thus, the structures become "transcendental" in a nearly Kantian sense: some are necessary conditions for others. So, for instance, a horizon of possibilities is necessary for the consciousness that things might change.

As referenced in the previous section, these structures become crucial for understanding the other across difference in contemporary phenomenological psychopathology. Ratcliffe, for example, claims that the experiential life of persons with major depressive disorder ("MDD") is not structured in the ordinary way by a horizon of possibilities. This form of consciousness is thus essentially different, in this respect, from "normal" consciousness. However, he needn't – and does not – stick only to a deficit model, though it is where the account begins. Transcendental Phenomenology grounds a method by which we can supply a positive signification to this apparent lack. If the horizon of possibilities is absent (conspicuously so, according to Ratcliffe), and if the horizon of possibilities is necessary for any sense that things could change, then those experiences without a background of possibility will include a sense that what is given (and perhaps experience itself) is unchanging, inevitable, perhaps even eternal. He uses this kind of transcendental analysis to make sense of claims made by persons who suffer MDD that nothing can change for them or, sometimes, that nothing can change at all, for anyone (Ratcliffe, 2013 & 2015). Many thinkers in phenomenological psychopathology apply the method in other cases as well. For instance, it is a popular position that schizophrenia involves a

---

[22] We must not mistake Husserl's understanding of "system" with a deductive, complete system. He attempts to uncover various strata and relations (e.g., of founding) between then, but does not attempt to construct a final system once and for all, nor to build a system on top of or in analogy to traditionally accepted logics.

lack of ordinary temporality. If the absence or "disturbance" refers to the future, that is, if one is not aware of the future in the ordinary ways, then this is often taken to undermine the "sense of agency" that one ordinarily has over their actions, since an anticipation-fulfillment structure is necessary for a sense of intentional action (Frith, 1992; Gallagher, 2005). If the disturbance refers to the past, then it is thought to imply an interruption of an ordinary "sense of ownership," since ownership requires that an experience be situated amongst my others over time (Martin & Pacherie, 2013; Metzinger, 2003; Zahavi, 2005). These states are then likened to the familiar by a series of analogies and similarities. For instance, non-pathological, or psychologically "normal" individuals undergo a number of experiences every day that lack the sense of agency that usually accompanies intentional action. One common example is perception: within perception, even if we agree with phenomenologists, pragmatists, and enactivists that it essentially involves a form of action, we do not typically experience the objects perceived as if they were created by us. The "made phenomena" that Jaspers references, and specifically what is known as thought insertion, can then be understood in these terms: thought insertion is like thought, in that its objects are intentions, willings, desires, fears, etc., but is like perception in that these objects are taken as existing independently of me, as created not by me but in some other way.[23]

Not all – perhaps not even most – phenomenological psychopathologists explicitly leverage Husserl and Stein. Popular figures also include Heidegger and Merleau-Ponty, and occasionally Scheler. These thinkers do not utilize empathy in the same way as do Husserl and Stein, though they do often trace back their methods to Jaspers. Heideggerians tend to make use of being-in-the-world and, rather than empathy, being-with. Merleau-Pontians utilize

---

[23] Something like this position is articulated by Ratcliffe & Wilkinson (2015) in "Thought Insertion Clarified."

intercorporeity to further articulate bodily being in the world and our relation to others.[24] In all cases, however, there is a kind of community assumed between self and others, some fundamental structures uncovered, and their interconnections and implications laid out.

The method is as follows. Begin with a set of transcendental structures of subjectivity, or "existentialia," or what have you. Then, encounter an other such that empathy – or being-with, or intercorporeity – fails, that is, encounter an other such that, relative to some phenomena, the ordinarily fundamental relation between persons does not afford any understanding. Next, negatively circumscribe the phenomena in question by saying what they are not; negate, subtract, remove (Jaspers' method ends here). Finally, plug that negation into your system of transcendental structures – essentially the structure of ideal subjectivity – and see what other features of that structure would apparently be disrupted, or if any positivity (such as corporealization) should be produced.

The Question of the Intersubjective Validity of Transcendental Phenomenology

Transcendental Phenomenology, especially in early Husserl, conceives of its method as a kind of self-reflection. This aspect of phenomenology is reflected in the title of Husserl's *Cartesian Mediations*: he is interested, like Descartes, in discovering the foundations for knowledge by a process a meditation. Throughout the first four books of *Cartesian Mediations*, Husserl sticks to a method of self-reflection to the extent that, by the fifth book, he is concerned that people will view phenomenology and its method as necessarily solipsistic, despite his frequent reminders (typically footnoted) that the first four books make implicit use of intersubjectivity and a theory of empathy, articulated in the fifth. Even before the *Cartesian Mediations*, Husserl engages in this kind of apparently solipsistic method of self-reflection, and

---

[24] I do not mean here to reduce being-with, intercorporeity, or empathy to one another.

the *Meditations* was not altogether successful in ultimately dismissing the charge. This is in part

due to misreadings of the work, but also because the fifth meditation is indeed problematic.[25]

The question is: If phenomenology proceeds from a method of self-reflection, abstraction

though it may be, with what validity can its findings be imported to others? Is it not *merely* self-

reflection, the products of which remain with the self? I think not, in general, and the

overwhelming majority of phenomenologists (if not all) will agree with me. However, I here

attempt to demonstrate that psychopathology broadly, and specific forms thereof in particular,

challenge Husserl's solution to the question. Furthermore, Heidegger offers a different solution

than does Husserl, and through Heidegger's influence Husserl comes quite close to the answer

offered in the former's *Being and Time* in his own *The Crisis of European Sciences and

Transcendental Phenomenology*. In order to explicate these solutions – that of early Husserl

(through, at least, *Cartesian Mediations*), on the one hand, and that of Heidegger and *Crisis*

Husserl, on the other – requires an understanding of the ultimate basis of world-experience at

play in each. In the first case, this is sensation; in the second, it is practical engagement. I then

move to introduce a dilemma that challenges the intersubjective validity of transcendental

phenomenology specifically in cases of psychopathology that appear to involve radical

difference.

Sensation and Similarity

Husserl regards sensation as the ultimate basis of experience. In this sense, he strives to

be a good empiricist, following thinkers such as Hume, whose copy principle was noted above,

and Heraclitus, who "prizes" "the things of which there can be sight, hearing, and learning"

(Wheelwright 1959, p. 19). For Husserl, sensation is an abstractive moment of perception, and

---

[25] I again recommend Robert W. Jordan's article, "On Not Living in the Primordial World," for resolving some of these misconceptions and addressing the problematic nature of the fifth meditation itself.

perception is temporal. Perception requires not only sensation (as a kind of substrate) but also

temporal synthesis of sensation, insofar as sensation only becomes *meaningful* against a temporal

horizon of both prior sensation and future possibilities. Thus perception ultimately refers to

sensation, while the very meaning of any given abstractive instant of sensation refers to the

retentional/protentional structure of consciousness that constitutes perception. Perception as a

meaningful act depends on a prior sensitivity, by which hyletic data are given to consciousness,

and temporal synthesis of this data.[26] This hyletic data, or what is sensitively impressed, in time

passes into the past and becomes the foundation out of which meaning is constituted.

Correlatively, the possibilities that make up the horizon of possibilities are, at the most

fundamental level, possibilities of future sensation, i.e., that which *may* be given as hyletic data

to sensation (even if these data are already *understood* from a temporal perspective). Sensation

then is a fundamental form of experience insofar as it is what provides the basic content of

experience, and sensible data is also the basic content from which all other content is derived.

Crucially, for sensation to mean anything requires reference to a horizon of the past: the

meaning of an occurrent sensory impression requires that past sensation be apprehended along

with the hyletic data presented at the present moment. The latter is presented as "now," the

former as a past now, a now that has been modified into the form of "past," and, similarly, a

horizon of the future (that is, what something is is in part a function of what something could be).

In any moment out of which the phenomenologist may begin his method and abstract from some

particular set of data, with its correspondent present, her sensation has already been incorporated

into a meaningful, temporal act that can in principle be grasped. That is, Husserl's notion of

---

[26] This givenness is sometimes called "primal impression," but again this is an abstraction, and it must be remembered that sensation is not separable from retention and protention, as the tripartite "protention-impression-retention" structure might easily be misinterpreted to imply.

sensation is not discontinuous with perception, contrary to the claim made by certain thinkers –

reductionists, some perception theorists and cognitive scientists among them – that sensation is a

kind of intrinsically meaningless building block which, when considered in combination with

other moments of sensation, add up to perception. Husserl follows Brentano here, in claiming

that intentional (or psychological) relations cannot be understood in terms of external causality, a

causality which simply combines various elements in order to understand the product as an

effect. Instead, sensation is unified over time, and this unification constitutes the meaning of

sensation at *each* moment; since sensation is an abstraction from perception, the meaning of

perception conditions our sensory experience, and insofar as sensory experience undergoes

temporal synthesis (and thus can be meaningfully thought at all), it is already incorporated into

the very constitution of perceptual acts. For Husserl, perception is not just a product of a number

of sensational states or contents standing in causal relations to one another.

Though sensation is a fundamental (though abstract) stratum of perceptual experience in

general (in an intransitive sense, but also in the transitive: experience… of the world,

experience… of ourselves), it does not directly underlie *all* forms of experience. Certainly,

mathematical entities, for instance, are not given to us by sensation (temporal synthesis does not

transform any sensory data into a perception of, e.g., *the number 2 itself*). And the foundation of

the perception of others (here, people or egos), as Husserl and Stein both make clear in their

account of empathy, is not exhausted by sense – the most interesting aspect of others for the

phenomenologist, i.e., their experiential being, is not given by sensation alone but by various

associative syntheses (including coupling and transfer of meaning). Yet all this, for (early)

Husserl, bottoms out in sensation. It is from sensation that we can abstract, according to specific

methods, mathematical concepts which we can only subsequently demonstrate. Once a proof is

produced, however, it is also demonstrated that the mathematical concepts or entities can be

infinitely re-produced, that they do not essentially depend on the time of their original

production, and thus that they are eternal.[27] We can be aware of the other on the basis of our

perception of their body, subtended by what is given of that body to sensation, in addition to our

own sensible experience, which reveals my own body as inseparable from a conscious life.

Though there are sophisticated forms of experience irreducible to sensation – Husserl does not

attempt to *reduce* experience to sensation – they are nevertheless founded on sensation and its

corresponding hyletic data. This seems to be part of what Husserl has in mind when he makes

claims such as

> […] every mode of consciousness involves its possibilities of an uncovering of what is
>
> intended, its possibilities of becoming converted into either fulfilling or disillusioning
>
> experiences of what is meant, and moreover, (as regards the genesis of the consciousness)
>
> points back to such experiences of the same intended object or a similar one. (Husserl, 1931,
>
> p. 106)

The way that Husserl hopes to avoid solipsism in his *Cartesian Meditations* depends on

his account of sensation. As mentioned previously, the first four books of the *Meditations*

presuppose his findings from the fifth, in which he attempts to address the charge of solipsism by

performing a "reduction of transcendental experience to the sphere of ownness" (Husserl, 1931,

p. 92). He writes:

> As regards method, a prime requirement for proceeding correctly here is that first of all we
>
> carry out, *inside the universal transcendental sphere, a peculiar kind of epoché* with respect
>
> to our theme. For the present we exclude from the thematic field everything now in question:

---

[27] I am indifferent as to whether this is an accurate account of mathematics; it is Husserl's, though intensely oversimplified.

we *disregard all constitutional effects of intentionality relating immediately or mediately to other subjectivity* and delimit first of all the total nexus of that actual and potential intentionality in which the ego constitutes *within himself a peculiar ownness*. (Husserl, 1931, p. 93)

What interests Husserl at this point is the way in which I myself could ever distinguish between myself, what is peculiarly mine, and an other, that is, how I could distinguish between my Ego and an other Ego.[28] Within this reduction appears my own intentionality, Husserl believes, directed towards the other. That is, the act of empathy is *my own* or, as Stein writes, with empathy "we are dealing with an act which is primordial as present experience though non-primordial in content" (Stein, 1989, p. 10). The primordiality of the act refers to the way in which empathy presents its object, that is, the other, through sensation and mediating syntheses, while the non-primordiality of the content refers to this very mediation, i.e., that the being of the foreign Ego specifically as ego is not given as hyletic data to sensation (though its givenness is founded on such data). In attempting to understand the way that one first constitutes, for oneself, something other than oneself, Husserl, while maintaining that the experience of the other is direct in the sense that "the other is himself there before us 'in person," clearly maintains that "properly speaking, neither the other Ego himself, nor his subjective processes or his appearances themselves, nor anything else belonging to his own essence, becomes given in our experience originally" (Husserl, 1931, p. 109). Instead, the other is given mediately, "making present to

---

[28] In a footnote, Dorion Cairns (trans.) remarks that Husserl later appended the following comment: "inside the universal transcendental sphere – "peculiar epoché". But it is misleading when the text goes on to say: "in that we exclude from the theoretical field everything now in question, in that we <disregard> all constitutional effects that relate immediately or mediately to other subjectivity," etc. The question after all concerns, not other men, but the manner in which the ego (as the transcendental onlooker experiences him transcendentally) constitutes within himself the distinction between Ego and Other Ego – a difference, however, that presents itself first of all in the phenomenon, "world" : as the difference between my human ego (my Ego in the usual sense) and the other human ego (the other Ego <likewise in the usual sense>)." (Husserl, 1931, p. 93)

consciousness a 'there too', which nevertheless is not itself there and can never become an 'itself-there' […] we have here, accordingly, a kind of *making 'co-present'*, a kind of *'appresentation'*" (Husserl, 1931, p. 109).

Through empathy, the other is first constituted as a kind of "mirroring" of my ego, as a "second" ego, which is "not simply there and strictly presented; rather he is constituted as "alter ego" […] The "Other", according to his constituted sense, points to me myself" (Husserl, 1931, p. 94). We saw earlier that empathy is founded on the perception of expressive behavior and certain associative syntheses, and here we see that it is constituted precisely as an alter ego, that is, as something *like* me *but not* me. Husserl thinks that he can account for the constitution of this form of alterity – a strange sort of analogue[29] – by the reduction to the sphere of ownness, which ultimately leaves sensation intact.

When we thus abstract, *we retain coherent stratum of the phenomenon world*, a stratum of the phenomenon that is the correlate of continuously harmonious, continuing world-experience. *Despite* our abstraction, we can *go on continuously in our experiencing intuition*, while remaining exclusively in the aforesaid stratum. (Husserl, 1931, p. 96).

This "phenomenon," or "primordial," world is what Husserl calls the founding stratum, and is prior to the world constituted as "objective." For this, we must first have some experience of others, such that the world can be given as in excess to my own intentionality, becoming something shared.

Within the primordial world our own body is "*uniquely* singled out" precisely as an animated body, to which "I ascribe *fields of sensation*." I further know my body through *kinesthesias* in the mode "I am doing," that is, in accordance with an "I can" (in contrast with a

---

[29] "the other is a "mirroring" of my own self and yet not a mirroring proper, an analogue of my own self and yet again not an analogue in the usual sense." (Husserl, 1931, p. 94)

Kantian "I think"). This founds the consciousness of my body as a psychophysical unity. He then takes the term "Other," referring to "alter" as his hint for discovering our intentionality of the other: we experience the Other as like oneself, a psychophysical unity of sensation and animation, but nevertheless distinct from oneself. What is presented of the other, within the reduction to ownness, is their body. But this body is given as an *animate organism*. The only other body given as such a thing, within the primordial world, is my own body. And so, the body of the other "apprehended as an animate organism," according to Husserl, "must have derived this sense by an *apperceptive transfer from my animate organism*" (Husserl, 1931, p. 110). The condition on which this can occur cannot be direct, as the sensation and animation of the other ("I can") cannot be given to me in the same way as my own. And so, crucially,

It is clear from the very beginning that only a similarity connecting, within my primordial sphere, that body over there with my body can serve as the motivational basis for the "*analogizing" apprehension* of that body as another animate organism (Husserl, 1931, p. 111).

Husserl notes, although it is sometimes overlooked, that this apprehension (which he calls an "assimilative apperception") is non-inferential, distancing his view from a popular theory of empathy as analogical inference. Instead, the apperception "points back to a 'primal instituting" (Husserl, 1931, p. 111), of a kind that is familiar in everyday experience. This appresentation of the other as an animated organism *like me*, i.e., the making co-present of subjective facts in the perceptual presence of a body, is instituted by "pairing," which is "*a primal form of that passive synthesis* which we designate as "*association*", in contrast to passive synthesis of 'identification." (Husserl, 1931, p. 112). Intentionally speaking, when two data are given together as similar, then when one is presented the other is co-intended, that is, they are given as

a "pair," and there occurs an "overlaying of each with the objective sense of the other," what Husserl calls a "mutual transfer of sense." What is given in sensation of the body of the other, due to the data's similarity to what is given in sensation of my own body, becomes perceived, through mediating synthesis, as the same kind of thing as my own body.

For an ordinary example, if I see an unfinished table, what is given gains its meaning and significance from other similar perceptions, that is, from other tables that I have experienced. In my experience, it is unsafe to place a glass of water directly on the surface of an unfinished table. I do not *infer* that this table, being like those, is also unsuitable for such a purpose. Rather, I experience this table precisely *as like* those others; those others form a horizon against which this table appears as the type of thing it is, and, being perceived as like those, it is also perceived as – not reasoned to be – unsuitable for placing a glass of water directly upon. Likewise, when I perceive the body of the other, it is perceived as similar to my own, and thus as the same kind of thing. My body is known to me as a psychophysical unity, and it is against this awareness that the meaning of the body of the other that I perceive becomes constituted: this is a body-like-mine, i.e., a psychophysical unity. As readers may recall from the above section on empathy, it is precisely the fact that the psychical aspects of the other can never be given primordially but only appresented, whereas my own can be made present to me primordially, that truly constitutes the other, given in this way through pairing and transfer of sense as an "alter ego," as distinct from myself.

In what does the relevant similarity consist, that motivates pairing and gets empathy off the ground? It helps to understand the mode in which empathy, according to Husserl, can be verified. Indeed, this is also an important development of the concept relative to Jaspers, who operationalizes empathy without explicitly spelling out how we could ever evidence the claim

that our empathetic findings are accurate, or true. In ordinary perception, verification comes in the form of making-present. Yet, the subjective side of the other can never be made present in this sense. According to Husserl, empathy has its own form of verification. I quote at length:

> Every experience points to further experiences that would fulfil and verify the appresented horizons, which include, in the form of non-intuitive anticipations, potentially verifiable syntheses of harmonious further experience. Regarding experience of someone else, it is clear that its fulfillingly verifying continuation can ensue *only by means of new appresentations that proceed in a synthetically harmonious fashion*, and only by virtue of the manner in which *these appresentations owe their existence-value to their motivational connection with the* changing *presentations proper, within my ownness*, that continually appertain to them.
> (Husserl, 1931, p. 114)

And,

> The experienced animate organism of another continues to prove itself as actually an animate organism solely in its changing but incessantly *harmonious "behavior"*. Such *harmonious* behavior (as having a physical side that indicates something psychic appresentatively) must present itself fulfillingly in original experience, and do so throughout the continuous change in behavior from phase to phase. The organism becomes experienced as a pseudo-organism, precisely if there is something discordant about its behavior. (Husserl, 1931, p. 114)

And finally,

> Whatever can become presented, and evidently verified, *originally* – is something *I* am; or else it belongs to me as peculiarly my own. Whatever, by virtue thereof, is experienced in that founded manner which characterizes a primordially unfulfillable experience – an experience that does not give something itself originally but that consistently verifies something

54

indicated – is "other". *It is therefore conceivable only as an analogue of something included in my peculiar ownness* [italics mine]. Because of its sense-constitution, it occurs necessary as an "*intentional modification*" of that Ego of mine which is the first to be Objectivated, or as an intentional modification of my primordial "world": the Other as phenomenologically a "modification" of myself. (Husserl, 1931, p. 115)

The way that we verify our empathetic intentions is through the perception of harmonious behavior, harmony being understood as bearing between a motivational context, on the one hand, and the totality of one's expressive behavior, on the other, such that one's actions are interpretable by relation to the motivational structure and to other action. Insofar as perceived behavior is apparently either self-inconsistent or inconsistent with a given motivational structure, then the behavior is seen as "pseudo" behavior; we suspect that what we perceive is not the action of an other, that there is no other here "in person." Chatbots, or artificial intelligences that are designed to convincingly carry on a conversation with a human interlocutor, exemplify the point nicely. Perhaps the currently most popular form of the Turing Test places a human being into a discussion with either a chatbot or another human, and it is up to the first human to determine whether their interlocutor is a chatbot or not. As anyone who has interacted with a chatbot knows, they are often fairly convincing, up to a point. Then, inconsistencies begin to appear. We wonder "why would someone say that," or we notice that something said now is difficult to make sense of given something said previously. When their behavior becomes disharmonious – when we cannot understand the motivations behind their expressions, or cannot make sense of the behavior as a consistent whole – then we cease to see the chatbot as possibly human; we recognize that it is only a pseudo-consciousness, to adapt Husserl's term.

But the motivational structure that serves as a key referent to perceiving one's behavior empathetically and verifying those perceptions is never itself directly perceived: motivation is a subject-concept that can, when it comes to the other, only be appresented. The final block quote above is illuminating: the motivational structure of the other, indeed our perception of the other in general, is conceivable only in relation to myself. I need not import my own motivations or other subjective states directly to the other, as if mine and theirs were identical, but rather my own structures and states can be modified, to account for the distinction between us. I modify my own Ego, and, at bottom, my own primordial world, and it is insofar as the behavior of the other can be consistently made sense of in relation to this modification of myself that empathetic perception gains its verification. Similarity to myself, similarity to what is "peculiarly mine," born out as harmonious behavior, is the similarity that undergirds pairing. So, if I see someone perform actions that I can understand in terms of myself and modifications thereof, then I can see them as indicating someone, that is, as an ego *like* me *but not* me. On the other hand, if their behavior is disharmonious, that is, internally inconsistent on the basis of modifications of my own motivational structure, then I question whether they are indeed an other (they "become experienced as a pseudo-organism"). If it could be demonstrated that there is some fundamental similarity that pertains between us, some way to make *some* sense of the other in terms of myself, then the other may be accepted *as* an other, but also as otherwise beyond understanding.

And so, as with Jaspers, we can see that empathy is founded on the perception of expressive behavior and some relevant similarity – something experiential that is shared between myself and the other – and that where empathy is undermined, including by the denial of one of these conditions, there occurs a threat of incomprehensibility. Yet, it is precisely because of the structural analysis of the Husserlian phenomenologist, and because of the importation of my own

structures onto the other – so long as this allows me to make sense of their behavior – that incomprehensibility no longer threatens a purely negative method, i.e., a pure deficit model. Instead, the deficit posited can be interpreted within a framework of underlying similarities (e.g., we say the other is *different* in some way, but similar in other, perhaps more fundamental ways), and in this way we can understand the deficit systemically. Though a deficit is attributed to the other, or even perhaps just a difference between us,[30] this deficit may have positive implications for the remainder of the system, as when Ratcliffe, for example, takes the absence of the ordinary horizon of possibilities for people who suffer MDD to both imply and explain the reported, positive phenomenon of corporealization.

Practicality and Inclusion

As noted, many contemporary phenomenological psychopathologists prefer figures other than those emphasized thus far. They will often cite Jaspers, and perhaps refer to Husserl, but tend to explicitly anchor their work in Heidegger, Merleau-Ponty, or other recognized major figures. Though such thinkers undoubtedly owe much to Husserl and Stein (though the latter has been historically under-recognized), there are certain distinctions within their systems that make a difference for my analysis. Furthermore, as already mentioned, Husserl's *The Crisis of European Sciences and Transcendental Philosophy* complicates his earlier thought and presents a position influenced by and similar to Heidegger's. I here indicate and briefly comment on some of these positions, focusing on the ways they differ from what has been presented here so far, so

---

[30] Here, the Husserl/Stein account has an edge on the Jasperian. The difference between myself and the other is more explicitly relational; rather than proposing to understand the other explicitly by attributing to them some deficit, the former try to understand our relation to the other in terms of the possibility of comprehending them while accepting that "the man himself" is something above any comprehension. However, radical differences – that is, differences that are a matter of modifications that aren't simply, in Jaspers terms, a matter of degree or causality – would seemingly have to either be or appear as negation of the familiar.

that moving forward I might better do justice to Husserl's ever-shifting thought, and so that I can address as wide an audience as this project permits.

Whereas Husserl takes the ultimate basis of experience to be sensation, Heidegger argues that this basis is instead our practical engagement with the world: our primary relationship to the world is through practical activity structured by Dasein's being as care, and, accordingly, the fundamental form of consciousness (understood intentionally, as consciousness-of) is our circumspective awareness in and of the world. Even sensation, according to Heidegger, is already organized by our practical concern and activity; meaning is not just a matter of temporality and association but rather a matter of use and value. The first relation to the other is not empathy according to Heidegger, as empathy already presupposes being-in-the-world as structured by care; empathy is a form of perception, and all perception presupposes practical engagement. Instead, our relation to the other is conceived of as *Mitsein* [being-with], which is revealed as an essential feature of our existence, i.e., an "existentiale" of Dasein, through the fact that we exist in a world always already structured by and implying the existence of others. The tools we use, the projects we pursue, the values we accept, and even the perception of the world abstracted from use and reformulated as the objective-scientific world, are already conditioned by the other. We first, inauthentically, exist in the mode of *Das Man* [the they], meaning (in part) that we first exist by taking up the world of others. Our values are the values handed to us by others, our tools the ones they make, our concepts historically generated, our projects what others expect of us… We are engaged in the world *with* others; others engage us in our practices. The other is thus foundational for Dasein, though Dasein, according to Heidegger, can abstract away from this foundation and instead exist "authentically," accepting the ultimate responsibility of its values and actions for itself: always conditioned by the other, Dasein reflects and decides,

58

and realizes by its reflection that, within its conditioning, it was always Dasein that implicitly endorsed (or rejected) what it was given.

For Heidegger, our understanding of being is structured by care, our care is structured by the others with whom we are in community, and thus the community informs our understanding of being, and care informs our understanding of others. The relevant similarity which undergirds an understanding of the other is thus not that of harmonious behavior construed as making sense of perceptual similarities between myself and the other as presented to me by sensation. Rather, we first understand the other by sharing a common world, that is, a world already structured by care, which we engage with the other, and in which our practices already presuppose community. We understand the other by doing what they do, by sharing their values, by using their tools, by participating in their (or, precisely, *our*) world.

According to the *Cartesian Mediations,* once we abstract to our sphere of ownness, we discover a "primordial world," a mere *Nature*, pre-objective and consisting of that which is given by our senses. This world, through the syntheses that Husserl analyzes there (it is not practical here for me to present these considerations), takes on the sense of a universal world, a *Nature* that must be in principle accessible to all others, as a shared sensational world. For Heidegger, however, even this world – any perception whatsoever – presupposes a certain practical engagement, and practical engagement is already communal, structured as it is by the care and concern of others, *the they*. Thus, for Heidegger, the fundamental appearance of the world is shaped by our concern, which is conditioned by others; the world that appears is not necessarily a universal one. Insofar as two individuals do not share worlds, then the question arises as to how they can understand one another. Two modes suggest themselves: first, they can participate in each other's worlds, immersing oneself in the practices and concerns that organize the other's

59

world and very being; second, they can negate their own practices, in an authentic movement by which they reflect on their being and accept or reject that which they have thus far only implicitly, non-thematically, become.[31]

Contemporary Heideggerian approaches to psychopathology tend to focus on existentialia, modifying them as the Husserlian modifies "transcendental structures of subjectivity." Typically, the method involves an explication of some ordinary world (a "normal" or "home" world), a posited point of departure from that ordinary world – spelled out in terms of alterations to the ordinary "care" structure, or the world, or altered "mood" (a common translation for Heideggerian *Befindlichkeit*) – and theoretical correlates of that altered world. For instance, Ratcliffe explores altered mood and the ways that it affects the at-home-ness within a world for the subject of altered mood, or how it results in a changed orientation to the world such that ordinarily accepted beliefs or processes are interrupted, or even become non-sensical.

In *The Crisis*, Husserl adopts a similar position, endorsing the life-world as the foundation of intersubjective understanding.[32] Husserl returns to the life-world from which his *Meditations* had abstracted, and the similarity between my own subjectivity and behavior is transformed from a perceptual-analogical model to a thoroughly normative one. This is first seen in his genetic accounts of the objective-scientific and mathematical world-views. Through certain practices, thoroughly historical and communal, we constitute new modes of givenness of the world, coming to understand the world through specific methods and in specific ways. These

---

[31] This latter may be preferable where participation is not possible or desirable. But it is also problematically negative, in the same sense as is the epistemologically limited deficit method.

[32] It should be noted that, even at the time of the *Cartesian Meditations*, Husserl takes the life-world – the world of actual, concrete, naïve experience – to be irreducible. His reductions are mere abstractions and, as Jordan tells us, we do not *live* in the primordial world. Indeed, the "transcendental clue" from which he derives his thought on empathy and his answer to the challenge of solipsism is our ordinary, everyday, life-world experience of other people. The fundamental datum out of which his theory is developed, and in which his considerations must be anchored (that is, the relevant phenomenon or "thing itself" to which Husserl promises to return) is thus a life-world phenomenon. This is his warrant.

practices thus found such experience. When it comes to our understanding of others, Husserl calls into question some of the findings of the *Meditations*. We understand the other by participating with them in the co-constitution of the world. At its most basic level, the other is perceived as making sense of a world, as do I, and as she whose collaboration is a necessary condition of my experience of the world in its ordinary sense.

The Dilemma

Phenomenological insights are grounded on some form of intersubjectivity which structures experience according to certain conditions. For Husserl, our ordinary experience of the world is structured by the appearance of others, in the form of bodies "like mine": the body of the other appears as the same kind of thing as my own body, and through pairing and a transfer of sense, the other is apperceived (or appresented) as a subject of experience like me. This is necessary for the constitution of the world as objective, that is, for the constitution of the ordinary, every-day world that we live in, which serves as a point of departure for all phenomenological reflections, regardless of how solipsistic they might otherwise seem. Self-reflection presupposes empathetic relations to others. For Heidegger and *Crisis* Husserl, this relation to the other and to a common world comes to the foreground. It is our inclusion in a common world that grounds not only our own Being (as an understanding of being) and phenomenological self-reflection, but the intersubjective validity of our claims: insofar as our reflections are always situated within and structured by some world, it is that world that conditions understanding. For Heidegger, a world is ultimately structured by care – by our own but, primarily, by the care of others, the former being possible only within the latter – and first appears through practical concern, i.e., circumspection. Those with radically different practices than our own, with whom we do not share care or concern, actualize other worlds, and the

61

possibility of an interworld understanding depends on the possibility of inclusion of ourselves in the practices and care of the other, and vice versa.

Phenomenological claims are intersubjectively valid – they can be applied to others beyond myself – insofar as I am conditioned by others, that is, insofar as some relation to the other is constitutive of the phenomena upon which I reflect. The dilemma is as follows. If (1) our relation to the other, and our fundamental awareness thereof, is based on sensation, as it is in Husserl through the *Meditations*, then harmonious perception is foundational for an understanding of the other, but (a) can be disrupted in certain circumstances and (b) what counts as harmony is based on *my own* possibilities, which do not exhaust (or perhaps even touch) the possibilities of the other. If, on the other hand, (2) our primary relation to the other is one of practical co-engagement in a shared world, as it is for Heidegger and, in a modified form, *Crisis* Husserl, then (a) radically different praxis implies ununderstandability (to echo Jaspers' term), and (b) in principle practice can interrupt understanding, conditioning incomprehensibility, i.e., ununderstanding.

What allows us to know the other, on Husserl's account, is sensation and similarity – perception and resemblance. When we perceive the other, through the perception of their bodily expression, we perceive them as *like me but not me*: a body *like mine*, expression *like mine*, subjectivity *like me*, but distinct. Our empathetic intentions are ultimately grounded on not just a static perception, a similarity at a time, but on harmonious perception. The expressive behavior of the other must cohere within itself – we must be able to make sense of the behavior as a whole – and with an unseen motivational field ascertained by modifications of my own.[33] If we assume

---

[33] Of course, some one action may not cohere with others, but the final arbiter is motivation: if various motivations can be made sense of together, as of the transitions between them, then the actions may seem quite incongruous amongst themselves but nevertheless coherent in relation to the motivational field.

that similarity holds, and that motivations are shared, then empathy may be a fine way of understanding, even perceiving, the other. But in the case of disruption of harmonious behavior, or the perception thereof, we begin to doubt our understanding of the other and, if the disruption is radical enough, we doubt even that we are face to face with a real other, as opposed to a "pseudo-organism." In psychopathology, such disruptions are possible. Indeed, where we understand the other as exhibiting "disorganized movement or behavior,"[34] we have admitted to a kind of ununderstandability, a disruption of the ground of understanding, and we have understood the other in a strictly negative manner as *disorganized*, *not organized*. Furthermore, when we understand the other as delusional, hallucinatory, or dissociative, we understand the other as no longer a legitimate locus of sensation. Importantly, where we cannot find reason for one's actions or beliefs, that is, where we cannot discover a motivational context out of which to make sense of the other, philosophers and scientists alike (e.g., cognitive scientists, psychologists, psychiatrists) are quick to attribute irrationality: if I cannot find a reason, there must be no reason. This is a justifiable move if the other must be understood in terms of motivational elements personally familiar to me; what is foreign, in the sense that it cannot be understood in terms of modifications of myself, will be unintelligible.

The question of the validity of empathy comes down to the ground for the assumption that similarity holds and that motivation is shared. In the *Meditations* and related work, this assumption is founded solely in perception: I see a movement like my own, and so I see it as expressive and as corresponding to a familiar motivational field, even if modified. Later Husserl and Heidegger can be seen as articulating this ground to be a kind of world, e.g., the shared life-world in Husserl. But the very constitution of such a world itself depends on a fundamental

---

[34] This is one positive symptom of schizophrenia, for example, listed in the DSM-5 (American Psychiatric Association, 2013).

relation to the other. In the broadest sense of 'world,' we have Husserl's "universal" life-world, which is the familiar world in which we live. But this life-world is constituted as objective and intersubjective precisely due to our empathetic relations with others; without such a relation, we are left with a solipsistic world of personal perception; the world appears as inseparable from me, as dependent while also transcendent.[35] Through empathy, we come to understand others as like but distinct from me, and only *following this accomplishment* does the world takes on the sense of an intersubjectively shared world. Worldly validities – what can be directly evidenced in experience – depend on the community of all possible subjects, to be confirmed by harmonious perception. Yet, insofar as the grounds of empathy are undermined, the other is not even known as other; a failure of empathy is exclusion from the community of co-constitutors of the world, exclusion even from the class of 'actual' organisms (i.e., relegation to the "pseudo-"). Ordinarily, such a failure is uncommon and easily corrected. In the context of psychopathology, however, it can become endemic or even essential to certain diagnoses or symptomology. Disorganized or disrupted movement, irrationality, hallucination, delusion, or radically abnormal experience generally threaten the conditions of empathy and thus exclude the person to whom they are attributed from the intersubjective community of valid world co-constitution.[36] The attempt to

---

[35] In the sense that Bob Jordan speaks of Husserl's "immanent transcendencies": the world is *immanent* insofar as it is perceived by me and the body, given as actually present, and variable with my bodily movements, but transcendent insofar as any intention from the solipsistic perspective intends more than what is actually given in the form of a horizon of possibilities or potentialities, and, correspondingly, that "objects" perceived solipsistically have their own series of potentialities.

[36] In the 1923/24 lectures (Part 2, section 2, chapter 1, lecture 34), Husserl takes the possibility of insanity simply to be the empirically real possibility of a disruption of universal harmony: "And surely the possibility is open that a human being's harmonious perceptual stream could be transformed into a senseless jumble, into a swarm of appearances. But what does this mean other than that a human being, and ultimately every human being, could become *insane*?" (2019, p. 259). He goes on to argue that, in cases of insanity, "there could no longer be any talk of this world, of "existing things," of experienced or (in freely initiated acts of perception) experiencable things in general […], there also could no longer be any talk of my lived-body, and none, therefore, of animals and of human beings – and so none of human beings whose concordant unfolding experiences, experiences that constitute an *actual* world, I could invoke" (261, italics mine), and "then "world" becomes a title for infinitely many phantasy-possibilities which are, taken together, *empty*" (261, italics his). The world constituted for the insane person is not an *actual* world, it is disharmonious and senseless; it cannot be understood by empathy if empathy necessarily references a coherent motivational framework. In *The Origin of Geometry*

understand the other by modifications of certain elements of my own experience while importing the remainder of my field is groundless in such cases, since the possibility of experiencing the other as relevantly "like-me" has been undercut.

In the Heideggerian case the problem becomes more obvious: if someone does not participate in the same practices as I do, if their actions are not organized by a familiar form of care (essentially taking the place of Husserl's "motivational framework," but at what Heidegger takes to be a lower level), then they will not be understandable to me. This could be remedied, for Heidegger, by participating in the world of the other, but this is not always possible. In psychopathology (though much of the following holds across cultural difference), we often struggle to understand what some experience is actually like; much of our literature attempts, in a variety of ways, simply to describe the phenomena in question. Analogy is often employed, as may be predicted given the methodological and theoretical foundations explicated so far, but this is often unsatisfactory. Importantly, the psychopathologist typically cannot simply undergo a similar phenomenon as that which we hope to understand.

Furthermore, if our understanding of others is conditioned by practice, there is no reason to suppose that certain practices could not undermine understanding. As I elaborate in the conclusion that immediately follows, some such practices can be seen to occur within psychiatric contexts.

Conclusion and Implications

Karl Jaspers, often credited with establishing the field of psychopathology, takes some form of "empathy" to be crucial to its method. For Jaspers, as we also see in Husserl, empathy depends on perception of expressive behavior and a form of personal familiarity with the

---

(1970, pp. 358-9), he echoes this sentiment at a different level, the level of language, where children and the psychologically abnormal are excluded from normal communicative society.

relevant phenomena. His theory of empathy, while somewhat difficult to pin down, is distinct from the "projective" theory of Theodor Lipps prominent at the time and taken up by current-day simulationism. When the conditions of empathy are undermined, according to Jaspers, our only recourse to understanding the other is through a deficit method, by which we strive only to say what some relevant phenomenon *is not*. This is inadequate for understanding what the experience of the other actually is, or is like, and has consequences for even a causal investigation into it; we have not determined the phenomenon to be explained. Accordingly, contemporary psychopathologists – whether they explicitly move past Jaspers or not – do not typically content themselves with deficit models. I have argued that transcendental phenomenology, particularly in Husserlian and Heideggerian forms, allows for the negative signification of deficit models to take on a positive significance: insofar as we presuppose a shared form of consciousness, that is, we presuppose certain structures of consciousness or existentialia, we can then trace the remainder of the system in light of a posited deficit, through imagination as well as transcendental argument.

Husserlian Phenomenology, and phenomenological psychopathology in this positive, structuralist sense, depends for its intersubjective validity on the discovery of "eide," or structural invariants of a given phenomenon. Eide are discovered through the method of eidetic reduction, which leverages a free variation of possibilities: a factual phenomenon is taken as an exemplar and initial evidence, and aspects of the phenomenon are imaginatively varied until the phenomenon no longer counts as the same kind of thing.[37] This process thus discloses what is

---

[37] The method of eidetic reduction is detailed by Husserl and Heidegger in the Encyclopedia Britannica article *Phenomenology* (1997). It involves four crucial features, but imaginative variation is the only one directly relevant to my considerations here; I do not currently challenge, at least explicitly, and believe I can indeed accommodate all other features (that is, they cause me no trouble).

essential to the phenomenon – its form, so to say, or sense; it articulates its "internal horizon of meaning."

Conceivable possibilities (i.e., those that can be imagined in the variation) are constrained by what can be given in original intuition, that is, by what can be given to oneself as self-evident and present. Original intuition is thus dependent on one's own experiences and capacities, and also one's community. One's communal being is an accomplishment which depends on understanding certain phenomena as indicating or partially constituting (depending on where in Husserl's work one looks) an Other ego like but distinct from me: an "alter ego." One knows the other through an act of empathy according to which expressive behavior is interpreted in relation to my own movements and psychical experiences, but distinguished from my own possibilities both in virtue of being given to me as simultaneous with my actual experience (which would be an impossibility for one and the same subject, to experience this-here and that-there, both *now*) yet as not original: the other is appresented, not presented. It is through this empathetic accomplishment that community is established, as a community of subjects of which I am simply one unique case. The encounter with the other, within the community, allows for me to conceive of new possibilities as modifications of my own experience that would be impossible were I a solipsistic subject, as well as by constituting new senses for already familiar experience, e.g., that of the world as shared.

Confirmation of these possibilities depends on the harmonious perception of expressive behavior of others, an understanding of which already depends on similarity to my own experience; it must be articulable as a modification of my own experience, and in particular must cohere with a motivational framework modified from my own. Confirmation, or fulfillment, of our intentions directed towards foreign experience, whether theoretical, perceptual, or

67

imaginative, depends on empathy as the act through which the other is given. If empathy is undermined, however, then confirmation of such intentions would be impossible: empathy is the perceptual act through which the other is appresented and is thus necessary for the harmonious perception thereof. Empathy depends on the perception of expressive behavior, which consists of a straightforward perception of bodily movement to which a sense of psychical activity is transferred through its being paired with my own bodily movements and their inseparable psychic correlates, and on a certain familiarity: without engaging in similar bodily practices, or without having undergone similar psychic phenomena, the 'perception' of the other is reduced to a kind of imagination or even projection. Without this familiarity, there is no tethering, so to speak, of an understanding of the other to modifications of oneself. And, with empathy being undermined, there falls the possibility of confirmation of related intentions.

If eide can only be discovered and confirmed by variations on one's own possibilities, and these variations and possibilities are constrained by one's capacities and community, and empathy can be undermined through radical differences in the practices or experiences of two people, then eide can only factor in familiar possibilities and cannot be adequately applied to unfamiliar subjectivity. The loophole is that eide hold for any *conceivable* phenomenon of such and such a type, but conceivability is ultimately constrained by the limits of empathy, due to the way that empathy conditions imagination (it constitutes new senses of phenomena and makes conceivable modifications of myself only possible if there are (and I am aware of) other subjects *like but not identical to* me). Any importation of transcendental structures – structural invariants of consciousness, the eide of experience – to what cannot be understood by reference to modifications of oneself is thus illicit. Precisely, it is illicit when empathy is undermined or

otherwise not possible. What cannot be conceived of in terms of the familiar and its conceivable modifications is thus, for the method, inconceivable.

Empathy is undermined in the psychiatric context, and in the broader case of "abnormality" in general. If some phenomenon is abnormal, it is essentially not familiar: it is not part of those ordinary practices and experiences that constitute a community, and therefore afford conception and confirmation.

It could be objected that the fundamental similarity supposed to hold in the case of empathy is that both persons are sources of animation and loci of sensation – we are both organisms – and that this cannot be undermined. But this sense of the other as an organism, even of myself as an organism, must first be constituted, and itself depends on a perceived similarity, pairing, and the correspondent transfer of sense.[38] In the first place, the perception of similarity is fraught with difficulties. Our own movements are not given to us in the same way as are the movements of others, not even considered as external objects in motion given to visual sensation. From our own perspective, our body does not even look like the body of the other. Something like a mirror would be necessary to see my movements as similar to the movements of the others, but would not the identification of my own body in a mirror involve many of the same difficulties?[39] Furthermore, pairing between my body and the body of the other is not just verified but instituted and sustained by harmonious perception: what separates the other's movements from mere motion is pairing with my own movements, which only holds as long as

---

[38] Alfred Schutz (1970, pp. 62-4) raises a similar objection, calling into question the perceived similarity between my body and the body of the other. While my point here is not entirely original, I do emphasize "outer perception" whereas Schutz emphasizes "inner perception." That is, he leverages a distinction between Leib and Koerpor, whereas I am pointing specifically to the way that my own body as object is, without presupposing a relation to the other, not given as similar to the body of the other.

[39] This problem could potentially be resolved by evoking a notion of convergence of action, the object perceived, and volition. That is, my reflection moves along with my volition, and with my own body both perceived (without the mirror) and kinesthetically.

the movements are coherently interpretable by reference to modifications of my own motivational framework (perhaps, minimally, a motivational framework in general) and its connections with those movements of mine meant to correspond to the other. What separates the other as a locus of sensation from mere blind mechanism is, again, perceiving them in such a way that I cannot make sense of their actions without an imposition of motivation, derived from what is familiar. In any case, this can all be undermined in cases of "disorganized behavior," and also in light of recent rejections of "basic emotions," "basic expression," and any sort of universal human practices.[40] In such cases, the motions in question cannot be confirmed by harmonious perception, and theoretical considerations usurp empathy, sacrificing their phenomenological basis of verification.

Furthermore, even if movements are perceived as similar – which is already, phenomenologically speaking, an accomplishment – they may not correspond at all to similar psychical states. Many persons with abnormal psychical states, and many people in general given their circumstances, adopt expressive movements that do not map their psychical states, for a variety of reasons. We can also invoke a deeper Wittgensteinian point, recalling his example (2009) of beetles in a box: we can see the box but we can't look into it to check. The only confirmation available is, to repeat, harmonious perception, but it seems that such a perception underdetermines the psychical phenomena. Moving forward, I will suggest that the "determining" factor is language, conceived of as discourse or conversation, and distinct from any Wittgensteinian or Husserlian formulation thereof.

---

[40] Disorganized behavior, as already cited, is a diagnostic symptom of the DSM-5; for a clear and convincing rejection of so called "basic emotions," see Giovanna Colombetti's *The Feeling Body* (2013). Regarding basic expressions, the literature broadly defines basic expressions as expressions of basic emotions, and so Colombetti's criticism of the latter carries over to the former.

Another point to be considered is that the other, conceived of as a locus of animation and sensation like but distinct from me, already supposes – somewhat covertly in Husserl's *Meditations* but explicitly in the *Crisis* – that the other and I already share a world, that we both act in and make sense of a world conceived of as the most fundamental stratum of experience and that the foundational datum for all of our considerations is that we do so *together*. However, in diagnoses of hallucination, delusion, or irrationality, the validity with which the other co-constitutes our shared world is denied: they do not constitute validities *for us*, that is, what can be given to original intuition, and cannot be understood through what Husserl calls world-validities. Ultimately, it is through the world as something shared – as something itself intersubjectively constituted – that we can first understand both ourselves and the other in the ways that we do.[41]

My last special concern here is that the psychiatric context itself can foster a kind of invalidation. The power dynamic between therapist or psychiatrist and patient, or scientist and subject, in which the former figure strives to listen to and responsibly observe the latter but ultimately subjects the latter to diagnoses and models (not to imply that doing so is ill-intended, or that it is not often, even usually, benevolent) is one in which the familiar has the final word over the unfamiliar, and where confirmation of the former's hypotheses can only proceed, and in a sense exclusively, from what is already known: from the possibilities afforded to the former by

---

[41] This leads Husserl to a paradox. In *The Crisis* (1970, p. 179), he puts the point as follows: "But precisely here lies the difficulty. Universal intersubjectivity, into which all objectivity, everything that exists at all, is resolved, can obviously be nothing other than mankind; and the latter is undeniably a component part of the world. How can a component part of the world, its human subjectivity, constitute the whole world, namely, constitute it as its intentional formation, one which has always already become what it is and continues to develop, formed by the universal interconnection of intentionally accomplished subjectivity, while the latter, the subjects accomplishing in cooperation, are themselves only a partial formation within the total accomplishment?" He attempts to resolve the difficulty through an account of functioning subjectivity, and specifically, functioning or transcendental intersubjectivity, for whom the world is present or co-present. I am indifferent as to whether he succeeds in avoiding a kind of circularity but note that in the pathological cases here under consideration, co-presence is jeopardized.

their own experiences, by the normal community, and by imaginative variations (constrained as they are!) thereof.

The critique has a number of implications. First, empathy is a legitimate method insofar as it applies to what can be made familiar. There are a number of methods for making the unfamiliar familiar. Jaspers suggests that the psychiatrist studies poetry, history, and literature. There are also technological methods of familiarizing oneself with otherwise foreign phenomena, such as augmented reality devices meant to, for example, disrupt ordinary visual perception in a way suspected to mimic autism (e.g., Qin et al., 2014; Mikropoulos et al., 2020), we can produce a number of illusions, such as the rubber hand illusion, that may be relevant for understanding pathology phenomena. But in all such cases, how can we know that the method has gotten it right? I suspect that, for the phenomenologist considered so far, the answer will be harmonious perception – does the new experience allow us to make sense, motivationally, of what we can observe – the expressive behavior – of the other? This runs into a number of difficulties already discussed, and I again suggest, as I will attempt to show in the next chapter, that, for a true "confirmation," communication is the ultimate recourse.

The difficulties exposed in this chapter rest on a theory of intersubjectivity that is grounded in a certain structuralist conception of expression: expressive behavior, and properly speaking perceptually similar bodily movement, is the ground of the pairing between myself and the other, and binds us together in a kind of quasi-identification (what Husserl says is an analogue but not "in the usual sense"). So tethered, the other can only be understood in relation to oneself; radical difference, i.e., differences that cannot be articulated in terms of my own imagined possibilities (however imaginative they may be, and however common this kind of difference may turn out to be), cannot be accommodated. A theory of expression that can

accommodate difference is required for an adequate understanding of abnormality, even if, according to such a theory, our understanding will always be – in some way – inadequate; what we need is a theory that reconceives of our fundamental relationship to the other and in doing so does justice to the phenomena and, more importantly, as I will proceed to demonstrate, does justice to the other.

Chapter 2

Expression: The Interruption of
Sensibility and Freedom by Demand

Introduction

Expression is a central concept for phenomenological and enactivist approaches to the

problem of intersubjectivity: it is taken as the foundation for empathy and various forms of direct

(social) perception, and performs crucial functions for embodied, intersubjective co-regulation as

specified by dynamical systems theory approaches. The sense of the term, 'expression', is not

univocal, but typically includes some claim about the relationship between mind and body (and,

specifically, bodily action), and is more and more considered in its socio-cultural dimensions and

conditioning.

In this chapter, I consider two concepts of expression: the first, that of Edmund Husserl,

and the second, that of his student Emmanuel Levinas. The Husserlian notion is an early and still

influential one within phenomenology and applications of phenomenology to fields such as

cognitive science and psychopathology. The Levinasian notion, on the other hand, has been

largely overlooked. I explicate both notions of expression, and demonstrate that the latter

provides unique insight into the nature of subjectivity and the interpersonal relationship.

Specifically, the Levinasian notion allows us to address many of the same problems as

Husserlian expression, but also allows us to recognize the distinctly – and thoroughly – ethical

character of intersubjectivity.[1]

For each concept of expression, I provide the conditions out of which expressive

phenomena arise, indicate some problems that the concept is meant to solve in each thinker's

---

[1] The term 'intersubjectivity' is not found in Levinas, though it is a term quite familiar to Husserlians. I adopt the term since it is the one used by my target audience, who know Husserl better than Levinas, to describe a relationship between one (e.g., myself) and the Other. I use the term in a broad sense as referring to any such relation, though I must recognize that doing so is not unproblematic.

system of thought, specify the concept, and indicate the solutions it affords. I then compare the accounts and consider some implications, objections and limitations. My central claim is that Levinas' concept of expression affords us an articulation of the ethical dimension of interpersonal experience that allows us to conceptualize the relationship between ethics and objectivity and can more readily serve as a foundation for phenomenological ethics than its Husserlian counterpart. This is fruitful not only for phenomenology but also its interdisciplinary applications, such as phenomenological cognitive science and psychopathology.

Husserl

In his *Cartesian Meditations*, Husserl takes up the project of establishing a basis for the sense of objectivity found in experience as well as a foundation for a corresponding objective knowledge. The possibility of objective knowledge depends on overcoming 'the objection that phenomenology entails solipsism' (Husserl 1960, p. 89). In other words, the objection claims that the phenomenologist cannot countenance the existence and sense of other consciousness and, since phenomenology demonstrates that what appears is in the first instance inseparable from its appearance, it follows that whatever appears to consciousness is incapable of being constituted as transcending the individual subject to which it appears.[2] Husserl explicitly takes up a 'solipsistic reduction' (Husserl 1960, p. 155), or a 'reduction to the sphere of one's ownness' (Husserl 1960, p. 92), in order to disclose the resources available to the subject for constituting for itself the sense and existence of others, and there and elsewhere he demonstrates the foundational role that the consciousness of others plays for the constitution of objective nature.

---

[2] Riceour and Schutz both emphasize the importance of the *existence* of others for Husserl's project, and use this interpretation to found their critiques thereof. Others, such as Carr (1973; 1974) and Costelloe (1998), argue that Husserl's problem in the Cartesian meditations is not a 'straightforward' 'problem of solipsism,' meaning that Husserl is not concerned with establishing the *existence* of others but only with the constitution of the *sense*: 'other.' They thus offer refutations of Riceour and Schutz that perhaps apply to my interpretation here. I consider this latter interpretation below.

Ultimately, the foundation of our consciousness of the other and, thus, the constitution of objectivity, depends on what Husserl conceives of as expression.

Husserlian phenomenology can be characterized as an investigation into one's own experience in order to discover universal forms of any possible experience. In this way, phenomenology is an egology: it explicates the structures and features of the subject of experience. But it also strives to move beyond this individual subject in order to found a broader knowledge of both subjectivity in general, whether my own or that of another, as well as that which subjectivity constitutes for itself, e.g., things, the world, and objective nature. At this stage, Husserl's *Cartesian Meditations* earn their name: phenomenology's foundation is the cogito, but the goal is to discover within the cogito – by initiating a kind of investigation into the nature of the cogito missed by Descartes – a foundation for all other knowledge (Husserl 1960, pp. 155-7). Though Husserl admits that the reflecting philosopher is never really alone, the reduction of all appearance to phenomena appearing in some intentional mode or other – the bracketing of the actuality of that which appears in favor of an intentional analysis for which constitution and correlation become the primary functions – requires philosophical reflection to take a subject hypostasized by itself as its starting point. Husserl begins with the hypothesis of solipsism to demonstrate the inadequacy of solipsism, not proceeding by a kind of proof by contradiction but rather by demonstrating that solipsistic experience contains within itself that which is necessary for its own transcendence.

At the heart of his response to the objection of solipsism is Husserl's concept of expression. In our own lives, our bodily and psychic states and actions are separable only in abstraction: they are lived as unified, and only in reflection can we describe this unity as psycho-physical. Specifically, our bodily movements *express* our subjectivity. Through a process called

'bodily coupling' there occurs a 'transfer of sense' from experience of one's own movements to the perceived movements of the other. Initially, the bodily movements of another (at first a spatio-temporal thing) are recognized as similar to my own. There thus occurs a synthetic achievement by which my movements and the other's are constituted as belonging to the same type. Next, as my bodily movements belong to the type '*expression*,' for which the essential characteristic is psycho-physical unity, the bodily movements of the other become constituted as one aspect (physical) of a psychophysical unity. Like my own movements, they take on the sense of expressing subjectivity. However, since I live through my own psychic life in union with my bodily actions but, on the other hand, only directly perceive the bodily action of the other and *not* their subjectivity, the subjectivity of the other thus expressed belongs to a distinct subject: not to the ego, but to an *alter-ego* (Husserl 1960, pp. 90, 94, 100). This intentional act, founded on expression and, accordingly, on bodily coupling and transfer of sense, by which the other is given as an alter-ego to the ego, based only on what appears immediately for the ego (structures of its own subjectivity and spatio-temporal movement), Husserl calls "empathy."

It is not that the movements of the other are given as formal, empty expressions of subjectivity: the similarities that found empathy pertain to the specific content of subjective experience. The movements of the other are made sense of in terms of one's own, contentful experience to the extent that similarity pertains between them. Perception of the other's movements thus takes on a more or less determinate character, as perception of <this> or <that>; movement *as* <this>, movement *as* <that>. That is, not only can we perceive that the other expresses subjectivity but also (according to Husserl as well as many contemporary phenomenologists, especially in the context of philosophical cognitive science and psychopathology) the particular states that the other undergoes. A smile, as resembling my own

smile, and in familiar circumstances, has a specific, familiar meaning: I see the other as happy, or

insincere, mischievous, etc.[3]

Through empathy, the ego constitutes the other, as alter-ego, for itself. The things of the

world which had previously appeared to consciousness now take on the sense of being given not

only to one's own possible intentional states, but also possibly to the other, or even to the other

and not me. That is, the constitution of the other founds a sense that the world is not exhausted

by me but, rather, is shared. The things of the world take on a new sense of independence,

becoming 'objects,' and the world gains its sense of objectivity.

Levinas

Levinas disagrees that the ego contains within itself the resources necessary for

overcoming solipsism, which he often calls "separation" or "atheism." Yet, in *Totality and*

Infinity, he performs an abstraction analogous to Husserl's solipsistic reduction: he considers the

human being in the position of solipsism, or 'alone,' and offers an account of the way in which

she can escape such a position. As he notes, "In the separated being," which I have so far

referred to as solipsistic, "the door to the outside must be at the same time open and closed"

(Levinas, 1969, p. 148).[4] Much of what proceeds from this abstraction can be understood as an

explication, first, of the closure of the separated being achieved by sensibility, and only then of

the openness revealed by the expression of the Other, which appears at the limit of the sensible.[5]

---

[3] For more on this point, see Husserl 1960, p. 119, where Husserl affirms that 'an appresentation of someone else continually furnishes new appresentational contents – that is to say, brings the changing contents of the other to definite notice'. Additionally, there is some debate in contemporary phenomenological and enactivist cognitive science as to whether, in addition to seeing *that* the other is happy (for example), we also see *the other's happiness*. Husserl appears to be of two minds on this question. It is not, in this context, crucial that I take a stance on this, but – against recent take-up of Levinas' thought in these fields, such as by Overgaard (2003) – Levinas does not seem obviously in agreement with a direct perception theory that maintains one can *see* the other's mental states, at least not in any unmediated, sensory way.

[4] Note how different an understanding of the separated being as having a *door* is from Husserl's well-known formulation of the subject as a monad that has "no doors" but rather "windows (which are acts of empathy)" (HUA XIV, 260). In Husserl, monads can experience one another, but do not welcome. For Levinas, the reverse is true.

In *Otherwise than Being*, this openness and closure are presented in a more holistic and dynamic way, but the relationship between these texts is complicated.

Sensibility

Levinas maintains that "sensibility enacts the very separation of being – separated and independent" (Levinas 1969, p. 138). Carrying out a kind of 'solipsistic reduction' that aims to consider the ego precisely in its separation from Others means, then, that sensibility is our starting point.[6] Levinas does not tend to provide his readers with clear, stable, individual definitions of that which he considers. Instead, he tends to define and redefine, to refer one concept to another and then, later, to still others, introducing complications at each turn. Sensibility is variously defined by Levinas, and always in reference to other central concepts. In *Totality and Infinity*, the central concepts for understanding sensibility include enjoyment, contentment, the body, life, egoism and hedonism. In section II, part B, subsection 4, titled "Sensibility," Levinas claims that sensibility is "the *mode* of enjoyment" (Levinas 1969, p. 135; italics in original), later describing it as "the instance of enjoyment" (Levinas 1969, p. 136), and even asserts, simply, that "Sensibility *is* enjoyment" (Levinas 1969, p. 136; italics mine). To make sense of these characterizations, we must first say something of enjoyment.[7]

---

[5] In English, "Autre" and "Autrui" are both translated as "other." Autre refers to whatever is not the ego but refers back to it, back to sensibility, hedonism, power, and a system constructed by the ego. Autrui refers to the *human* other. L'autre is not typically seen as "other enough" by Levinas and does not correspond to transcendence. Autrui, however, is the other that Levinas has in mind when he talks about the social or interpersonal relation. I here follow Lingis's convention (in *Totality and Infinity*) of translating "autre" as "other" with a lower-case "o," and Autrui as "Other" with a capital "O."

[6] This is a crucial aspect of the entire strategy of *Totality and Infinity*. Levinas writes that "Revelation is discourse; in order to welcome revelation a being apt for this role of interlocutor, a separated being, is required" (1969, p. 77). That is, without a separated being, transcendence is not possible. Or, were the subject all-inclusive, there could be no radical alterity. Or again, without finitude, there could be no idea of infinity, formulated as a thought that thinks more than itself, and produced in the subject's encounter with the Other. Since sensibility enacts separation, sensibility becomes the first major consideration of the separated being.

[7] Kas Saghafi has pointed out to me that "enjoyment" translates *jouir*, which, due to its sexual meaning, underscores and emphasizes the relation to the body.

According to Levinas, enjoyment – not intentionality (understood as consciousness-of, and of which Levinas is, perhaps in part unfairly, dismissive) nor practical circumspection – characterizes our fundamental relationship with the world. The sensitive being is the body (Levinas 1969, p. 136), and the body is simultaneously a closure and openness to the world: the body has needs, and, through sensibility, relates to the world that supports it, making possible the satisfaction of those needs. Enjoyment, however, cannot be understood merely in terms of a means-end practical relationship, where the things that we encounter in the world are given to us first as tools defined by a practical finality, that is, merely as things with which to satisfy our ends. Rather, Levinas claims, the sensitive being comes to take a stance towards this very process of satisfaction, making the satisfaction of needs into its own end, rather than only a means: we take pleasure in satisfying our needs, "we live from 'good soup'" (Levinas 1969, p. 110); "The need for food does not have existence as its goal, but food" (Levinas 1969, p. 134). That is, we do not only eat to live, and to assert that the meaning of eating derives first from a kind of means-end relation is to ignore the relationship the living, feeling body actually has to food, and the significance of eating for the living body.[8]

The way that we relate to the world through enjoyment is characterized precisely by sensibility. It is as a sensitive body that one both has needs and can relate itself to a world, *feeling* the world and the very relationship between them. Levinas writes:

The sensibility we are describing starting with enjoyment of the element does not belong to the order of thought but to that of sentiment, that is, the affectivity wherein the egoism of the I pulsates. One does not know, one lives sensible qualities: the green of these leaves, the red of this sunset. (Levinas 1969, p. 135)

---

[8] Indeed, for Levinas, "Food can be interpreted as an implement only in a world of exploitation" (1969, p. 134).

Sensibility is thus a matter of the affective body, both in its capacity to be affected by and to affect worldly things. Rather than knowing an object, one lives sensible qualities, and it is the latter, according to Levinas, that ultimately grounds the former. Adriaan Peperzak characterizes Levinas' sensibility accordingly, demonstrating its divergence from 'traditional conceptions of the senses':

> *Sensibility* is the name for that dimension of the human subject thanks to which it is able to have this pleasurable commerce with the elements (108-14/134-40). Against the traditional conception of the senses as the means through which we are able to know things, Levinas shows that the basic sensibility, in which all other intentions and relations are rooted, is an affective commerce with the elements, a naïve and spontaneous feeling at home in a world that has not yet taken the form of an order of things, objects, instruments, and rational relationships. (Peperzak 1993, p. 156)

The language of *Otherwise than Being* emphasizes sensibility as "vulnerability, exposure to outrage, to wounding" (Levinas 1981, p. 15), which allows us to make sense of certain claims in *Totality and Infinity* that seem to imply, despite Levinas' occasional equation of enjoyment with happiness, that sensibility does not refer solely to happiness but to an entire affectivity: not only happiness but also despair and suffering, as "a failing of happiness" (Levinas 1969, p. 115). To sense is to relate oneself to something that "delights or saddens [life]" (Levinas 1969, p. 112). The body is that through which, in virtue of sensibility, the ego becomes constituted as a being separated from but dwelling within a world which matters to it, in which it takes a stance towards its own activity not only in terms of finality but in terms of happiness and suffering, joy and sorrow, pleasure and pain. It is in these terms, on a Levinasian picture, that the ego is constituted by sensibility, as it is in these terms that the embodied being relates to the world and to itself.

Furthermore, the egoism of the separated being is a hedonism. The human being, as a sensible creature, discovers the world through its enjoyment of the world, that is, its vulnerability to a world that fills its life with happiness or misery. It is in these terms that the world becomes meaningful, and so we can see that hedonism is both a moral and ontological term, describing a certain motivational character of sensibility as well as characterizing the manner in which the world is disclosed.[9] Sensibility is in part a motivational basis that conditions the very being of the ego. Peperzak, as one of few commentators who provides an in-depth consideration of this hedonism, is again useful:

> This self-identity of Me is more than a logical tautology; it is the concrete activity of self-identification through which I establish myself as inhabitant and owner of my world. The concrete way of my being what I am – in the supposition that we can make an abstraction from all encounters with other people – is the egoism of my enjoying, ruling, and transforming the world according with my needs.[10] (Peperzak 1993, p. 136)

This hedonism is a first, inadequate morality at the heart of the ego's very existence. Hedonism is not inadequate for being in some way "bad" or owing to some badness, , since it is a condition of the ego's existence, as built into the very sensibility that constitutes the ego. Hedonism, when considered from the perspective of separation, is not yet a transgression against any Other, "Not against the Others," "but entirely deaf to the Other" (Levinas 1969, p. 134).[11] At this level of

---

[9] Levinas often has Heideggerian philosophy in mind when he writes of ontology, though he seems to also lump Hegel into this category, as well as many specific claims made by Husserl.

[10] Peperzak also ties this hedonism to the central project of *Totality and Infinity*, specified in footnote 6 above: "The enjoyment of a corporeal and terrestrial existence is constitutive for any ego: the I establishes itself as a self through the absorption of elements, things, or events or by submitting them to the I's domination and possession. Without this appropriating and hedonic egocentrism, there would be no relationship to other persons because this relation presupposes a basic level of individual independence, even if further analyses will show the relativity of this independence" (1993, pp. 24-5).

abstraction, hedonism is unavoidable. Peperzak writes that "In this still solitary dimension, the law of life is: Enjoy life and enjoy the earth as much as possible" (Peperzak 1993, p. 36).

To say that hedonism is, in the ego's separation, unavoidable, is to invoke what Levinas refers to as sensibility's *contentment*. Contentment does not here mean that all one's bodily needs are satisfied. Rather, in the context of this dissertation, we read contentment as a claim that the sensible being (contra Husserl) does *not* contain in itself the resources necessary for transcending solipsism: the sensible being does not – cannot – constitute the Other for itself. As such, the world of sensibility is inherently a world of hedonism, disclosed in terms of enjoyment. Levinas formulates "the permanent truth of hedonist moralities" precisely in terms of the contentment of egoism:

…to not seek behind the satisfaction of need, an order relative to which alone satisfaction would acquire a value; to take satisfaction, which is the very meaning of pleasure, as a term. […] need is naïve. In enjoyment I am absolutely for myself. Egoist without reference to the Other, I am alone without solitude, innocently egoist and alone. (Levinas 1969, p. 134)

Sensibility thus constitutes a separated and hedonistic being, i.e., the ego, whose very way of being is structured by a content enjoyment, an enjoyment that does not search beneath itself, nor that which it enjoys, for its very condition. As content, the ego knows no Other: all that it encounters it encounters as for itself, to be appropriated as part of "the same". At this stage, there is a certain resemblance – though extremely limited – between the condition of the separated being in Levinas and the solipsistic ego in Husserl: without the Other, the world as present to the ego does not transcend the ego itself; all presence refers back to the ego. The major difference, of

---

[11] Since Levinas takes expression to be speech, he also speaks of separation as involving 'deafness.' I am aware that this term might be considered an ableist slur, especially here where ethics itself depends on 'hearing' the other. This term does not here refer to literal deafness, of course, but I recognize that it remains problematic.

course, is that for Husserl this is a matter of the world presenting itself to actual and possible intentional acts: the world does not transcend the ego's own *intentionality*. But for Levinas, the world is first present in terms of enjoyment, which is not understood in terms of an intentionality in which an act and an object become correlated, but rather in terms of the body as an exposure to the world, and the very living through of the world by that body. Indeed, Levinas maintains that intentionality, thought, cognition, representation, truth and even meaning rest on the prior condition of sensibility as enjoyment.

Similarly, a problem of objectivity arises: if the world does not transcend the ego, then experience of the world is a subjectivism. The problem is no longer merely epistemological, however, but thoroughly ethical: at its rock-bottom foundation of sensibility, the ego already adheres to a specific morality, i.e., hedonism. Anticipating a solution to the problem, the Other not only founds objectivity in Levinas, but also reveals the necessary inadequacy of hedonist morality. Indeed, the solution, too, partially resembles Husserl's own, in which the constitution of the Other reveals to me a perspective similar to but distinct from my own, to which the world may also appear. Peperzak notes that "Objects are born when I place things in the perspective of other persons…" (Peperzak 1993, p. 165). For Levinas, however, the ego cannot constitute the Other *for itself*, and the Other is not constituted on the basis of a perceived similarity to myself, or even at the level of perception at all. Rather, the approach of the Other is a matter of the disruption of sensibility, of calling the ego into question. The previous quote continues… "…*Detached from their hedonistic and egocentric function*, those things receive an intersubjective meaning and existence" (Peperzak 1993, p. 165; italics mine). We must now ask, as we did for Husserl, for the nature of our relationship to the Other such as to found objectivity. But now, this question bears with it a distinctly ethical significance. As in Husserl, the answer to

this question – what is our fundamental relation to the Other, on which objectivity and, for

Levinas, a refutation of egoistic morality, depend? – begins with expression.

Expression

Levinas offers a more radical notion of expression than does Husserl, for whom

expression refers to a kind of behavioristic sign: an intuitive datum interpreted in relation to

one's own familiar horizons of experience, gaining its sense as signifying a subjectivity similar

to but distinct from my own through the achievement of synthetic acts in which meaning is

transferred from one(-self) to the other. The other, as revealed to empathy, on the Husserlian

picture, is constituted by the ego and for the ego as a modification of the ego, that is, as alter-ego,

"conceivable only as an analogue of something included in my peculiar ownness' (Husserl 1960,

p. 115). The Levinasian concept does not follow this trajectory. While Levinas could also accept

Husserl's claim that the Other is never given originally (Husserl 1960, pp. 114-5), he insists that

any conceptualization of the Other, insofar as all concepts – all thought – refer back to myself, is

inherently inadequate. While skeptical of any 'appresentation' of the Other in the Husserlian

sense, and maintaining that the Other is she who precisely does *not* appear in experience (even or

especially as an 'alter-ego' correlated with my own through a unity of our expressive acts in

understanding), Levinas speaks of a distinct kind of *presence* of the Other:

> The presence of the Other, or expression, source of all signification, is not contemplated as an
>
> intelligible essence, but is heard as language, and thereby is effectuated exteriorly.
>
> Expression, or the face, overflows images, which are always immanent to my thought, as
>
> though they came from me. (Levinas 1969, p. 297)

This presence (which is perhaps equivalent to what is called 'the trace' in *Otherwise than Being*

(Levinas 1981, p. 12, and throughout)) is not the presence of an intuition given to perception or a

sign interpreted by an understanding always situated within its familiar horizons: the face, at times identified with expression, is *not* the empirical face accessible to vision, and, as Levinas states in his interviews with Phillippe Nemo,

> The best way of encountering the Other is not even to notice the color of his eyes! When one observes the color of the eyes one is not in the social relationship with the Other. The relation with the face can surely be dominated by perception, but what is specifically the face is what cannot be reduced to that. (Levinas 1985, p. 86)

Instead of perception and understanding, the relation to the Other is language or speech, which proposes, and thus precedes, a (common) world to us. We do not first perceive or understand the Other, but, rather, our first relation to the Other is discourse; it is on the basis of conversation that I may subsequently come to understand her.[12] Expression is a form of saying which is the condition of anything at all being said.

The face is that through which the Other expresses herself, and as such I begin with the face. First, we must clarify that the 'face', for Levinas, is not the empirical face, and does not strictly correspond to it. Rather, "face" is Levinas' term for the entire body, insofar as it is expressive: the face is the expressive body. Peperzak writes that "The word 'face' can be replaced by 'expression' or 'word' or 'speech' (la parole)" (Peperzak 1993, p. 142), and Barbara Jane Davy points out that "Expression is not merely verbal for Levinas, because the significance even of actual speech is not just words but meaning" (Davy 2007, p. 50). She goes on to cite a

---

[12] Levinas sometimes speaks of the relationship to the other (even substitution; see *Chapter 5* below) as "maternal," or describes the self in relation to the other as "the maternal psyche" or the psyche as "the maternal body"; embodiment in contact with the other, vulnerable to the other and bearing an asymmetric responsibility, is maternity (Levinas, 1981, pp. 67, 71, 75-9; throughout). Interestingly, Diane Perpich (2008) provides a broadly developmental reading of the way that the other (she emphasizes the father figure, rather than mother, but I believe the idea of maternity is here the appropriate one) brings us into the world. Richard Cohen (in Gantt, 2002) argues that maternity is key to understanding the psyche in its specifically psychological sense.

useful passage in which Levinas explicitly states that "the whole body – a hand or a curve of the shoulder – can express as the face" (Levinas 1969, p. 262).

In what way does the face "overwhelm images," or escape domination by perception? There is indeed an ambiguity at play. *Otherwise than Being* makes clear, and *Totality and Infinity* raises the question of, the fact that it is within sensibility that we encounter the body and expression of the Other. Yet, it is crucial for maintaining the alterity or exteriority of the Other, i.e., for arguing that the Other is irreducible to the ego's familiar domain, that she nevertheless escape this sensible manifestation. In arguing that Levinas maintains a kind of direct perception theory, according to which the Other is intuitively present to perception, Søren Overgaard cites Levinas' claim that "In expression the manifestation and the manifested coincide" (Overgaard, 2005, p. 268, citing Levinas 1969, p. 296), taking this as evidence that the Other is indeed, in a certain qualified sense, "extremely" or "personally" present in their expression. The relevant passage continues:

…the manifested attends its own manifestation and hence remains exterior to every image one would retain of it, presents itself in the sense that we say of someone that he presents himself by stating his name, which permits evoking him, even though he remains always the source of his own presence. (Levinas 1969, p. 296)

The presence of the Other in their manifestation is not as something perceived or apperceived; rather, it is as attendance to manifestation, source of presence, or "source of meaning" (Levinas 1969, p. 51; Overgaard, 2005, p. 269; Davy, 2007 makes a similar claim). The Other is not something expressed alongside their manifestation, as something coinciding with it *in* perception – they are not *perceived* together – but rather the Other is she who expresses herself. Attendance is not something perceived, but the fact that the Other can defend, rearticulate, double-down-on

or un-say her word, calling into question any interpretation of that word by the ego, "overflowing [any] idea a thought would carry away from it" (Levinas 1969, p. 51). The Other escapes their word, their manifestation, but coincides with it as she who speaks that very word. Reinforcing this point on manifestation and the distinct form of presence achieved by expression, Levinas writes that "expression does not manifest the presence of being by referring from the sign to the signified; it presents the signifier. The signifier, he who gives the sign, is not signified" (Levinas 1969, p. 182).

Though expression is equated with speech, it is not simply one sign among others, signifying the Other in the way apperception would present an alter-ego through the perception of a familiar gesture. In *Ethics and Infinity*, Levinas states that "In discourse I have always distinguished, in fact, between the *saying* and the *said*… the *saying* is the fact that before the face I do not simply remain there contemplating it, I respond to it" (Levinas 1985, p. 88). This distinction is a central theme in *Otherwise than Being*, and is perhaps the primary way that Levinas there approaches the topic of expression. Though every saying refers also to something said (and vice versa), it is *saying,* considered as irreducible to the said, that characterizes speech as expression.

William Simmons clearly formulates this distinction so as to highlight the irreducibility of the saying to the said, distinguishing Levinas' approach from "traditional theories of expression." He writes that "In the traditional view, language originates with the speaker. The speaker intends to speak, formulates thoughts into words, then expresses them. The ego is pre-eminent" (Simmons 1999, 88). Note that Husserl's view of expression in the *Logical Investigations*, *Analyses Concerning Passive and Active Synthesis*, *Ideas I*, and *Husserliana* XI, XIX, XXVI (references taken from Moran & Cohen, 2012, pp. 116-118) corresponds to this

traditional view: there, before he becomes primarily concerned with our experience of other people in the face of the objection by solipsism, expression refers to a sign that delivers a meaning already present for ("internal to") the ego. Simmons continues:

> Levinas, on the other hand, emphasizes the role of the addressee. The focus is thus shifted from the ego to the Other. 'The activity of speaking robs the subject of its central position; it is the depositing of a subject without refuge. The speaking subject is no longer by and for itself; it is for the other. (Simmons 1999, p. 88; citing Peperzak 1993, p. 221)

What is said refers to a sign produced alongside an act of expression, a *something* or *content* picked out by language. As Simmons and Levinas both emphasize, it is the said that carries the "as" structure characteristic of intentionality: it specifies 'a this as that.' Traditionally, this issues from the subject considered as speaker, and is an exercise of his power. However, any *said* presuppose *saying*, in which the Other approaches the ego, addressing them in expression. Simmons writes that "Before any speech, before any intention to speak, there is 'an exposure of the ego to the other, the non-indifference to another', which is not a simple 'intention to address a message" (Simmons 1999, p. 88; citing Levinas 1981, p. 48). It is in speech construed as *saying* that the Other opens onto the ego, exposing herself to question and response and, crucially, imposing herself on the ego by demanding response. It is risky to construe saying even as an *act* (that, for instance, delivers a said), since this risks covering up its ethical significance – that saying is a relation with an irreducible other, who affects me in my passivity by calling me to responsibility – with categorical/ontological terms.

Expression is not just a communication of content, but communication as such, discourse: addressing the ego as an interlocutor called to respond. It is this aspect of speech *as saying* which, furthermore, indicates the way the Other always escapes, eludes, or overflows what is

said, as well as perception; as saying, as approach, as expression and a call to response, the Other stands behind their word and at its defense. The Other, revealed to the ego through expression, is not captured by the ego in that expression, and is thus incessant. Expression, and the corresponding inadequacy of any idea that attempts to capture it as signifying the Other in something said, can be characterized by the incessancy with which saying always undergirds and resists the said which presupposes it. The said carries with it a foundation for ontology, as specifying a world in terms of signs, specifying 'a this as that', but the saying which addresses the ego as an interlocutor, demanding from him a response, is – *qua* demand, approach, imposition – already ethical. The Other approaches us through our sensibility, signifying the world in sensible signs, but always standing behind those signs as *someone* exposed to the world; language exposes us to the Other, exposes the exposure of the Other. Founding thought of the world beyond our own naïve sensibility is one sense in which Levinas can think "Ethics *as* first philosophy" (Levinas, 1989).

As in Husserl, one problem with conceiving the ego as a separated being is the problem of objectivity: the world is inseparable from the ego to which it is given. A newly articulated problem is that of hedonism: the way that the world is given to the separated being is through enjoyment, and the ego is thus a kind of hedonist. Expression resolves these issues by calling the naïvity of sensibility into question: the incessancy with which the Other can always challenge the ego disrupts a sensibility which is content in itself – it demonstrates a failure of sensibility to establish the existence it had taken for granted. Insofar as speech is demand and imposition, it calls into question the ego's freedom to pursue for itself whatever contributes to its happiness: no longer are the only questions posed to the ego by itself those concerned with its own enjoyment, e.g., how best to enjoy worldly things for itself, or when, i.e., hedonistic questions, but now also

include whether the ego ought enjoy these things at all. The hungry Other who approaches the ego and expresses herself calls into question the ego's right to their food, thus instituting a kind of rationality that Levinas takes to be foundational of all rationality. In his introduction to *Otherwise than Being*, Alphonso Lingis leverages Levinas' frequent formulation of being-for-another, which we might call 'ethical subjectivity' (or, perhaps more radical and simple, just 'subjectivity'), writing that "To acknowledge the imperative force of another […] is, materially, to give sustenance to another, 'to give to the other [Autrui] the bread from one's own mouth" (Levinas, 1981, p. xxviii).

Peperzak writes that "Face, speech, and expression are the concrete manners by which the irreducibility of the Other comes to the fore and surprises me, disrupts my world, accuses and refuses my egoism" (Peperzak 1993, p. 142). As disrupting sensibility, simultaneously resisting perception and making a demand on the ego – disrupting my world and accusing me – expression founds objectivity over naïve sensibility and ethics over hedonist morality.

Juxtaposition

The major, fundamental differences between Husserl's and Levinas' concepts of expression have to do with the ego's power over – or one might say right to – the Other, and the foundational role of similarity as opposed to difference. For Husserl, our relation to the other is founded on a basic similarity pertaining between us, which allows the ego to conceive of or otherwise direct itself towards the other, through intentional acts and synthetic achievements of its own, to the other conceived of in the ego's terms: the other is an alter-ego, distinct from but similar to myself, and known by me on the basis of that similarity. There is established, through empathy as a perception of expressive behavior, an increase in the ego's power: it constitutes both the other and the world for itself, through its own perception and experience. On the other

hand, for Levinas, expression presupposes disruption and difference, rather than similarity – the calling into question of the ego's sensibility and their right to the world; our fundamental relation to the Other is not a relation with an alter-ego, like myself, but rather with an interlocutor who calls me into question.

> The alterity of the Other does not depend on any quality that would distinguish him from me, for a distinction of this nature would precisely imply between us that community of genus which already nullifies alterity… The relation between the Other and me, which draws forth in his expression, issues neither in number nor in concept. The Other remains infinitely transcendent, infinitely foreign; his face in which his epiphany is produced and which appeals to me breaks with the world that can be common to us, whose virtualities are inscribed in our *nature* and developed by our existence. Speech proceeds from absolute difference. (Levinas 1969, p. 194)

The Husserlian conception of expression has been and will continue to be a useful one, and it is in part its usefulness that has earned it a prominent place in phenomenological thought. In phenomenology, it allows us to conceive of the other in familiar terms, founding understanding on an easily recognizable basis. In phenomenological science more broadly, the familiar structures of the ego founded by phenomenology – e.g., time-form, identifying synthesis, and so on – provide a basis for investigating others and communities. For instance, in cognitive science the intentional structures of the ego are assumed and extended to others in order to articulate particular modes of thinking; similar holds true for phenomenological approaches to affectivity. In psychopathology, it is typically these familiar structures of the ego, supposed to apply in broad strokes to others, that allows for explanations by modification and deficit: schizophrenia becomes ordinary experience *minus* some aspect of time-form, or

depression becomes ordinary experience *except* with a modification of <familiar structure determined by phenomenology as egology>…

One difficulty, however, is that similarity must be presupposed to get such an approach off the ground. The relevant similarities pertain not only to the material body in its shape or form, but also in expression, and thus are subject to culture, situation, and the languages that serve us well enough in everyday life. For these reasons, Jaspers, for example, was skeptical of the possibility of intercultural psychology, and thought some psychopathological experiences were simply "un-understandable" (Jaspers, 1997). In cases where the foundational similarity for Husserlian expression is undermined or otherwise cannot be presupposed, then, the method amounts to a crude and inaccurate projection. This is especially troublesome in psychopathology, where what is thematic is not similarity but difference, and where it seems that an application of one's own familiar structures to the other is especially suspect. Another, deeper concern is the question of confirming presuppositions of similarity, even when they hold. We assume a relationship, modeled on oneself, between movements and psychical states, but Husserlian phenomenology itself accepts that the psychic states of the other can never be given originarily. That is, our intentional acts directed to the other cannot be fulfilled in the ordinary way. Instead, it is through ongoing experience of the other – harmonious perception – that we come to confirm, though never conclusively, our understanding of the other.

All of this neglects another presupposition, however, which is precisely the presupposition of an originary communication between self and Other, a form of expression by which the Other expresses herself, regardless of my interpretation of that expression. It is on the basis of the Other actually coming into contact with oneself, precisely as someone else irreducible to me, that I can begin to make sense of the Other as someone like me. This is the

93

bed-rock foundation of harmonious perception, and of any interpretation of the other whatsoever; this is the basic fact on which intersubjectivity stands. To reiterate a point made earlier: First is communication, regardless of what is communicated and its recognizability, and only on that basis can there be understanding. Similarity is not a simple fact of which the individual is aware, but rather this awareness must be produced by coming to terms with the Other, and this means coming to terms with her on her terms, not my own.

This point is all the more important when we consider not only the foundational role played by the Other for the constitution of the world as objective, but also for the constitution of particular objectivities: it is, on both accounts, the Other who allows one to formulate specific things as objective, providing a concrete perspective on what might otherwise appear only to oneself. But here, similarity is only useful insofar as one is dealing with a real object, and of course one cannot know if that is the case without at least some Others confirming our own experience. It is through expressing herself that the other confirms or denies our experience, and it is specifically expression as *saying*, rather than as something said (which I can already recognize or appropriate as my own), that founds an understanding of what exists as an object and what is perhaps only idiosyncratic. At the level of the said, or expression in Husserl's sense, there is already a kind of validity of my own experience of the world that Levinas calls into question, by making the similarity that founds that very expression dependent on contact with an Other whose expression is not reducible to the familiar.

A fruitful way of formulating the difference between these approaches is in terms of a distinction between recollection and teaching. On the Husserlian approach, our awareness of the other is a kind of recollection, or maieutics (Levinas, 1969, p. 51 and throughout). We retrieve something from our own experience – a relationship between my own bodily movements and my

own mental life – and then apply it to something in perception – the bodily movements of the other. We understand the other in terms of ourselves, though of course in Husserl there remains the possibility of modification, negation and recombination, and the necessary numerical distinction between myself and the other: the alter-ego is similar to me but is not me, owing to my inability to experience her psychic life originarily; they are *another* ego. Likewise, when the other proposes something to us, we make sense of it by reference to what we have already experienced; the meaning of phenomena is always situated within the context of what the ego has already undergone. For Levinas, however, the Other disrupts my experience and calls it into question: their first act is not necessarily the confirmation of what I already know, with the addition of a sense of objectivity, but rather a demonstration of the inadequacy of the individual perspective and a calling of attention to that which escapes it. In the case of concrete, worldly things, the Other, through proposing the world in language, can teach us to attend to what we may ourselves have missed, what we may ourselves have not already experienced. This is not a project of making the implicit explicit but, rather, of learning, teaching. Whereas maieutics is "to receive nothing of the Other but what is in me" (Levinas, 1969, p. 43), teaching "comes from the exterior and brings me more than I contain" (Levinas, 1969, p. 51). In maieutics, expression is already interpreted by a meaning – a transfer of sense – determined within the egos own experience, and is then transformable as a modification of that egocentric meaning. For Levinas, on the other hand, "Discourse [or expression] is not simply a modification of intuition, but an original relation with exterior being… It is the production of meaning" (Levinas, 1969, p. 66; brackets mine).

> To approach the Other in conversation is to welcome his expression, in which at each instant
> he overflows the idea a thought would carry away from it. It is therefore to *receive* from the

Other beyond the capacity of the I, which means exactly: to have the idea of infinity. But this

also means: to be taught" (Levinas, 1969, p. 51).

Building on this passage, Davy claims that "Conversation with the Other that is welcomed, as

when one is open to learning from the Other, is a teaching" (Davy, 2007, p. 57). The appropriate

response to the Other, not only in order to challenge hedonism but also for the establishment of

objectivity, is therefore not first understanding or perception, not first a recognition of similarity,

but, rather, to welcome the Other across difference.

An Objection

It might be argued that this reading of Husserl is simply misled. One way this can be

done is by arguing that Husserl, in his fifth *Cartesian Meditation*, is not concerned with the

existence of the other but, instead, the very *sense*: 'other'. David Carr (1973; 1974) and Timothy

M. Costelloe (1998) take this position. They orient this discussion around Schutz' criticism of

Husserl, in which he claims that rather than accounting for the existence of the other, he

illegitimately presupposes it. The thrust of Carr and Costelloe's position is that Husserl begins

with the rich, everyday world in which we encounter others, and takes such an encounter – and

the very meaning of 'other' – as a datum to be explained within the phenomenological reduction,

pointing out that the "bracketing" characteristic of such a reduction is not skepticism or doubt

(akin to Descartes') but, rather, only a putting existence aside in order to understand the

production of meaning, or sense, and the ego's contribution to that production. Solipsism is a

problem from within the phenomenological reduction, but not from the world with which

phenomenology begins. What is in question for Husserl, according to Carr and Costelloe, is how

the ego can have a sense of the other, and what precisely this sense amounts to.

I first point out a difference between Husserl and Levinas when it comes to existence, and then turn to consider the sense of the other in Husserl.

Supposing for argument's sake, that Husserl is indeed not interested in the existence of the other, then this existence is left an enigma. It is often claimed that the traditional problem of solipsism does not occur within phenomenology or, otherwise, that phenomenology resolves it. Yet, if the traditional question concerns the existence of the other, and that very existence is bracketed by Husserl, then phenomenology does not answer the question. Indeed, since the sense of the other is something accomplished by the intentionality and synthetic achievements of the ego (even if passive), it seems that the ego could constitute the other *in the absence of any actual other*. We in fact do this all the time, seeing subjectivity in inanimate objects or even in the anonymous machinations of the cosmos. An intuitive and perhaps familiar example is that of a chatbot, which we may, for a time, think is a real interlocutor, but in fact is not. In such a case, we may constitute the sense 'other' in the absence of any actually existing other. While Husserl's approach is explanatory in such cases, demonstrating the production of such a sense, it nevertheless does not amount to a solution of the problem of solipsism, as the existence of the other is not even necessary for such a production. I take the process of bracketing to not be tantamount to blindly accepting everyday experience, and as such this existence of the other still calls for some kind of explanation, if the non-existence of the other would compromise the very achievement of objectivity meant to be carried out by the ego producing for itself a "sense" of this potentially non-existent thing.

On the other hand, for Levinas, the ego cannot constitute the Other for itself, since the Other is not an alter-ego nor reducible to anything familiar. He often formulates this position in terms of Descartes' idea of infinity, which the finite being cannot produce for itself and thus

must receive from something exterior to itself; for Descartes', this exteriority is God, and for Levinas it is the Other. Levinas thus takes a metaphysical orientation to the Other, and imposes a kind of transcendental requirement: proximity to or contact with an actually existing Other, through expression across difference, is necessary for Levinas' understanding of the sense 'other' (if we can be forgiven this language), in opposition to a concept of the other as alter-ego. Contact with an actually existent other is necessary for calling the ego into question, demonstrating its sensibility and the hedonism inseparable from it to both be inadequate. We can see in Husserl that such a disruption does not occur, as the other takes on its sense from the ego itself.

We might turn the above counterexamples against Levinas, showing that he too does not solve the traditional problem of solipsism. In that regard, then, the theories would coincide. It is important to note that Levinas does not formulate his project in terms of the problem of solipsism, but his project does entail some response to the problem: as opposed to Husserl's take-up of the problem, Levinas wants to affirm a kind of separation between self and other while nevertheless overcoming this separation in ethical (rather than perceptual) terms. In any case, when we compare Husserl and Levinas, the Other and her role in experience is of much more central concern for Levinas; the major reason for this is that, for Levinas, an actually existing other is necessary not only for establishing a sense of objectivity in experience, but also for founding a genuine ethics (that is, the Other is necessary for knowing we are not "duped" by ethics).

There is also a problem if we restrict ourselves only to a question of the *sense* of the 'other,' as Carr and Costelloe would urge interpreters of Husserl's fifth meditation to do. Levinas demonstrates an otherwise neglected dimension of the interpersonal relationship: the relation

between oneself and the Other is inherently, fundamentally (if we agree with him), an ethical one. The sense of the Other as an alter-ego is inadequate to this ethical dimension, in at least two ways. First, the sense 'other', in Husserl, does not call into question the egocentric morality. Indeed, for Husserl, ethics is often understood as a higher level concern to be addressed by additional acts of the ego, or as either subordinated to epistemology or equiprimordial with it – that is, epistemology itself being 'ethical', but in both cases immediately pertaining to the relationships between the ego and itself, the ego and its knowledge and understanding, the ego and its own exercise of freedom – freedom from bias on the part of the reflecting phenomenologist (Donohoe, 2004; Siles i Borrás, 2010). Ethics in Husserl does not primarily refer to the Other but is born of egology, referring to understanding, comprehension, and the meaning of concepts discovered by the phenomenologist; the path to the other follows the unchallenged egocentric trajectory of Husserlian expression and empathy outlined above. If what is essential in ethics, in addition to a (secondary) relation to an other (who may or may not really exist – it is, on Carr and Costelloe's view, a bracketed and therefore open question) is a primary reference back to the ego, then it is not clear how a non-egocentric ethics, hedonistic or otherwise, could be founded.[13] Even the great universalist principles of traditional ethics bottom out in the ego's own activity, as when Kant, for example, speaks of a rationality that gives the moral law to itself.

Second, the sense of the 'other' as alter-ego is inadequate to the way in which the Other overflows any idea we may have of her, subordinating the Other to a concept. In that sense, the phenomenological project of determining the sense of the 'other,' whatever that sense turns out

---

[13] Hedonism does not exhaust egocentric ethics. An ethics that refers back to the ego *may* be able to overcome hedonism, but nevertheless does not found itself on a relationship with the Other and, as such, is not properly speaking an "ethics"; the self is always privileged in some way, and altruism and sacrifice become impossible.

to be, will be inadequate to the way in which the Other actually manifests herself to our sensibility, i.e., as disrupting it, as non-containable by the ego. This disruption is the fundamental ethical moment at which one comes into real contact with an Other, and any sense that covers up this moment – even bracketing the existence of the Other – covers up the distinctively ethical aspect of the *experience (as-) of the Other*. This noncontainability, this refusal of static sense and the inadequacy of the conception of the Other as 'alter-ego', is the 'trace' left by the Other *in* the ego – the "other in the same" or the "one-for-the-other" of *Otherwise than Being* – necessary for the upheaval of egocentric morality.

To sum up my response to the objection that Husserl is not interested in the existence of the other, but only its sense: it may be the case that neither Husserl nor Levinas solve the problem of solipsism, but the transcendental move made by Levinas according to which the existence of the Other is necessary for ethics would then justify the existence of the Other if we indeed believe in the actuality of ethics. For Levinas, the possibility of altruism and even a small deed performed despite-one-self is all the proof needed of the actuality of ethical relationships, and thus too for the existence of the Other. Further explication on this point is outside the scope of this chapter. Additionally, the very sense of the other in Husserlian phenomenology is an inadequate one, as it does not account for the way that the Other overflows any idea we may have of her, since this overflowing is not a matter of numerical distinction between us but instead refers to the foundational ethical dimension of the interpersonal relationship. So restricting the question of solipsism in Husserl to a question of the sense of the 'other' does not obviously salvage his concept of expression, if the concept is meant to found an objectivity that must already refer back to ethics.

Conclusion

I have here explicated two concepts of expression. For Husserl, expression is founded on a perceptual similarity between myself and the other, and a transfer of sense from my own expressive behaviors, which always correspond to my subjectivity, and the expressive behaviors of other bodies, which are then constituted appresentatively as corresponding to a subjectivity like but distinct from my own. For Levinas, expression is speech, language, discourse, and the other's attendance to her own manifestation. Rather than similarity, difference is presupposed: the Other does not first appear in perception as similar to me, but rather disrupts my sensibility and calls hedonist morality into question.

Both concepts are meant to found the objectivity of the world as well as particular objectivities *in* the world. I have argued that, regarding the former, Levinas indicates an ethical dimension to objective-world-constitution, in virtue of the ethical nature of the interpersonal relationship. I here note that the prospect of such an ethical dimension is receiving increased attention in phenomenology and cognitive science. For instance, Colombetti (2009) attempts to set out a beginning of enactivist ethics based on the supposition that our relations to others are inherently ethical. Krueger (2019) develops a notion of the "between" taken from Watsuji's *Rinrigaku*, or "Ethics" (1996) in order to understand a kind of consciousness distinct from the Husserlian phenomenology for the purpose of understanding certain cases of psychopathology. Overgaard (2003) makes a serious attempt to reconcile Husserl with the ethical dimension pointed out by Levinas. De Jaegher (2019) points out an overemphasis on rational thought and knowledge and instead argues for the centrality of love and other ethical affectivity for a proper understanding of intersubjectivity. Levinasian ethics, in particular, is being taken up more and more in the areas of psychotherapy and clinical psychology, as a reprimand against a distanced

objectivization of patients by psychological theory and understanding, the latter being widely influenced by Jaspers's broadly Husserlian theory of empathy (see, for example, Gantt, 2000; Gantt & Williams, 2002; Nortvedt, 2008; Sass, 2019; Tsang, 2017). Interest in Levinas is now increasing in other areas of phenomenological ethics, as well, such as animal and environmental ethics (Atterton, 2011; Crowe, 2008; Davy, 2007). In all of these areas, difference becomes problematized over similarity, and ethics privileged above – even as generative of – knowledge. It should be no surprise, then, that Levinas is gaining traction in those areas where understanding and knowledge must apparently be grounded in ethics across difference. Even the Husserlian approach, I have suggested here, covertly assumes the kind of communication thematized by Levinas, and so even adherents to a broadly Husserlian methodology stand to benefit from Levinas' formulation of expression.

Levinasian expression founds the possibility of truly expansive objective knowledge, as the basis of our relationship with the Other is already one of teaching, that is, of an introduction to the ego of that which it did not already contain, rather than the Husserlian model for which the basis of understanding must always refer back to the ego itself. I would suggest that the way to understand this kind of expansion mechanically – if one is so inclined to do so and recognizes such an approach's inherent risk of subordinating the other to concepts – is to explore the intersubjective dynamics of attention regulation; for Levinas, the Other teaches attention.

Chapter 3

Encounter:
How expression teaches attention for the world and other,
with implications for psychopathology

Introduction

Expression is at the heart of predominant approaches to intersubjectivity, including

within cognitive science, psychology, psychiatry and psychopathology. Prominent theoretical

conceptualizations of intersubjectivity and intersubjective sense-making especially germane to

these areas include Direct (Social) Perception Theory (DsPT), Interaction Theory (IT), and

Participatory Sense-Making. Each approach attempts to explicate the situation or context,

meaning, and what we might very broadly call the "mechanisms" of expression and how we

come to understand one another through expression. Yet, as a consequence of the argument from

Chapter 1, there are some expressions that are seemingly – on standard phenomenological

accounts, from Jaspers to Husserl, to Heidegger and contemporary thinkers – ununderstandable:

they do not adequately articulate the other or her mind, and leave it a mystery as to her

experience. This occurs, according to Jaspers and implicitly for Husserl, when there is too radical

a divergence from myself and my own phenomenological, psychological, or existential structures

and the real or apparent being of the other. Accordingly, it might be thought that the Levinasian

concept of expression, committed as it is to both separation and the absolute alterity of the other,

might be deeply committed to a form of ununderstandability; indeed it may be difficult to see the

Levinasian approach as lending anything new to intelligibility, and opponents and proponents of

Levinas alike often charge his position with a kind of mysticism, deeply committed to what

cannot be understood (Derrida, 1978; Janicaud, 1991). It has been argued – and to some small

extent addressed by Levinas – that a Levinasian ethics, with its concept of expression, would

result in a kind of relativism or (epistemological) nihilism: if expression signifies beyond any

familiar horizon, it will be inherently ununderstandable. These assessments pick out something right, but this is neither a complete picture nor a reference to anything mystic.

Recent enactivist approaches to intersubjectivity often endorse some version of direct social perception theory, interaction theory, and/or participatory sense-making, each of which offers some form of rejection to the ununderstandability thesis (UT). The basic strategy is to reject any kind of radical inaccessibility of other minds. Direct social perception theory maintains that expressions partly constitute the mental states with which they are associated, e.g., there is happiness *in* the smile, and since expressions are perceptible, so are the states they express; there is accordingly no problem of radically inaccessible minds other than my own: far from being inaccessible, other minds are available to perception. Interaction theory maintains that intersubjective perception or, more broadly, understanding is established within and dependent upon interaction between agents, operant prior even to the acquisition of concepts: the view holds that perception of the other's expression as meaningful is originally non-conceptual, pre-linguistic, embodied, and conditioned by the situation of multiple agents interacting with one another as they navigate a shared social and material world. Participatory sense-making maintains that interpersonal interactions take on an autonomy of their own, such that meaning is generated by the interaction itself and, more radically, that individuals are only constituted through interaction in the first place; individuals and individual meaning are thus products of an interaction that already incorporates a plurality of persons.

Given the fact that most theorists find the ununderstandability thesis to be either untenable or, more often, simply unacceptable, it would be a blow to my approach were it committed to UT, at least in any deeper way than are, as I contend, these other approaches. Accordingly, I here elaborate my position so as to reject UT in a way distinct from those options

just presented. Rather than denying any separation between individuals, my position presupposes separation, and understands expression as a relation across that separation. I consider the theory of natural pedagogy to show how learning can nevertheless occur across a kind of separation or inaccessibility.

I begin by briefly revisiting the ununderstandability thesis' motivations and raising the concern that Levinas may be committed to some form thereof. I then turn to the relationship between expression and attention, or *teaching/learning*, in order to articulate a form of understanding even across the radical difference demanded by Levinas. This can be seen as an extension of Levinas' thought on learning and understanding (the key insights are contained within Levinas, but only infrequently articulated in any detail). My goal is to express these insights in a way that is both accessible and helpful to contemporary thinkers working on such issues. I make use of the theory of natural pedagogy for this purpose, as this theory and Levinas' thought, I hope to show, converge on certain relevant insights. I also suggest that dynamical systems theory has something important to offer us, and conclusions can be drawn about the way in which a certain form of attention should – and, when done well, does – guide psychiatric practice. The first upshot of these analyses is the formulation of a unique objection to UT: the Levinasian denies UT in a way that is distinct from (but not necessarily incompatible with; for more on this relation, see *Chapter 4* below) enactivist approaches to intersubjective understanding. This difference allows me to draw some initial implications for psychopathology.

Ununderstandability

Jaspers claims that certain psychopathology symptoms are beyond empathy and are therefore ununderstandable. "Made phenomena," for example, are entirely inaccessible to those who do not undergo them and cannot be psychologically understood positively, i.e., as what they

are; they can only be negatively circumscribed from the outside (Jaspers, 1963, p. 580). This so-called "ununderstandability thesis," or "UT," is largely considered to be theoretically or practically unsatisfactory. Many thinkers attempt to refute UT by arguing for a continuity or structural relationship between pathological and non-pathological consciousness. This is of course true in phenomenological psychopathology, when researchers offer structural explanations of pathology experience and its relationship to "normal" consciousness (Fuchs, 2013; Gallagher, 2015; Ratcliffe, 2014): pathological experience *is like* non-pathological experience, *except* […]. Another common strategy consists in appealing to some specific, ordinary, everyday phenomenon and constructing an explanation of how a feature of that familiar phenomenon enhances our understanding of some (aspect of) pathological phenomenon. Like the previous, this strategy is meant to undermine the radical separation that UT seemingly depends on and which, within the Jasperian mode of thought, perhaps inevitably prevents understanding. For example, one might use the experience of being shoved to provide an analysis of the way in which a sense of ownership (e.g., it is *my* body that is moving) can come apart from a sense of agency (e.g., it was *not* me who is (causally) responsible for the movement), and then apply this distinction to psychopathology phenomena, e.g., to explaining thought insertion (a "made phenomenon"): thought insertion, on such an analysis, would be the dissolution of the sense of ownership and/or agency with respect to one's own thoughts, a dissolution we might ordinarily be familiar with, at least in broad kind, from certain bodily experiences (these approaches vary: Shaun Gallagher (2015), for instance, maintains that thought insertion involves a loss of the sense of agency, while others, such as Bortolotti & Broome (2008) argue for a lack of the sense of ownership as well as "authorship"). Other strategies, such as those proffered by direct

perception and interaction theorists, likewise depend on rejecting a radical difference or separation between familiar and alien states, or between oneself and others.

The Levinasian approach, however, is indeed committed to a certain form of radical difference between self and other, a separation between them such that any concept, all of which are founded in some way or other on familiarity, could only be inadequately applied to the other. On traditional models, and certainly on the Jasperian picture, understanding is seemingly inseparable from some reference to familiar horizons and known elements. There is thus a resulting concern when we consider the Levinasian perspective through the lens of these traditional conceptualizations of understanding: if the other is she who *cannot* be made to appear within familiar horizons, she who is refractory to concepts (Levinas, 1969, p. 40) and she whom the familiar can only present by dissimulation, then is not the Levinasian committed to UT?

The Other, for Levinas, is she who exceeds reference to the known, and, thus, would exceed understanding (traditionally conceived). Yet, no small portion of Levinas' task, not only in those works that explicitly deal with understanding and intelligibility (e.g., "Transcendence and Intelligibility" in Levinas, 1996), can be conceived of as articulating a different form of understanding, one that can withstand the kind of distance introduced by him between self and other. No doubt, this is not simply a change in mechanism yielding the same finished product as the traditional notion. The task, rather, is to articulate some form of sense or meaning *prior to* the appropriation of the other by the same, a form of understanding that is born of separation rather than its negation, and which does not strive to totalize or systematize the other by reducing them to knowledge and familiarity, that is reducing their very alterity. Despite the separation insisted upon by Levinas, he is equally insistent that it is only in virtue of a form of contact with the other – a contact all too often or perhaps altogether missed by traditional phenomenology and its

107

contemporary inheritors in psychopathology – that our attention can be drawn to that which is not already known, that which we do not already contain, that is, that we can genuinely attend to the other. The Levinasian perspective, accordingly, strives to undermine UT without denying the separation often (but incorrectly) taken as the guarantor of incomprehensibility, promising a kind of "contact across distance" (Levinas, 1969, p. 172) as the ultimate foundation as well as first instance of (i.e., primary) intersubjective understanding. The linguistic relation in a personal encounter accomplishes contact across distance.

Levinas is often taken to imply, particularly within psychotherapy discourses, a kind of "unsettled" therapy (Rossiter, 2011) "without foundations" (Loewenthal, 2011). Such characterizations often instantiate only an incomplete understanding, offered by proponents and opponents alike, who tend to variously attribute to Levinas a corresponding commitment to some form of relativism (no general truths are possible, since truth is determined by each particular other), or nihilism (there simply is no fact of the matter when it comes to the other). The various contributors to *Psychology for the Other* (eds. Gantt & Williams, 2002) resist these latter attributions, arguing throughout that, despite Levinas' qualms with knowledge and method, Levinasian psychotherapy is not thereby doomed to relativism, nihilism, or otherwise to providing no methodological guidance whatsoever, and that resisting these purported commitments is crucial for developing an effective and coherent Levinasian approach. It is my position, largely in agreement with these insights of *Psychology for the Other*, that a Levinasian psychotherapy and psychopathology alike, be it clinical or theoretical, need not be committed to a complete lack of method, at least not if we understand 'method' in a broad sense. Nevertheless, we must also attempt to capture the sense in which it is true that Levinas is, in specific ways, opposed to foundations and knowing when it comes to the other. His philosophy is one that

resists systematization, refusing to subordinate intersubjective or, perhaps more appropriately, (inter-)personal relations to rules and principles. In the following section, I develop an understanding of expression and attention that sets out a *kind* of "foundation" for psychotherapy and -pathology, while nevertheless striving to respect the anarchic aspect of Levinas' thought.

Encounter

Expression and Attention (Learning)

There are at least two major ways that Levinas speaks of attention in relation to expression: the way that the other attends to her own manifestation, and the way that, in doing so, she directs or teaches my attention The way that the other attends her own manifestation, that is, *expresses herself*, was thematized in *Chapter 2* above, so will not be belabored here. However, one point that must be added is that the way the other attends to her own manifestation, and even the mere fact that she does so, is an integral part of a Levinasian theory of learning: without expression, in which the Other attends to her own manifestation, the world is a kind of "anarchy," without principle for distinguishing phenomena from mere apparition. Apparition is Levinas' term for an appearance without reference to that which appears, that is, without reference from the phenomenon to an existent. An appearance is not yet a phenomenon, but that which appears as merely subjective and without reference to any "thing" at all. A phenomenon proper, on the other hand, is a complex kind of appearance referring from consciousness to a thing-itself, implying a subject-object relation (whatever ontological status might pertain to each term). The object of the phenomenon must appear *as* some kind of object; this is not the case for appearance, in which what appears is merely lived through according to a subjective and indeterminate sense, referring specifically to the individual bodily sensibility through which it appears at all. What is needed to overcome the anarchic world is a principle for

distinguishing phenomena from pure sensible appearances (and this is primarily a distinction of *sense* construed in terms of the way in which the world appears). According to Levinas, "Speech introduces a principle into this anarchy" (1969, p. 98). In speech, phenomena are placed into propositions, according to which they gain signification, which is the form of meaning defined by a relation of *this* to *that*, specifically the form of thinking "X *as* Y." It is this "as" structure that defines signification and picks out what appears as a specific 'what' that appears. On the one side, signification issues from speech, and on the other side it introduces a principle – this 'as' structure – according to which phenomena can refer beyond myself and, instead, to other phenomena, or to the Other who expresses them. It is insofar as the other attends to her own manifestation that she embodies this very principle, or, rather, that she introduces – in the way that we say of someone that she introduces herself – both herself and a principle into anarchy.

It is better to understand speech not only in terms of propositional language, but as communication more generally. As seen in *Chapter 2* above, expression refers to the entire body, and so even non-linguistic bodily expression can serve to indicate the "what" behind a mere appearance and can present the perspective of the other for whom it appears. Within appearances, we do indeed perceive differences – our sensory field is not a uniform one but rather one in which differences matter to us – and expression in general, not just propositional expression, can serve to indicate differences to others. The idea of attentional direction and the theory of natural pedagogy (explored below) can help us to better understand how bodily expression can introduce significance into experience, that is, refer the things given to my perception beyond their mere appearances for me.

Not only does the other attend to her own manifestation in expressing herself, but she also directs, commands, or teaches my own attention. In expressing herself, the Other proposes a

world, providing phenomena with significance. Providing phenomena with significance by expressing them is also called thematization: what is spoken of becomes a theme for consciousness. No longer are phenomena revealed only to a sensibility that, in an important sense, lies *beneath* consciousness (understood as intentionality, as *consciousness-of*) for Levinas, but rather phenomena can now be *explicitly thought* as well. In Husserlian terminology, we can say that the Other places phenomena into a horizon of thought; in Heideggerian terms, the relation to the Other who expresses herself is prior to and founds the moment of disclosure. Indeed, much of Levinas' critique of these two figures consists in showing that intentionality – consciousness' being always related to an object, and always within a horizon – and disclosure do not constitute the foundational moments of consciousness but are themselves founded on a relation to the other. (It is in this sense, if we keep in mind that "ethics" for Levinas refers to such a relation to the other, that we can understand Levinas' move to make ontology secondary to ethics, i.e., ethics as first philosophy.) The other introduces a principle to anarchy by expressing herself, which is simultaneously an act by which my attention is directed to a world thus thematized. That the other expresses herself is what projects me toward a world at all, by establishing a principle of phenomena, while the specific ways that her expressions command my attention direct me toward specific features of that world.

The position here is that expression is teaching and as such also the foundation for learning from the other, that is, attention: expression teaches by directing our attention to that which we were not already aware, or simply, expression teaches attention. In the first place, to the other who expresses herself (the saying) and, consequent to this, to that which is thematized by that expression, given as significant (the said). The distinction between primary and secondary intersubjectivity originating in Trevarthan (1979) and Trevarthan & Hubley (1978) is

111

helpful here, though I will be reformulating the former. The first form of attention is attention to the other who expresses herself, or primary intersubjectivity. The second form deals with attending to that which is picked out by or attended to by the Other and, in doing so, better understanding the other, i.e., secondary intersubjectivity.

Primary and secondary intersubjectivity, however, are predominantly understood as referring to our understanding of or direction toward the other, that is, forms of intentionality. Primary intersubjectivity, for instance, refers to the way the other comes to be constituted as other *for us*. But for Levinas, constitution is not the primary issue. The relationship to the other is one that cannot be understood primarily in terms of consciousness. That is, though the other is constituted for us in a variety of ways, this all depends on the prior fact of communication – of expression – from the other to me, a fact which cannot be reduced to its appearance to my consciousness. The primary form of intersubjectivity on this Levinasian model is a form of contact. Contact with an other who is not reducible to their appearance for me founds what is *typically* known as "primary intersubjectivity," or *consciousness* of the other as other, and makes secondary intersubjectivity possible by accomplishing joint attention. It is by encountering the other on their own terms, that is, encountering an other who expresses herself and draws my attention, regardless of the specific content to which my attention is drawn, that this contact is accomplished. This encounter is asymmetric, in that the other disrupts me and directs my attention; it is the asymmetry between teacher and student who, nevertheless, can based on this relation subsequently come to understand each other, that I first enter into a relationship with both the other as such as well as an intersubjective, shared world.

Natural Pedagogy

The position resembles, and can gain elucidation through, what is known as the "theory of natural pedagogy," developed by Gergely & Csibra (2009, 2013; Csibra & Gergely 2011). According to the theory of natural pedagogy, human beings quickly and efficiently learn generic information from one another through a certain form of communication. Gergely & Csibra argue that human beings in general (though they emphasize infants) are receptive to indications that they are being addressed by others in communication, and that being addressed by an other makes a key difference for learning. They write that

> The most unique proposal of the theory of natural pedagogy is the hypothesis that the information extracted from the other's ostensive-referential communication is encoded and represented qualitatively differently from the interpretation of the same behavior when it is observed being performed in a noncommunicative context. (Gergely & Csibra, 2013, p. 128)

Specifically, when one merely observes the other and their actions, they are directed toward the 'here-and-now': they see the performance as relating to current objects, events, or perhaps even goals, specific to this unique situation, perhaps even idiosyncratic. To use their example of a man opening an unfamiliar bottle through a series of specific movements: it is unclear which of these are necessary for opening the bottle and which are merely superfluous behavior, and it is unclear if this bottle itself uniquely requires such methods (Csibra & Gergely 2011, p. 1149).

If the other, in this case the man opening a bottle, first grabs one's attention, however – primarily by establishing eye contact and/or employing directed speech, both of which "make it manifest that a [person] is being addressed" (Gergely & Csibra, 2009, p. 150) – there is a shift in our understanding of the actions performed: we take the action to express, first, the other who addresses us, that is, the action is no longer something I observe in a detached way but instead as

113

a demonstration *for me* (Gergely & Csibra, 2009), and, second, the action no longer refers to or imparts understanding of specific features of the here-and-now but rather generic and cultural information. Teaching and learning refer to address. For instance, in the example of the man opening a bottle, if he establishes eye-contact and addresses me by speech, there is a shift in my attention such that I am now more likely – and, as is especially clear in the infant case, more capable of – perceiving his expression as teaching me something *general*: this is how you open this kind of bottle (and, Gergely and Csibra add, that this specific kind of bottle-opening practice is cultural). In so far as learning and knowledge trade in generality or universality, and such an address institutes the transition from particularity to generality, natural pedagogy describes the way that the other, encountered in a certain way (that is, being addressed by the other), institutes a kind of epistemic principle: being addressed by the other is foundational for knowledge acquisition.

The theory of natural pedagogy (which has been shown experimentally), particularly when coupled with a Levinasian perspective according to which the other who approaches us addresses us prior to being constituted *by me* as other, is then useful for further specifying primary and secondary intersubjectivity. There is a way of attending to another that is non-communicative, non-personal, and this way corresponds to observing their actions as specific to the here-and-now. General knowledge, at this stage, is impossible and, given its observational mode and objectifying attitude, this mode seems hardly worth the name "intersubjectivity" at all. There is then another form of interaction, which truly deserves the name "primary intersubjectivity," since it involves a communicative relationship to an Other as an Other, as one who addresses me in communication (Levinas may call this "saying"), and this makes possible a

secondary intersubjectivity in which we jointly attend to features of the world, making sense in general terms ("the said").

That the Other's expression, as addressing me, teaches attention to worldly things (objects, practices, cultural things) so as to teach something general about them beyond the here-and-now or, as we might say, beyond the mere sensibility that constitutes the very here-and-now-ness of the here-and-now, is an overlap between my interpretation of Levinas and the theory of natural pedagogy. (Another somewhat inexplicable place of overlap is that both views exclude animals from the account, claiming that such a dynamic is at play only for human beings, with Gergely & Csibra going as far as to argue that natural pedagogy is a unique human evolutionary adaptation. I am here neutral on these further points.) What is crucial is that a certain, important kind of knowledge – perhaps *the* form of knowledge, rather than mere acquaintance with a plurality of contingent particulars – depends on a communicative relation in which the Other directs my attention through a personal address.

Focusing on infants, Gergely & Csibra write that:

The evidence we reviewed here indicated that infants are also prepared to learn generic kind-relevant information directly and from a specific source that is not available to other species: from benevolent communicators who manifest generic knowledge 'for' them that would be difficult (if not impossible) to acquire without such support. (Gergely & Csibra, 2009)

That is, the other addresses me and draws my attention, *thematizing* that which they attend to for me, allowing for general learning of that which I perhaps *could not* have learned without such an address. Striking a similar chord, Levinas writes:

As an attendance of being at its own presence, speech is a teaching. Teaching does not simply transmit an abstract and general content already *common* to me and the Other. […] Speech

first founds community by *giving*, by presenting the phenomenon as given; and it gives by

thematizing. (Levinas, 1969, p. 98)

I see these two approaches as deeply compatible. The theory of natural pedagogy not only allows

us to make the Levinasian hypothesis more readily understandable and specific, but also provides

a naturalized account of how such a hypothesis could be realized in cognitive and evolutionary

terms. I do not maintain that such an account is necessary for the justification of a philosophical

position, but philosophical positions – especially ones that hope to intervene in empirical debates

– must be amenable to such forms of explanation, and I point readers to whom such an

explanation is important or helpful to the work of Gergely & Csibra.

One major difference to note between my reading of Levinas and Gergely & Csibra's

theory of natural pedagogy has to do with the role of perception. On their view, there are certain

perceptual cues that signify that one is being addressed by an other, and as such it is at the level

of perception that one identifies a teacher and gains the capacity for learning. This construal of

the way that the other affects myself is consistent with an interactional or exchange view, in

which student is just as active as the teacher, though in a different way. On my view, however, it

is not perception but, rather, the interruption of perception that establishes address. That is, the

gaze or directed speech of the other does not grab my attention as a special kind of perceptual

object, but rather precisely insofar as it grabs my attention in a way that perceptual objects do

not: it points beyond what I perceive to someone else who directs my attention through my very

own perception. In this reading, encounter is distinct from interaction, in that the former is

asymmetric and has to do with the way the other commands my attention from outside of my

own powers. The encounter consists, initially, in being acted upon by the other, and only in

making sense of this act do I become active myself. The encounter refers to my own passivity in

116

the face of the other, beyond even a capacity to receive the other. It is precisely as interruption of perception that nevertheless leaves its trace in perception that the other addresses us and draws our attention in a way distinct from perceptual objects: the latter appear to our sensibility and are taken up by it in enjoyment, appropriated, and passivity becomes activity; the other interrupts the contentment of enjoyment, redirecting my attention to what is beyond my own activity.

Dynamical Systems Theory

Recently, dynamical systems theory approaches to cognition, especially those within "4E" (enactive, embodied, extended, embedded) traditions, have also emphasized the ways that our relations to others condition or undergird our understanding of each other and the world. Though such approaches are not univocal, they generally agree that our understanding of ourselves, each other, and the world depends on the concrete relations we find ourselves in, and the ways in which our own behavior and thought is modified by and likewise modifies the behavior and thought of others. Some examples, offered by Joel Krueger (2014, 2010), include musical improvisation and breastfeeding. In improvisation, each musician's performance is shaped by, and likewise shapes, the performances of the other musicians, as well as standing in a dynamic relation to the ensemble's overall performance. In breastfeeding, when infants pause, this prompts mothers to "instinctively jiggle the infant, or exhibit similar tactile behavior, as a prompt to resume feeding. And it seems to work […]" (Krueger, 2010, p. 10).

I here understand such dynamics primarily in terms of attention. The notes played or the rhythm structuring the music, etc., draw the musician's attention to the music and to their own instrument and capabilities in ways that would not have occurred without this relation. The infant, by pausing, draws the mother's attention in such a way as to solicit a response that, in turn, redirects the infant's attention. In explicating the unique contributions of Watsuji to

117

phenomenological and cognitive discourse, Krueger goes as far as to write that "basic attentional processes *by which individuals constitute intentional objects* may be modulated by the sociocultural contexts in which these processes arise" (Krueger, 2020, p. 765; italics mine). Within discussions on secondary intersubjectivity that emphasize interactive or coordinative dynamics, the emphasis is often on joint attention, which is sometimes described as necessary for secondary intersubjectivity (e.g., Gallagher, 2012). A dynamical systems theory approach to understanding the ways that we make sense of one another and of the world, then, is already largely and perhaps fundamentally a study of attentional dynamics, whether it explicitly takes on this moniker or not.

There has been, regrettably in my view, an overemphasis in the literature, however, on similarity, mimicry, and reciprocity in the interpersonal relationship as studied by dynamical systems theory. Simulation theory in particular bakes this bias into its conception of interpersonal understanding: we understand the other by simulating them, that is, by some kind of mimicry (whether explicitly conscious or pre-personal varies). The discovery of so-called "mirror neurons" has intensified this idea, as simulation theorists typically understand mirror neuron activity as simulation at the functional level, and/or the basis of mimicry in general. Phenomenologists too, albeit in somewhat different ways, also rely on similarity, mimicry, and reciprocity in their explanations of intersubjectivity: Husserl and the perception of similarities between myself and the other; Heidegger's account of Dasein caught up in a similarity with *das Man*; Merleau-Ponty's emphasis on reciprocity in his concept of intercorporeity. This reliance sometimes translates into overemphasis in contemporary phenomenological work, as well (for instance, as I discuss in depth in the remaining chapters, see Froese & Krueger, 2021). These approaches orient themselves around infants mimicking their caregivers (for example by

replicating facial expressions and gestures), or participation and reciprocity (according to which to understand social practices properly or to truly understand those who engage in them requires us to participate in those practices, at least to some extent, and to construct the meaning of those practices together, in a give and take relationship). This includes the participatory-sense making literature (which takes its name from De Jaegher and Di Paolo's 2007 article) and appeals to the role of "common-sense" and the loss thereof in psychopathology, e.g., in de Haan & Fuchs (2010) or Stanghellini (2000, 2001 & 2004). Yet, two points can be immediately made against this emphasis.

First, it is not at all obvious what role *mimicry* must play in attentional dynamics and learning. The dynamic relationship between an infant and their mother during breastfeeding seems to imply there is indeed an *asymmetry* and *divergence* of action that pushes the interaction forward. The infant does not do as the mother does, nor vice versa. Rather, they modulate one another's attention without mimicry, and the situation is shaped by a plurality of action types, not simply a multiplication of same-kind tokens.[1] To take the music example, the dynamic relationship may be productive of synchrony (or, sometimes, certain forms of asynchrony), but only very rarely mimicry: only on rare occasions, typically for emphasis, is it desirable for all string instruments, for example, to play the same melody. This technique is indeed good for emphasis precisely because it is not usually applied.

It may be the case that mimicry demonstrates some form of understanding of the other especially germane to questions of theory of mind, simulation, and perhaps even empathy, but mimicry does not characterize intersubjectivity as such, nor does it seem necessary even for establishing shared understanding. Suppose, for example, that one witnesses a horrific racist act.

---

[1] Similar points have been made by De Jaegher and Froese 2009, and Gallagher 2020.

It does not seem that one must imitate features of that act, nor even *truly* understand the victim's experiential world, in order for it to be meaningful (and repugnant) in the first place.

Second, mimicry already presupposes attention: to mimic a face, one must already attend to it. Infants selectively attend to human faces (Lewkowicz & Hansen-Tift, 2012; Tsang, Atagi & Johnson, 2018). Here, we can apply the insights of the theory of natural pedagogy. It is as addressed by an other that the infant (and the rest of us) learn from the other, and it is only as addressed by the other that mimicry makes sense: we mimic *the other*, whereas we do not, at least to the same extent and at the same developmental stages, tend to mimic inanimate objects or even images or recordings of others (Shimada & Hiraki, 2006). Catching the infant's gaze, or the infant catching ours (we gaze back), addressing the infant in "baby-speak" (indicated as significant cues by Gergely & Csibra), these are instances of the first grabbing-of-attention that then makes mimicry an attractive possibility (and, sometimes, productive of further understanding).

These points reject symmetry of the interpersonal relationship in the sense of simulation or mimicry, and specify asymmetry in terms of attention: the infant attends to the other who addresses them. Reciprocity, which characterizes intersubjectivity as a give-and-take relation in which two or more parties all contribute, may well factor in to sense-making, in the traditional senses of primary and secondary intersubjectivity, but they presuppose the asymmetry accomplished by address and its direction of attention. We learn from the other as addressed by the other, as one for whom the other communicates; there is at the heart of attentional dynamics an asymmetric teacher/student relationship. Of course, the student might also come to teach the teacher – both directions can be the case, even simultaneously. But the possibility of symmetry is not what defines the intersubjective relation. Instead, address to one from the other, grabbing

one's attention in the right way, is really what is primary when it comes to intersubjectivity: this asymmetric direction of attention is what first institutes a relation to the other as other, founding understanding of the other and the world. Mimicry is neither fundamental nor strictly necessary for the intersubjective relationship nor for understanding.

Ethics

Lastly, there is an ethical component to all of this (and indeed, teaching is "ethical"). Consider care-giving. The caregiver responds to the infant even when they do not know what the infant wants or needs. First there is an address, and then there is understanding. The infant draws the caregiver's attention in this way or that, and the caregiver responds accordingly, e.g., by feeding the infant, changing their diaper, initiating play, and so on. This of course feeds back into infant behavior – once the baby is burped, which is what they wanted all along, for example, the crying ceases. It is indeed through the search for something that will relieve the infant's cries that the caregiver comes to understand what it is that the infant expresses. In this case, the infant is the teacher, and the caregiver a student. This does not articulate a structure of a developmental phase to be overcome, but rather characterizes any interaction with another: first, our attention is directed to the other or to that which they attend, and it is based on this shift of attention instituted in us *by* the other that we come to understand.

Use of the caregiver example does not load ethics where it does not belong. What I am referring to is a general structure of interpersonal dynamics: we attend to the other, and then we respond. This is the asymmetry, and this call for response is the demand (or perhaps necessity) of ethics. Expression is tied to responsibility and, as such, is ethical on Levinas' account. This is due to expression's simultaneous negative and positive moments. This was explored in *Chapter 2* above but should be briefly restated here: In the negative, expression is an interruption of

sensibility's contentment. That is, it calls into question my naïve freedom and basic hedonistic morality precisely by making me answerable to someone: an other presents herself to me, which calls into question the subjectivism according to which the world is *for-me*. Again, this is a kind of ethical version of the Husserlian epistemological task whereby the other demonstrates to me, by being constituted as another perspective on the world, that the world is not exhausted by my own intentionality. But, for Levinas, it is not a matter of constitution, but rather of interruption: the other is not constituted by me in the encounter but is only dissimulated by the application of familiar concepts to an other that has already made herself present. It is this making oneself present – manifesting oneself – that interrupts my contentment. Expression disrupts the free-play of constitution. The other side of the same coin, i.e., the positive moment, is that expression is always a demand, command, invitation. That is, it presents the speaker and solicits response. This solicitation is inseparable from a kind of responsibility: I am responsible to the other in the sense that, having spoken to me, she has made herself present as she to be answered. My freedom no longer has a right to proceed according to its own spontaneity, but rather is now responsible for justifying itself. That is, if I am to act in any way at all, there is now an other present to whom justification becomes possible, invited, demanded. Of course, a particular other may not actually demand justification of us on each and every of our acts – such a state of affairs might be maddening, though I believe Levinas would not shirk the idea that the demands of ethics may be overwhelming and maddening – but the point remains that our freedom becomes enmeshed with practices of justification, that is, justification becomes required or unnecessary, as soon as an other to whom we would offer or deny such justification presents herself to me.[2]

---

[2] Though ethical responsibility is overwhelming, on a Levinasian account, it is also the case that, for the sake of justice (that is, extending our responsibility to *all* others, and not just one), we develop customs and practices to guide our responses. Occupying a certain social role, for instance that of a teacher, comes with a certain set of expected and prescribed responses: answer emails in a timely fashion; help students develop good study habits; foster a respectful

The caregiver is just one example in which an other, here the infant, expresses herself and makes a demand on my freedom, requiring a response and, specifically, requiring me to act according to that expression or otherwise justify myself (indeed, it is easy to imagine a toddler, being allowed to cry without assistance, asking why their parent is not helping them, and even in the infant case others often get wrapped up into this questioning, asking the parent "why are you just letting her cry like that?"). The structure remains intact in any number of examples: when the bassist gets a guitar player's attention with eye contact and an overt movement of the head, he solicits a certain kind of response from the guitarist. When one knocks on your door, a response is invited, solicited (we say: "no soliciters"), demanded. One asks a question which calls for an answer, one says "no really, it's okay" which calls for consideration or continuation, and so on.

One can of course be inattentive or neglectful to the other who expresses herself as command, but this is no objection to the position. First, attention must already have been redirected by the expression of the other for one to be properly accused of ignoring the other. Second, we already consider inattention to be an ethical failure. Not only is this confirmed by our practices by which we accuse and charge people of neglect, but for Levinas it is more deeply inexcusable to leave the other hungry, even if we are not aware of their hunger. To cite inattentiveness is to "give reasons, when it is food the other requires" (Perpich, 2008, p. 89), that is, to behave unethically. Third, in part because expression is defined by the way that it interrupts our contentment and perception, Levinas would maintain that an absolute inattention – encountering the other yet nevertheless remaining utterly unaffected by her – is an impossibility.

---

classroom environment, and so on. The demands made on a given person are infinite, but our commonly accepted practices allow for a degree of consistency and practicality that, when responsive to the actual concrete demands made by others, can help to foster justice for all.

We may indeed not care at all about the plight of the other (which would be our ethical failing), but Levinas maintains that "The being that expresses itself imposes itself, but does so precisely by appealing to me with its destitution and its nudity – its hunger – without my being able to be deaf to that appeal" (Levinas, 1969, p. 200). As Diane Perpich puts the point, "This moment, where deafness eludes me, is the moment of responsibility. It binds expression to responsibility in Levinas' account and constitutes both as the essence of language (Levinas, 1969, p. 200)" (Perpich, 2008, p. 96). It must be noted that, for Levinas, murder, too, is an impossibility. Of course, he does not deny the banal fact of killing, but rather maintains that murder is an ethical impossibility, which is of a different order than that of physically destroying someone's body. Likewise, it is an ethical impossibility to remain unaffected by the other who expresses herself: if she indeed expresses herself, we are already responsible to that expression, whether she manages to "get through" to us or not. What is primary is the expression, which enacts an affection, but not that affection itself. The affection is only primary in the sense that it is the trace of the expression, and as such is our first (though inadequate) *own* resource – produced by the other but inside me – for turning toward and conceptualizing the other who expresses herself.

The point here is that learning is not an unethical engagement. It is founded on an asymmetric relationship with an other who addresses us and, in doing so, solicits, demands, or warrants some response. This becomes salient by the way that the other grabs our attention, and in some cases attending to this shift in attention might be all that is demanded; in other cases, the demand may be to formulate the world in specific ways for theoretical or practical purposes; it may even be a purely material demand, a demand for food, shelter, clothing, with specific formulation of the world being only instrumental (even if necessary) for this purpose. It is this relationship with the other, accomplished by a personal encounter to which demand is

ubiquitous, that founds understanding of the other and of the world. We can understand aspects

of the encounter through natural pedagogy theory, and we can model some general affects that

the other has on one's attention through studying attentional dynamics – and in these ways we

can come to understand how understanding itself is produced – but these tools must recognize

the inherently ethical nature of their target phenomena.

Understanding

One of the most important consequences of a Levinasian approach to intersubjectivity,

with its emphasis on the face of the other as commanding or directing attention, is that it founds

understanding not on a perceived similarity between one and the other but, rather, on a

communication between one and the other, construed as an address – a proximity and contact –

that commands me prior to my understanding its content. As Levinas states in his interviews with

Richard Cohen, *Ethics and Infinity*, "The face is signification, and signification without context"

(Levinas, 1985, p. 86). This turns our ordinary conception of understanding, for which the

ultimate basis is the *understandum's* being situated in a familiar horizon, upside down: the first

foundation, according to this theory, is the address to the one by the other, which signifies across

the *difference* between myself and what is beyond me – indeed because of this gap – and only

once introduced to consciousness (Levinas calls the experience of the face of the other

"experience par excellence," as it "does not come from our a priori depths" (Levinas, 1969, p.

196)) does the communication begin to become understood in terms of familiarity and horizonal

structures. As is seen in the example of the crying infant, this coming to be understood in the

traditional sense, involving the situating of expression into familiar horizons, is not immediate,

and is oriented according to an ethical dimension. One implication is that the other's expression

is more than (literally undergirds and is *beyond*) the way that we make sense of it; it is more than

125

it appears to be *for me*, since the other's expression is not defined by the way that it appears to and is categorized by me, but rather in terms of the fact that the other acts upon me disruptively. (For Levinas, I am passive with regard to the encounter with the other, and even thought and understanding are considered ways that I bring my own being to bear upon invocation by the other as a form of activity. This passivity is the foundation of activity – it sets it up; receiving the expression of the other makes possible thought and action.) Another implication is that we can determine ahead of time neither what the other will demand of us nor our own response. Her expression comes from outside myself and threatens to disrupt any or all of my own projects, and to defy any principle (ethical or otherwise) I may have chosen for myself: I can be called to recognize her, but can also be called to feed or cloth; I might respond like a utilitarian, but this very relation might show the inadequacy of utilitarianism. In a sense, demand is a structural feature of expression, but one that always bears some content without being bound to any content in particular: it is a specific demand that conditions (and motivates) specific sets of response. The Other is singular and expresses herself singularly.

On the account pursued here, understanding is founded on communication, which is in the first place prior to understanding, as opposed to our understanding of the other being *first* founded on ourselves (with some traditional accounts going as far as to maintain that the other is constituted *as* an other *insofar as* they resemble me). Communication is initiated by the other with oneself only coming in second to make limited sense of that communication, i.e., understanding. Understanding stands in a dependence relation to communication, which sets it up and makes it possible. The establishing moment of communication is an asymmetric encounter in which the other addresses me and commands my attention.

Accordingly, to return to the ununderstandability thesis (UT), we can see that not only is a Levinasian approach not committed to UT, but it is actually resistant to it in a straightforward way, rejecting its implicit assumption: difference from my own experience does not commit us to ununderstandability, but rather is the foundation of all interpersonal understanding. Compare this to traditional Husserlian/Jasperian view, on which a failure to empathize with the other, that is, to approximate their own experience by what is already familiar to oneself, commits us to unintelligibility, i.e., UT.[3] On my view, we understand others across difference, not only despite but because of difference, by being exposed to an expression that introduces something new to consciousness, which is only subsequently approximated (or dissimulated) in terms of the familiar.

A prior form of understanding – we can call it ethical understanding – is the basis of understanding traditionally construed. Ethical understanding refers simultaneously to an orientation, a starting point, and a never-finished process: it is turning toward the other, having heard their cry, and realizing the demand to respond. What is understood, even if only implicitly, is the other as making some claim on me, but this understanding is precisely *not* interpretable as understanding an object. It is a personal understanding, what Levinas might call "revelation", of the other as other, which cannot be understood on the model of perception but only in terms of an ethical relationship. It is a starting point for understanding in the traditional sense, as it is only in responding to the other that thought becomes explicit, be it about the other, the self, worldly things, or the world itself. It is never-finished, since it is discourse: the ethical relationship refers to an accountability – a need to justify one's actions, including one's thoughts – *to the other* prior

---

[3] For Jaspers, understanding does indeed begin with the limit that is ununderstandability, but it is only through empathy – founded as it is on similarity rather than difference – that we come to understand. My proposal is different, since understanding does not presuppose similarity; in responding to the other, even before we know what they mean, we come to understand them. More on this in the next chapter.

to a lonely need to justify oneself to oneself, attaining rational self-satisfaction, in an abstract game of rationality detached from its interpersonal condition.

Implications

This has some implications for clinical and theoretical psychopathology. First, I disagree with positions, such as that of Giovanni Stanghellini's (2019), that take theoretical psychopathology to be the foundation for clinical psychopathology (and in this, I agree with, for example, those such as Edwin E. Gantt (2000, 2002) and Robert Kugelmann (2002)). Rather, it is in the encounter with the other who expresses herself to me that I can first understand anything at all about them, including the ways in which their experience differs from my own or from the norm. Any theorizing – any understanding, from the abstract all the way down to the most particular and concrete – is based on this encounter in which the other directs my attention (in this example, directs me to what comes to be known as their "pathology"). Psychopathological theory is grounded in encounters with the people theorized over, notably the clinical encounter.

Second, the understanding that comes out of such an encounter is particular and limited: it pertains to the other as they express themselves, the expression itself, and my limited ability to make sense of that expression based on familiarity. This limited sense means both that my understanding of the other is always inadequate, and that it does not generalize to still other others. In the face-to-face encounter with a singular other, my understanding is an ethical one insofar as my interpretation is motivated by the very demands she makes on me, rather than being motivated precisely by my familiarity with the familiar: my self-interest, my habits, norms, my own structures of subjectivity or those thematized for me by still other others. Furthermore, though I do bring the familiar to bear in making objective sense of the expression of the other – that is, in understanding what is said, the world she thematizes, the worldly implications of her

demand, and that which she tells me about herself – the sense thus produced does not straightforwardly translate to an insight about anyone else, no matter how similar their expressions might seem. Generalization might indeed be *necessary* for navigating a world occupied by a plurality of others (for Levinas, since justice requires that we consider not one other but all others, that is, that we "compare incomparables," justice indeed requires generalization (Levinas, 1981, p. 158)), but nevertheless must always be founded on and responsive to particular encounters and their unpredictable idiosyncrasies. Our generalizations must be responsive to each other whose expression is the basis of generalization and must be justifiable to each other. Otherwise, they risk being not only utterly dumbfounded, resulting from a limited perspective on phenomena not of our making for which one does not and cannot have the resources to adequately determine for oneself, but also risk being ethically unjustifiable as an appropriate response to the very author of that expression who speaks to us and in doing so founds the entire enterprise of understanding as at bottom ethical. Indeed this is another meaning of the famous "Ethics as first philosophy", and is exactly the point that politics, law, policy, and so on must always be held accountable to ethics as revealed in face-to-face encounters, or in other words, to the singular others that we aim to understand. Further, it is for Levinas to the very specific demands made by the other – as varied as all possible needs, desires, and whims, so long as we remember that possibility is not exhausted by one's own formulations – that our thought and action must orient themselves (though we can do a better or worse job of accomplishing this orientation ethically). In the context of psychopathology, this means that our generalizations – our diagnostic categories, our recommended treatments, our entire psychiatric style both in and out of the clinic – are always inadequations or dissimulations that can be

ruptured by any patient at any time, and for which the ultimate standard must always be ethics. Psychopathology must be held accountable to the patient.

Third, consequently, clinical practice, especially when it depends on such generalizations as it, arguably to some degree, must, can do a bad job of listening to the other. Indeed, when the practicing psychiatrist views the other first from the perspective of their own expertise and knowledge, the moment at which the other expresses herself across difference is buried and their expression becomes a series of too-few pieces to be fit into a jigsaw-like puzzle resembling I cannot know what but must hazard a guess. Where do I place the pieces, among those already familiar and positioned, to attain a diagnosis? *It is exactly here that lies the threat of the ununderstandability thesis: not separation or discontinuity or abnormality but, instead, over-reliance on one's own familiar horizons to the neglect of the other's expression.* Of course, this may not be what all psychiatrists do, and probably not even what any psychiatrist does all of the time, but nevertheless it is a special problem of foundational importance introduced here by reconsidering the very grounds on which psychopathology is possible at all.

Fourth, we can conceive of attention (and attention to attention) as guiding psychopathology practice and theory. My analysis agrees with Robert Kugelmann's (2002), when he writes, citing Devereux, that

> The Other who faces me and calls me to task is, according to this analysis, an exteriority that I cannot assimilate, an exteriority that I find affecting me within myself. The psychoanalytical anthropologist, George Devereux, writing on counter-transference, claims that the psychoanalyst "understands his patient *psychoanalytically* only insofar as he understands the disturbances his patient sets up within him. He says: 'And this I perceive' only in respect to those reverberations 'at himself" (Devereux, 1967, p. 301). (Kugelmann, 2002, pp. 139-40)

A concrete example of how this kind of affect recognition on the part of the psychiatrist is already performed concerns "praecox feeling." Praecox feeling, coined by Rümke (1958, 1990), refers to a specific affective configuration of the psychiatrist when encountering a person displaying schizophrenic symptomology, that is, expressing themselves in a way conceptualized as schizophrenic. The psychiatrist attends to the patient affectively, that is, they're affected by the patient's expression such as to pay attention to the way in which they are affected by that expression. Furthermore, it is the way in which the psychiatrist is affected that draws their attention to features of the encounter and aspects of the expression that they otherwise would not have noticed. This modification of attention which reveals the way that one is affected in the encounter does not totalize the other but rather indicates the difference between one and the other in such a way as to introduce to the psychiatrist something they did not already and could not by themselves contain: their affection by the other. In general, it is my position that the way in which the psychiatrist is affected by the other guides their attention, though it is not always as obvious as in the case of praecox feeling. Emphasizing the psychiatrist's affective response to the patient as its own text, and in a certain sense the only text available, institutes a distance between the psychiatrist's diagnoses (as interpretations) and the patient diagnosed, which I speculate would decrease the risk of harmful objectifications and cut one of the many legs out from under labeling stigma. To deny one's own affective responses to the other is to deny the trace of communication across the interpersonal distance that makes objectivity possible, depriving the psychiatrist of one of their only forms of contact with the patient. This contact denied, not realizing that affectivity reaches out toward the other before falling back on the self, the psychiatrist overlaps completely with himself, becoming accountable only to familiar horizons in which established theories and "objective" observations are recorded. Indeed, the real

131

distance is the incommunicable distance between psychiatrist and patient, now incommunicable only due to the denial of affectivity that records communication across interpersonal space. Objectivity cannot be better insured by denying contact (at a distance) with she who is to be understood.

Furthermore, if we understand dynamical systems theory approaches to intersubjectivity as tracking the ways in which we affect one another, but incorporate the Levinasian insight that one's reaction to the other is always limited and never exhausts the expression that institutes that reaction, then we can understand such approaches as a useful conceptual apparatus for formalizing these changes and thus explicitly and systematically drawing our attention to general patterns warranting that attention. That is, if we understand, in general (given all the caveats just mentioned with regard to generalities), the ways in which psychiatrists are affected by patients, we can produce a kind of affective-attentional schematic, emphasizing avenues of research that might productively lend themselves to theoretical and practical applications, just as praecox feeling is taken to be indicative of a certain theoretical and practical trajectory. Simply put: praecox feeling draws our attention to certain aspects of the other as she expresses herself, through the way that expression affects the psychiatrist, so as to guide research and practice; dynamical systems theory, as tracking the ways in which the other affects the psychiatrist, can generalize this structure without totalizing the other, that is, without needing to define the other exhaustively in the role she plays in the interaffective system in question.

Chapter 4

Ethics and Understanding:
From Ethical Asymmetry to Doxic Symmetry
with Levinas and Enactivism

Introduction

Levinas and enactivism have a good deal to offer one another. Enactivism began with an understanding of the organism as an autopoietic system that makes sense of its environment through its own activity oriented by a self-interest in survival or perpetuating (re-producing) itself. One track of development for this tradition has been the exploration of human intersubjectivity. It has been argued in empirical and conceptual terms that relations with other human persons are an important part of our understanding of ourselves, others, and the meaningful world. Two important and closely related approaches to social understanding in enactivism are interaction theory and participatory sense-making. Some enactivists have moved to exploring the specifically ethical dimension of human sociality. Two major attempts are *Ethical Know-How* by Francisco Varela and "Emotion and Ethics: An Inter-(en)active Approach" by Giovanna Colombetti and Steve Torrance (the latter of these works explicitly builds on participatory sense-making). Both attempt to move away from a self-centered and rationalistic conception of ethics, according to which the self relies on reflective principles in order to determine action, and to instead establish ethics on the foundation of affective relationships with others. (Due to this emphasis on affectivity and relationality, there has been a recent influx of quality work relating enactivism to care ethics, e.g., , Juan M Loaiza, 2019; Petr Urban 2014 & 2015.)

Enactivist approaches to ethics ultimately bottom-out on autopoiesis (or, more broadly, autonomy) and conceiving of subjectivity as being *for-interaction*, and accordingly fail to make sense of ethical responsibility, ethical experience in general and the experience and possibility of

altruism and self-sacrifice. By reconceiving primary intersubjectivity in terms of address – that

is, ethical expression that already implies responsibility – we can better understand the ethical

dimensions of interaction and participatory sense-making. This formulation requires us to

understand human intersubjectivity as involving *ethical asymmetry*, in a Levinasian sense.

Recognizing this asymmetry, enactivism can then provide a convincing story of how we move

from ethical asymmetry to a reciprocal give-and-take process of sense-making that finally results

in *doxic symmetry*, that is, in the establishment of mutual sense.

I argue that enactivism (and enactivist ethics in particular) cannot at present satisfactorily

make sense of ethical responsibility, ethical experience, and the experience and possibility of

altruism and self-sacrifice. I then offer a Levinasian corrective. Lastly, I demonstrate how

enactivism can help us to understand a transition that is recognized but underexplored by Levinas

which is crucial for understanding how we actually come to understand each other and act

ethically: a transition from *ethical asymmetry* through *reciprocity* to *doxic symmetry*. First, I

should say something about what I mean by responsibility, ethical experience, altruism and self-

sacrifice.

Responsibility

My use of the term "ethical responsibility" is not a particularly special one: I mean what

we ordinarily consider to be responsibility, that is, a kind of accountability that directs our

actions according to *should* or *ought* and is bound up with practices of blame, praise, and

accountability in general. I do not mean causal responsibility, as these are clearly distinct: one

can be causally responsible for something but not ethically responsible and, perhaps more

controversially, one can be ethically responsible for what one is not causally responsible for. One

example of the latter would be future-directed responsibility, in which I can be held responsible

for the future even if I am not responsible for the past. For example, I may be held responsible for helping to fight our current climate crisis even if I really had nothing to do with bringing it about. Levinas makes the distinction between ethical and causal responsibility even stronger, insisting that in ethical responsibility we are responsible for what is not our own deed, responsible even for the other's responsibility (Levinas, 1985, p. 96). In terms that should be familiar from previous chapters, the idea is that the other calls us to responsibility by their very expressing themselves to us; expression is a demand for some response or other, and for what Levinas calls the "justification of being". The other who is hungry, for example, may call me to feed him, and irrespective of whether I myself have brought about his hunger, I must now offer some response to this demand (we may choose to ignore the other but there is no response – even ignoring the call – that is somehow not a response at all; we may choose how to respond, but not whether we respond). Against phenomenology as an exploration of the "I-can", against existentialism in which I am responsible because I am free, and against any ethics based on any philosophy of freedom, Levinas states that "The freedom of another could never begin in my freedom, […] The responsibility for the other can not have begun in my commitment, in my decision" (Levinas, 1981, p. 10). (Indeed, a crucial point in the following is the inversion of an existentialist formulation: we are free because we are responsible, according to Levinas, rather than being responsible because we are free.)

One crucial aspect of responsibility for my purposes here is that responsibility must involve a dimension of being responsible *for the other*, and even despite ourselves. It has long been pointed out by ethicists that ethics is not necessarily easy, and that if what we should do was simply a matter of doing what benefitted us or of always deriving some private benefit from action, then ethics would collapse into a selfish egoism and relativism. Utilitarianism attempts to

135

counter this position by appealing to the greatest happiness *for all*, weighing the happiness of others as much as my own; deontology maintains that personal gain does not decide ethics, but instead appeals to a universal reason that takes consideration of the value of all rational beings. It is thus a defining feature of ethical responsibility that we are responsible not only for and to ourselves but also for and to others, even despite ourselves: an ethics that cannot make sense of the denial of my own interests is no ethics at all, but rather a Darwinian moral psychology and anthropological relativism that threatens a collapse of the *ought* into the brute existence of selfish inclination.

Experience, Altruism, and Self-Sacrifice

The notion of responsibility as responsibility *for others* pivots into questions of ethical experience and the possibility of altruism and self-sacrifice. Levinas is often taken as providing a kind of phenomenology of ethical experience. It is my position that this is certainly part of what he is doing, but only part; a discussion on this point is outside the scope of this work. But we do not need Levinas to make some very basic points about ethical experience; everyday experience is demonstrative. When we experience the other, we experience them as mattering independently of ourselves. That is, we experience them as having interests irreducible to our own and, sometimes, as (legitimately) calling our own self-interest into question. We are familiar with the experience of being called to justify ourselves, as well as being called to respond to others, and we are also familiar with the experience of our justifications and responses not being enough, even if we do not understand how or why. We offer justifications of our own actions and are surprised and even confused when these justifications are rejected; we may even respond to the poor other truly as best as we can and yet understand that this is still not enough. Furthermore, and returning to Levinas only for his presentation of an elementary fact of ethical experience, he

136

sums up his entire work as the possibility of a primordial "after you, sir", in which we put ourselves literally aside to allow another to pass before us. This is a simple and familiar point that demonstrates the possibility of altruism, that is, of acting in such a way as to put the interests of the other ahead of our own. A cynic might say that it brings us pleasure to perform such acts, but this is often foreign to our immediate experience of the situation and our own motivations, requiring speculative psychology, e.g., that we are motivated solely or primarily by pleasure, to justify itself; such acts are often not pleasurable and may even be irritating: we do not prefer to wait, we do not prefer thankless jobs, and we may even complain that no one says thank you anymore. Yet, even anticipating pain, we persist in acting despite ourselves in service of the other simply because it is good. And were altruism revealed to be selfish, we would only have thereby discovered a fact; the good is greater than being, Levinas maintains, because it is *better* (e.g., Levinas, 1981, p. 19). That is, even if we are selfish, goodness calls us to be better. I take self-sacrifice to be a kind of extreme altruism, action *for-the-other-despite-oneself* even unto death or destruction. Self-sacrifice can be the literal sacrificing of one's own life for the benefit of the other but can also be – when we speak in interactional terms – the destruction of an interaction or relationship when that destruction promotes the interest of the other precisely by putting an end to one's own interest. That is, though an interaction may benefit myself, self-sacrifice includes the possibility of terminating that interaction for the good of the other despite myself.

Autopoiesis

Enactivism maintains that sense-making is a cognitive-affective activity undertaken by embodied organisms embeded in their environment. In classical versions of enactivism, or "autopoietic enactivism" (Métais & Villalobos, 2020), the organism is described as autopoietic.

Importantly, there have been some recent attempts to "supplant" the concept of autopoiesis: on the one hand, some rely instead on autonomy and adaptivity as more generally applicable for defining organismic life (see Corris & Chemero, 2019), while other approaches, characterized as 'sensorimotor enactivism', propose a less central role for autopoiesis. The objections to be raised in this section apply straightforwardly to the first alternative (replacing autopoiesis with autonomy perhaps sharpens the point); the sensorimotor enactivist approach is an interesting alternative, but one I cannot consider here.

Autopoiesis refers to the way in which the organism is self-re-producing or self-maintaining. This concept is attributed to Maturana and Varela (1973) who specifically consider cellular autopoiesis: cells maintain a permeable membrane that distinguishes inner from outer and their inner states are adjusted responsively to the environment as well as to the overall state of the cell itself (there is thus a dynamic or circular causality: the parts determine the whole *and* the whole determines the parts). Evan Thompson (2007, p. 124) also notes that Maturana and Varela further propose that "all living systems are autopoietic systems and that all autopoietic systems are cognitive systems". Thompson takes up this idea in the appropriately named *Mind in Life*. In exposition, he writes:

> According to Maturana and Varela (1980), the relation between autopoiesis and cognition has two crucial features. First, the instantiation of the autopoietic organization in an actual, concrete system entails a cognitive relation between that system and its environment. Second, this cognitive relation reflects and is subordinated to the maintenance of autopoiesis.
>
> (Thompson, 2007, p. 124)

These are indeed crucial features.

First, cognition becomes a matter of embodied and embedded action – hence the name 'enactive cognition.' Enactivism maintains that in order to re-produce themselves, organisms sensitively respond to features of their environment as well as their own states. Cognition thus becomes understood precisely *as* this kind of sensitive activity, constituting an embodied understanding of oneself and one's environment, specifically an understanding that is value-laden. To use the most famous example, a bacterium swimming in a sugar gradient swims towards higher concentrations of sugar precisely because sugar has a value as *food* for the organism (Colombetti, 2017). Of course this is not what we typically understand as cognition, e.g., high-level thought processes or abstract reasoning (the question of transitioning to those processes from the low-level ones proposed here generates what is known as the "scaling-up" problem); instead, enactivists hold that cognition is, at its most basic, this kind of value-laden, sensitive activity. Cognition is accordingly rebranded as "sense-making", e.g., the bacterium actively makes sense of sugar *as* food (note here that significance, as argued previously, has an *as* structure, at least when described from the outside: for the bacterium that enjoys sugar, there is likely nothing similar to the high-level, propositional understanding often associated with human cognition; yet there is understanding in the form of the sensitive activity that makes sense of the sugar in terms of its value as food: understanding as an active, cognitive-affective relationship. Note the close association – perhaps even identity – between significance and value). It is this value-laden aspect that Giovanna Colombetti refers to as "primordial affectivity": it is by affecting and being affected by an environment that *matters* to the organism (specifically, to its self-re-production or survival) that the organism makes-sense of things. The bacterium thus understands, cognizes, or makes-sense of the environment: it organizes its own activity in response to an environment that matters to it (specifically, to its perpetuation). This is

of course an extremely simple example, but nevertheless demonstrates that a world – the environment as *meaningful* to the organism – is enacted by the organism's own, value-laden activity.[1]

Because sense-making is oriented by what matters to the organism, precisely in terms of self-re-production, cognition is accordingly *subordinated to the process of autopoiesis*. Insofar as affectivity is a primordial aspect of sense-making – and indeed, many enactivists hold that cognition and affectivity can be separated only abstractly, if that – then affectivity, too, is subordinated to autopoiesis. Autopoiesis as a fundamental drive or structure of sense-making invites comparison between enactivism and Spinoza: we could say that the essence of the organism (as is true, for Spinoza, of any finite mode) is its conatus, i.e., its striving to persist in being.

Enactivism maintains that the organism is characterized by autopoiesis (or autonomy and adaptability). The principle of autopoiesis maintains that cognition = sense-making = activity oriented according to the organism's capacity and drive to perpetuate its own existence.

Such an account subordinates sense-making to self-interest and, accordingly, cannot make sense of ethical responsibility, ethical experience, or the experience and possibility of altruism and self-sacrifice. Specifically, if sense-making *is* activity, and activity is oriented by perpetuating one's own existence, then there is a challenge for making sense independent of my own self-interest. Fortunately, however, enactivism does not stop here, because the individualistic account is clearly inadequate on enactivism's own terms, i.e., it is an inadequate account of sense-making: we do not understand the world only in terms of ourselves, or else we would collapse into the kind of solipsism warded off by Husserl's appeal to the constitution of a

---

[1] This position is largely congruent with Jakob Johann von Uexküll's formulation of the *Umwelt*. (In *The Open: Man and Animal* (2003, p. 46), Giorgio Agamben paraphrases von Uexküll's discussion of the Umwelt of a tick.)

world precisely as *that which is available to others* (see *Chapter 2*). Enactivism recognizes the phenomenological fact that we make sense of the world together. Accordingly, there is an intersubjective turn in enactivism.

Intersubjectivity

At least for human beings, sense-making cannot be solely a matter of our own activity. It has been shown, both empirically and conceptually, that other human beings play a crucial role in our understanding of ourselves, others, and the world (in addition to the positions already discussed, particularly in *Chapters 1* and *2*, see Piaget, 1995; Hegel, 1977, Carpendale & Lewis, 2004, and much of the work in the enactivist tradition). Shaun Gallagher, for example, builds on Trevarthan and Hubley (1978) and Trevarthan's (1979) concepts of primary and secondary intersubjectivity to explicate the ways that we come to understand one another and the world (Gallagher, 2009). Primary intersubjectivity refers to our basic awareness of other people as such and, according to Gallagher, this basic awareness necessarily involves interaction: it is by interacting with others, rather than observing them with indifference, that we relate to them as genuine others, and accordingly come to understand them as others.[2] Secondary intersubjectivity, on the other hand, describes the ways that we understand others through our shared relations with the world. In secondary intersubjectivity, we begin to recognize features of the world along with enhancing our understanding of others. Extending the argument of the previous chapter, we can easily see how this is largely a matter of joint attention and joint action: we can understand the other, in part, by directing ourselves towards what the other herself is directed towards. It is not

---

[2] Gallagher (e.g., 2009) supports this point by considering studies on infant interaction but argues that primary intersubjectivity is not simply an infantile, developmental stage to be overcome: rather, interaction remains the primary way in which we understand others throughout our lives, and thus interaction characterizes primary intersubjectivity.

simply a matter of perceiving the object that the other attends to, however, but also a matter of

joint action, as it is through action and interaction that we come to understand those objects.[3]

De Jaegher and Di Paolo (2007; also see De Jaegher, Di Paolo and Gallagher 2010) build

on this work, but also aim to go further by demonstrating precisely how interaction itself can

become primary: interaction theory, for them, is still self-centered insofar as it conceives of

interaction as that through which an individual comes to understand another individual, and

particularly in terms of agentive intentions (e.g., X intends (to do or accomplish) Y).[4] They take

this further by emphasizing the way in which interactions often take on a life of their own,

gaining their own kind of autonomy – interaction-autonomy – that emerges from but is

irreducible to the autonomy of the interacting agents, and which acts as a condition on those

agents. We can already see here a parallel to the autopoietic structure in which the parts

determine the whole but, reciprocally, the whole also feeds back on the parts, constraining and

adjusting them top-down. In a dyadic relation, for example, the dynamics of the interactions

themselves take on the characteristics of an emerging autotpoietic system, and therefore takes on

an autonomy that transcends either individual participant. It is this surplus of autonomy beyond

the autonomy of the participants that is productive of new meaning; rather than meaning being

something that pre-exists an encounter, and simply recognized through interaction; it is instead

produced by that very interaction. Since interaction-autonomy is in excess of individual

autonomy, meaning and social understanding can be said to emerge from individual

understanding yet nevertheless be something beyond it. It is perhaps for this very reason – the

---

[3] Eye-tracking research is helpful here, demonstrating how social understanding ordinarily involves following the other's gaze to objects in the environment. Furthermore, certain cases of pathological disruption to social understanding involve disrupted gaze-following (for an overview, see Koster & Fang 2015; specific sources listed in *Chapter 1* discussion of empathy deficits in autism research).

[4] I cannot here get into the question of whether De Jaegher and Di Paolo's is a *fair* characterization of interaction theory.

surplus of meaning and beyondness indexed to interaction itself – that Gallagher attributes a certain *transcendence* to interaction itself (Gallagher, 2014).

We can conceive of the idea that interaction takes on a life of its own almost literally, that is, by understanding interaction as constituting a higher-level autopoietic system. Importantly, there is a source of *value* in interaction-autonomy (as we will soon explore), since certain things *matter* for the continuation of interaction. Of course, interactions are in a sense more frivolous than organisms: they tend to be more short-lived (relative to the relevant organismic timescale) and only perpetuate themselves so far. In a certain sense, interaction-autonomy is still *answerable to* individual-autonomy: if individual autonomy is undermined, De Jaegher and Di Paolo (2007) tell us, the 'interaction' ceases to be a genuine, social interaction. On the other hand, some interactions do perpetuate themselves (or, perhaps more precisely, condition their interactors so as to perpetuate the interaction): we can think of highly stable and self-referential deeply embedded social institutions such as the economy or legal system (see Gallagher & Crisafi 2008), as well as smaller scale phenomena such as individual (e.g., romance, friendship, business) relationships (and the norms surrounding them) which condition individuals towards the maintenance of those relationships for better or, too often no matter how often, worse.

For the sake of clarity, and also because it will be of value later-on, I want to note two oft-cited mechanisms by which participatory sense-making operates to produce social understanding: imitation (or mimicry) and dynamic coupling. In imitation, we understand what others do by ourselves doing what they do. There is a rich empirical literature on imitation and understanding, particularly regarding neonate imitation, that is of growing interest in philosophical cognitive science. Imitation could be seen as one form of dynamic coupling, but a very specific one in which what I do agrees with or even simulates what is done by the other:

143

they smile → I smile → I understand something about smiling. Imitation as an approach to social

understanding may inherit some of the problems of simulation theory, but we can set this issue

aside (see e.g., Gallagher 2020 for more on problems with ST). Beyond imitation, dynamic

coupling more generally requires a kind of *attunement* of action between one person and the

other (imitation being only one form of attunement) such that one's actions are somehow

appropriate responses to the actions of the other. To borrow an example from Dan Zahavi, I

might understand my neighbor's fury not by imitating it but by responding with fear (Zahavi

2010). In this sense, interaction can be characterized by reciprocity as a kind of give-and-take

relation. Through dynamic coupling, we come to understand the intentions and states of the other

because those intentions and states are inseparable from the activity with which we are coupled:

they are part of the other's activity which is entangled with our own; insofar as we act in

response to the other and the other does the same, we arrive at insight into the total nexus of self,

other, and world.[5]

In turning towards intersubjectivity, enactivism promises a move away from the focus on

individual autopoiesis or autonomy, that is, from self-interest, to an approach that appreciates not

only the ways in which others affect us but also the ways in which our very relations to others

condition and constrain us both. Though interaction-autonomy depends on and is responsive to

individual autonomy (that is, when an interaction is guided by individual interests), it is not

exhausted by it; due to the partial independence of interaction-autonomy, we may end up with a

result that is broader than or even in partial conflict with our own interests: collaboration may

produce something greater than solitary labors, but may also involve compromise. This gets us

closer to understanding the possibility of acting against our own interests. However, it does not

---

[5] This might be seen as a cognitive expression of the observer effect in physics.

yet get us to understanding responsibility, altruism and self-sacrifice in the sense of acting against our own interests *for-the-other*, but rather only *for-the-interaction*. Our only relation to the other is here mediated by the interaction itself, and that interaction exceeds both myself and the other. I therefore relate to the other by relating to a third term that does not adequately represent the other's interest. To put it another way, participatory sense-making overcomes the subordination of sense-making to self-interest by subordinating self-interest to interaction-interest, which is always an inadequate composition of the other's interest *and also my own*. To act in the interest of the interaction is, even in compromise, to act at least in part for my own interest; and to act against the interaction – to bring about its end – can only be made sense of by reference to my own interest or the interest of the other *as related to me in interaction*. Dierckxsen (2020) makes a compelling case that our responsibility to others extends beyond any possible interaction or participation; indeed, we can hear the call of the other and respond to it without ourselves ever being recognized by or conditioning the actions or experience of the other, that is, without dynamic coupling. This ethical dimension of sociality is overlooked by interaction and participation and exceeds it.[6]

More precisely, what is missed is the fundamental asymmetry of my relationship to an other. The difficulty with enactivist sense-making, and specifically with participatory sense-making, is that on these views values and norms arise out of interaction, where interaction is construed as a coordination of my action and the action of the other. The problem is that, in conceiving of value as a product of interaction, we fail to recognize the independent value of the other, absolutely distinct from myself and from any relationship (including interaction) that may

---

[6] Dierckxen (2020, p. 111) writes: "ethical relations transcend natural relation in the sense that responsibility for others requires more than natural participation. It requires a sensibility for otherness which cannot simply be understood in terms of natural participation or failed natural participation, because it goes beyond that which any kind of participation could possibly achieve."

hold between us. The expression of the other reveals to me the other as independently valuable, independent even of that very revelation. In terms of responsibility, I am responsible for the other whether I recognize it or not, and prior to any activity or autonomy on my part. It is for this reason that Diane Perpich, for example, can talk about ethics in terms of "normativity without norms" (2008, p. 124) (though in this context, I would say *prior to* norms): it is the independent value of the other that motivates interaction and constitutes at least part of the binding force of the norms thus produced. Responsibility for the other is not a product of my own action, even in coordination with others, but rather characterizes the basic relation prior to and even despite my own action and interest. As such, enactivist sense-making misconstrues responsibility, as well as the possibilities of altruism and self-sacrifice, since these are all made to depend on coordinated activity at least partly invested by my own interest.

An important objection can be made that only through interacting with others, we come to recognize *their* interest, and thus can act on that interest specifically; such an objection might leverage primary intersubjectivity construed as a *direct* relation to an other. But if understanding is one with activity, and my activity is always my own – no matter how conditioned it may be by the other and our interaction – and if the interest of the other is also conditioned by our interaction, then my understanding of the other is always suffused by self-understanding and self-interest. This is not at all a description of facts, but rather of the resources available to an autopoietic understanding of interaction that relates parts precisely through mediation by the whole. Another way to make the point is in terms of asymmetry, which is rejected by Gallagher (we return to this point later): according to Gallagher, ethical meaning depends on mutual recognition, i.e., recognizing the other precisely as recognizing me. Yet, it is not at all obvious that ethical responsibility requires that the other recognize me nor, further, that I recognize that I

146

am recognized by the other. All it seems to require, from the perspective of Levinas [?], is that the other express herself as an other, and perhaps that I hear that expression. As soon as interaction is conceived of as symmetric (and not just reciprocal; again, more on this later), recognition of the other is already bound up with self-recognition and therefore with self-interest. As we will see, this situation is a real and common one, but it crucially neglects the fact of experience, that we can recognize the interests of the other before interaction, and that we can act on those interests without incorporating our own.

Another difficulty can be drawn from Levinas' concerns with Buber's I-Thou relationship. He charges the I-Thou relationship with formalism and claims that the relation "can unite man to things as much as man to man."

> The I-Thou formalism does not determine any concrete structure. The I-Thou is an event (Geschehen), a shock, a comprehension, but does not enable us to account for (except as an aberration, a fall, or a sickness) a life other than friendship: economy, the search for happiness, the representational relation with things. (Levinas, 1969, p. 68)

In one sense, participatory sense-making is indeed a formalism: it posits a formal structure that characterizes interpersonal relationships in general. But supported by phenomenological and empirical concerns, enactivism explores concrete, contentful relational dynamics on the basis of these insights. As such, the formalism charge doesn't quite stick.

However, the charge that the relationship can characterize man's relation to things just as much as to one another is appropriate. Using the notion of economy, we can understand enactivist interaction as transaction: a reciprocal exchange resulting in an independent third term, the interaction, an economy, in virtue of which the individual contributions gain their significance. The representational relation here is not representation in the sense in which the

147

term occurs in most current debates but instead refers to the fact that the relationship to the other is both mediate (mediated *by* the interaction/interaction-autonomy/coordination of activity) and inadequate. The search for happiness indicates that our own interest is wrapped up in the economy that determines the meaning of self and other: the interaction may exceed my interest but is nevertheless invested by it.

Objects stand in dynamic relations to us as well. Indeed, a standard articulation of the extended mind hypothesis maintains that objects can count as proper parts of cognitive systems precisely because they are dynamically coupled to states we ordinarily consider cognitive, that is, mental states traditionally conceived, i.e., memory, judgement, intentions, and so on (e.g., Chemero, 2009; Clark, 2008; Clark & Chalmers, 1998; Piredda, 2017; for the famous objection that dynamic coupling is distinct from constitution, Adams & Aizawa 2010).[7] There is also a growing literature on how affective states are extended qua coupling between objects and affective states traditionally conceived, e.g., between a musician and their instrument (Colombetti & Roberts, 2014), or a theater play and its viewer (Slaby, 2014). In short, we dynamically couple with objects: we act upon them, they resist through their own active productivity or power, this changes how we respond in turn, so on and so forth until the system formed between us is irreducible to either of its parts. We witness this in art, science, literature – any human endeavor. The very notion of phenomenological intentionality – inherited at least as inspiration by enactivism – is itself a notion of a dynamic, mutually conditioning and constraining relationship between subject and object.

---

[7] Clark (2008) states that coupling is not sufficient for extension, and that what is further required is a "self-stimulating loop." Specifically, the system must, through the coupling of its parts, self-stimulate or augment itself. That is, there must be effects within the system irreducible to any of its parts. This is only a sidenote for us because he and others go on to argue that relations between persons and objects can satisfy this condition.

What distinguishes the other from objects, in experience? Husserl distinguished the other as another ego like myself, but distinct from myself, and this distinctness largely amounted to the non-presentability of the other, although, how this *principle* of non-presentability, i.e., the principle that the other is never given originarily, emerges in experience, and in such a way as to distinguish between absent other-presentations and absent object-presentations – since, after all, objects are also never given "from all sides" (Husserl 1999, p. 27) – is unclear; it is plausible that there need be some *positive* aspect to this lack that motivates the principle, some positive aspect to the experience of the other that distinguishes such an experience from object-experience. Or else the distinction depends more crucially on similarity – the other, unlike the object, is *like me*, and also non-presentable (like the object) – but similarity has already been called into question. Contemporary phenomenologists and enactivists have pushed back against – or at least deflated – this non-presentability, without moving beyond its focus on the negative (the non-) to its positive significance (e.g., Smith, 2010 maintains that others are given non-originarily in the same way as are occluded sides of a given perceptual object; Krueger, 2018 maintains that certain aspects of the other's experience are not given to us originarily, but that others, namely the bodily aspects of emotional states, are given to us in the same mode as are objects, i.e., perception). Levinas articulates the positive significance (or presence) of this non-presentability that distinguishes the other from objects in experience (and not just in terms of what does *not* appear in experience) as ethical responsibility, which is present as an asymmetry. It is as responsible to and for others that they, others, are positively distinguished from objects in the first place.[8] The other presents to us an *ethical resistance*, not only disrupting one's action as an

---

[8] Of course, we can note that people sometimes personify objects, even to the point of feeling responsible to these objects. My goal is not to argue that such an error is impossible, but rather to point out that it is precisely responsibility revealed by expression that constitutes otherness phenomenally: it is only in taking the object as making demands on us that the object is personified. We can adapt Husserl's notion of harmonious perception (or continued confirmation) in

opposing force but, rather, as a demand for the justification of one's action; the other demands an ethical response, rather than a modification of my appropriative acts so as to overcome objective resistance.[9]

The claim that autopoietic enactivism subordinates self-interest to interaction-interest may seem abstract, but we can now pivot to enactivist approaches to ethics and see how, for these approaches, responsibility is formulated as responsibility-to-and-for-interaction, rather than to-and-for-the-other, and thus misses or misunderstands fundamental ethical explananda. There is nevertheless something right in these approaches and, later on, I will show that everything we have explored up to now is not irrelevant but rather can be productively modified to help us better understand ethics. Specifically, once we recognize the asymmetry of social interaction, we can not only make sense of responsibility, altruism, and self-sacrifice, but we can also see – through enactivist eyes – how the reciprocity of interaction establishes mutual sense and norms.

Ethical Know-How and Affectivity

The embodied, embedded, enactive conception of sense-making, with its move away from reflective rationality and towards embodied action, naturally leads to a prioritization of *know-how* (as opposed to *know-that*). A number of thinkers have taken up know-how as a central notion for understanding how we should live and what exactly we are doing when we live well (see, for example, DeSouza (2013), and, for a criticism, Dierckxsens (2020)). The relationship between know-how and ethics is the major theme of the appropriately titled *Ethical Know-How*

---

helping us to distinguish actual others from mere objects; the point is simply that responsibility is the difference that makes the difference.

[9] Diane Perpich (2008) provides a discussion of this topic. Though objects resist us, we still recapture them according to an idea. The mind reasserts its spontaneity over the object by constituting it for itself. The other, however, disrupts my spontaneity in a way that is not recuperated. She appeals to Levinas' consideration of the idea of infinity in Descartes. I can be the cause of my representations of finite modes, but not my idea of infinity; the idea of infinity is instead produced in contact with the other, exceeding every idea: "infinity does not enter into the idea of infinity" (p. 58).

(1992) by Francisco Varela. Varela associates know-what with knowledge and rationality, and know-how with wisdom and spontaneous, skillful action. He embraces a kind of virtue ethics, based on Confucianism, Taoism, and Buddhism, that focuses on developing certain abilities (he calls these, following Mencius, extension, attention, and intelligent awareness) to act wisely when it comes to conducting oneself practically in the world and to serving others through one's actions. In short, what we extend is our own understanding of a given situation; what we attend to is the situation itself and what *it* demands of us; what we are intelligently aware of are salient features of our own embodied and embedded activity.

Varela recognizes the importance of responding to the needs of others for ethics; an ethics built on one's own well-being and self-service is no real ethics at all. He goes as far as to claim – in a way that should half remind us of Levinas – that

> What we call "I" can be analyzed as arising out of our recursive linguistic abilities and their unique capacity for self-description and narration. […] If this narrative "I" is necessarily constituted through language, then it follows that this personal self is linked to life because language cannot but operate as a social phenomenon. In fact, one could go one step further: the selfless "I" is a bridge between the corporeal body which is common to all beings with nervous systems and the social dynamics in which humans live. My "I" is neither private nor public alone, but partakes of both. And so do the kind of narratives that go with it, such as values, habits, and preferences. In purely functionalist logic, "I" can be said to be *for* the interactions with others, *for* creating social life. (Varela, 1999, pp. 61-2)[10]

---

[10] There is apparently an affinity between this account and that of discursive psychology, specifically as presented by Gillet and Harré (2013), according to which "I" exist as a kind of intermediary between two worlds, one of traditional causal ontology and the other of human social life.

I say half because the I is, on this picture, neither for itself nor for the others, but instead for *interactions* with others. Here we see a kind of responsibility, a being-for, directed towards the interaction itself. We can see the entire argument concerning autopoiesis and motivation above spelled out here in explicit connection to ethics: a move from individual autopoiesis and self-service to a relational account that privileges the life of the interaction itself – a relation considered not simply as the way that two beings relate, but as itself a kind of third being which exceeds us both. The focus on self-concern is brought out by the fact that the "I" on this account is really an intermediary between the organism's low-level parts – neuro-biological facts concerning brain or hormonal states and patterns, for instance – and high-level "constraints imposed by global coherence on local interactions" (Varela, 1999, p. 61). The virtual self acts as an *interface* (Varela, 1999, p. 61) between the organism's parts and overall structure; it refers to the double or circular causality from low to high and high to low that characterizes the organism as a dynamic system.

Two questions. First, does the self get out of itself on this account? That is, if the self is for interaction, and interaction is the introduction of higher-level social behavior that must then be squared with lower-level organismic facts, can this help us to understand responsibility to and for others? Or, rather, does it only allow us to understand the "I" as self-serving, in this case serving to make coherent the relationship between my own higher and lower level processes, including those processes of mine constrained by social interaction? For Varela, the "I" is a social product in the sense that our self-organization is in part an "ongoing interpretative narrative" between lower and higher levels. As such, it is language dependent and so depends on relations to others. But on this account, the object(s) of the narrative are precisely *my processes,* or *me*, and there arises a question of whether the other becomes a mere causal input to a self-

serving being, a constraint on my self-organization. If the "I" refers to an organismic process by which I come to coincide with myself – by which I regulate myself – then is it possible to make sense of action against self-interest, against my own freedom, against egoism?

The second question, raised by Varela himself, is whether his account of the virtual self can resolve conflicts between scientific explanation and experience. On his view, science *seems to* (but perhaps does not really) show that consciousness can exist without the self, yet in experience we seem to have a persistent sense of self as a "top-down, centered, globally directing" (citing Dennett) "centralized center or agent" (Varela, 1999, p. 60). His view is an attempt to follow a middle path between these extremes, maintaining that there is no substantive, persistent entity, "the self," but nevertheless our own self-organization results in an interactional process that explains experience of the self: it is the "two-way movement between levels" that explains our experience of the self, in terms of describing, narrating, and accounting for this very movement. But I want to extend this question: does the virtual self account for our experience not only of ourselves but of others? Does navigation between my own lower and higher-level processes make room for the other as anything but a constraint imposed on self-organization?

Varela's answer to the first question appeals to a pragmatic approach grounded in Eastern meditation traditions and, specifically, to the association between emptiness and compassion: again in a double (or circular) causality, it is by cultivating compassion that we come to experience the self as "empty" (non-substantive, non-persistent, non-central) and it is in overcoming oneself that we cultivate compassion. Emptiness then is not a concept of nothingness or absence but rather refers to a "full" emptiness, a non-substantial yet affective orientation towards others. This "full emptiness" is seen in his definition of the virtual self as

a coherent global pattern that emerges from the activity of simple local components, which

seems to be centrally located, but is nowhere to be found, and yet is essential as a level of

interaction for the behavior of the whole. (Varela, 1999, p. 53)

It is the self as virtual that is meant to resolve the tension between science and experience, as per

the second question: the self is not to be found *anywhere* but instead is an interactional level

between those states or processes described by traditional science.

Ultimately, Varela's goal in this work is to understand how we can foster spontaneous

ethical action. To do this, he must overcome powers of the self that bind it to reflective principles

– reason, knowledge – and call the self into question. He does so by recommending a pragmatic

approach: meditative practice that discovers the emptiness of self and cultivates compassion.

Having displaced the self, he goes on to explain our sense of self in terms of a virtual self, that is,

a global pattern that emerges from simple local neuro-behavioral states. Subjectivity, thus

emptied of substantiality and persistence and instead emerging within and adjusting itself to

arising situations, becomes conceived of as *for*-interaction.

But if it is the self that calls the self into question, then have its powers been overcome?

Is it not the self-same entity that simultaneously demands and provides its own justification?

Does not the appeal to serving others refer back to the appeal of self-service, of the cultivation of

virtue, to the improvement of oneself? And is compassion at all a sufficient motivation for wise,

spontaneous, ethical action that opposes itself to rationality and knowledge, that is, to pre-

fabricated principles, or is compassion not just a new, affective principle, but a principle all the

same? Does it not, just as much as reason, depend on something pre-existing the specific

encounters with the others who calls us to service, that is, depend on something already existing

in me, whether that be a rational or affective capacity and disposition? Do the other, the social

interaction, the interaction between levels, and compassion always agree in their demands? If they do not, then who does ethics prioritize? Who should we prioritize? Virtue ethics clearly maintains a preference for the self – here, for the self's own interactional dynamics and affective cultivation – even if that self is virtual, this is still a self-orientation – but I maintain that ethics must uphold the priority of the other, or else it collapses into self-interest, perhaps a productive, pro-social self-interest, but action guided by one's own interests nonetheless. That is, without prioritizing the other *over* interaction, we fall into cynicism of altruism and self-sacrifice as we fall back onto principles prior to encounter.

Does the virtual self, a kind of narrative unity of change over time, account for (or, nonequivalently, contain) all there is to say of ethical experience or, by extension, the totality of our experience of the self? Is not part of the sense of the self, particularly if we are interested in ethics, a matter of our responsibility and accountability over time? We are responsible for far more than what we ourselves call into question, far more than we ourselves see a need to justify or account for. We tell stories about ourselves but others tell these stories too, and it is not obvious that our own have priority or privilege. The question of the individuation of the self is answered by Varela through a story of changing organizational dynamics carried out at a level connecting neuro-behavioral states and narrative. But our narratives are not only or primarily for ourselves; they are always offered to another with whom we speak: we offer them to the other. Varela recognizes this and calls the 'I' linguistic and social. But any good narrative must relate to its audience. When we tell stories of ourselves that interest only ourselves, concern only ourselves, satisfy only ourselves, we are rightly rebuffed or ignored. The individuation of the I as a narrative unity and as a creature held responsible over time for those actions corresponding to disparate neuro-behavioral states to which I now only bear, as my very existence itself, a

155

developmental relation of adjustment (see, for an example, predictive processing accounts), comes from the other who demands an act of justification: we are held accountable for organizing and presenting ourselves, not only by a self-interest that could only make sense of the good as the good-for-me (even if in the form of the good-for-an-interaction-that-includes-me), but by the interests of others who question and invest my interests.

More simply, there is a fact of experience that Varela's account of the virtual self as a form of self-overcoming does not acknowledge or explain. The other is revealed to the self through compassion – a form of orientation towards the other – but compassion is a cultivation of self and part and parcel with calling *oneself* into question. Varela states that the beginning of mindfulness/awareness practice meant to cultivate compassion is a certain freeing of oneself, resulting in the development of an interest in others: a warmth, or compassion, as an other-directed interest (Varela, 1999, pp. 66-7). Compassion thus becomes a central part of the pragmatics of ethical expertise, that is, the cultivation of one's own ethical dispositions. Yet, compassion is developed by the self and, even if directed to others, is still the self's interest in others, an interest produced by overcoming one's own other interests. There is a datum in experience that Levinas indicates by expression and responsibility, the experience of being called into question not by oneself but by the other. It is not our self-overcoming, then, that explains our experience of self (at least not in its entirety), nor is self or other disclosed in an interaction in which I bring my own developed powers – my cultivated expertise, or virtue, as capacities to extend, attend, and become intelligently aware – to bear. Rather, the self precisely as ethical is experienced as that which is called into question by the other. It is in the face of this demand for justification that we construct and offer ethical justifications (that is, that we regulate our holistic social behavior with individualistic neuro-behavioral states). Compassion and the emptiness of

self, then, should not be considered merely a development of my own cognition and cultivation of and by myself but rather refer to an other who calls me to compassion, who calls me to empty myself.

Varela's account recognizes the importance of overcoming the self in the pursuit of ethics, that this overcoming must be an overcoming of principles that determine the situation in advance, and he recognizes in affectivity a potential site of resistance to self and principle. But this account raises a number of questions that must be answered if it is to achieve its goals. As presented, I believe that it fails: the attempt to overcome the self still relies on the self, that is, the self is called into question by itself, jeopardizing the selflessness and overcoming of principle that ethics is supposed to imply. As such, it is still a form of egoism and still does not make sense of either ethical experience or self-experience, which must account for the ways in which we put ourselves aside in service of the other (that is, it must make sense of altruism and self-sacrifice), and of how we experience ourselves as responsible over time. But the question of compassion as an affective direction to the other starts on the right foot.

Emotion and Ethics

Giovanna Colombetti and Steve Torrance (2009) draw a close connection between ethics and affectivity, specifically emotions. They use this approach to show a kind of responsibility that exceeds personal responsibility and in doing so they implicitly demonstrate a kind of motivation that does not refer primarily to one's own persistence. According to Colombetti and Torrance, emotions are inherently ethical insofar as emotions – and affectivity in general – are value-laden: for the enactivist, affectivity is a primordial mode of sense-making in which the organism makes sense of the environment (enacting a meaningful world) in terms of what *matters* to it. They transition quickly from this claim that emotions matter to a claim that

157

emotions are, therefore, ethical in character, the idea seeming to be that what one should do, as well as what one is responsible for, depends on what matters.

Before continuing, I want to tie this claim into the discussion of solipsistic hedonism from *Chapter 2*. For Levinas, the organism indeed makes sense of the world in terms of what matters to the organism. This fact defines the first *morality* (which I distinguish from ethics: the former deals with morals, i.e., rules, principles, agendas; the latter deals with responsibility for others), which is hedonism. The organism is to enjoy the world as much as it can. This sort of mattering can be achieved by the organism alone, even by the sole bacterium swimming in a sugar gradient. On the other hand, it is only the introduction of the other that shows hedonist morality to be insufficient by introducing a responsibility that exceeds one's self-interest (i.e., by disrupting self-interest). That is to say, though affectivity deals with what matters to the organism that does not mean that affectivity is therefore *ethical* in the sense of one's responsibility to others. Rather, affectivity (or in Levinas, sensibility) can emerge from solipsism and engender hedonism, which is a valuing yet unethical perspective on the world.

The next claim made by Colombetti and Torrance, supposing emotions are ethical since they are value-laden, is that the emotional character of an interaction is in part determined by the interaction itself in a way that is irreducible to the actions or intentions of any particular agent; they embrace De Jaegher and Di Paolo's claim that interaction takes on a life of its own. They then go on to claim that responsibility should not be exclusively conceived of in individual terms but, rather, that interaction itself can bear ethical responsibility, since it produces its own emotional character or tone. But this seems to conflate ethical (or even moral) responsibility with causal responsibility.

Consider the following two situations presented by Colombetti and Torrance (2009): one interaction, between an elderly woman and her caregiver, utterly devoid of affection, genuine engagement, and, for lack of a better term, care; and another interaction in which the caregiver initiates and maintains eye-contact (the elderly woman reciprocates), speaks softly, connects with the elderly woman, and the overall interaction involves a sense of tenderness. First, I want to note that while we might prefer the warmer emotional tone of the latter interaction over the former, we nevertheless cannot reduce ethics to such a preference; sometimes ethics might require coldness. (The preference here is normative and hypothetical, presupposing what we already think such an interaction should look like, but then it is unclear how we can conceive of ethics in opposition to dominant norms, or how to think these norms come about.) While care ethics, for example, teaches that emotions are ethically relevant, it does not erase traditional theories of ethics according to which *other things matter*. Colombetti and Torrance, it should be noted, are *not* attempting to reduce ethics to emotions, but I raise these points simply to illustrate that an emotional account of ethics is incomplete (I do not believe they would object to this). We will see, in part II below, that the Levinasian position complicates enactivist ethics by conceiving of affectivity as more radically passive than enactivism has thus far, but passive in a way that structures interaction itself.

What they do insist on, however, is how the emotional character of and responsibility for interaction depends not only on the participants but also on the unfolding interaction itself. They suggest that their approach requires two shifts, the first of which is

…to see the ethical content or valuation of a given situation as emerging as much from the interaction of the participants as from the autonomous decision making or original authorship of the participants themselves. This shift implies a very different way in which

159

ethical appraisal is to be applied in such situations from the way appraisal is conventionally

applied. It constrains us to defocus (to a greater or lesser extent) from questions of individual

responsibility, exculpation, blame and praise, and encourages us to focus on the ethical

qualities of the interaction itself. (Colombetti & Torrance, 2009, p. 523)

 Colombetti and Torrance leverage interaction-autonomy (and the close association between

autonomy and responsibility) to argue that the coldness of the first interaction is not solely the

caregiver's responsibility. The coldness of the interaction is a relational fact, and the interaction

itself is partly *ethically* responsible. Put another way, responsibility is somewhere *between* the

caregiver and elderly woman. Indeed, to praise the second caregiver and blame the first

represents a "superficial ethical analysis" (522) whereas a sophisticated analysis focuses instead

on interaction.

The shift to interaction-autonomy conceived of as the autonomy of agents in an

interaction seems feasible, but the additional step to redistributing responsibility away from those

agents to the interaction itself is ethically untenable. This seems to be a backwards-facing, guilt

formulation of responsibility, implying a quick transition from causing or bringing about

emotions – that is, being causally responsible for (or even constitutive of) emotions – to *ethical*

responsibility. This may be too quick a transition from causal to ethical responsibility. Crucially,

we do not need Levinas to tell us that ethics is concerned, at least predominantly, with humans

and human responsibility. We may extend responsibility to or for animals, or the environment,

for example, but it seems strange to hold an interaction between individuals to be *ethically*

responsible over and above its participants; how would we hold an interaction accountable,

blame or praise the interaction itself, or demand that an interaction cultivate its own virtues;

would not these responsibilities and moral ascriptions distribute to the participants? (Why would

blaming the first caregiver, for example, only be possible on a superficial approach?) Even if we could properly hold an interaction accountable, to do so would amount to making an excuse, e.g., for a caregiver who should have acted better and whom we actually hold responsible for acting better, and also amounts to putting *some* of the responsibility on the neglected woman herself since the interaction is in part dependent on (though irreducible to) her own action. The ethical consequences of displacing responsibility by appealing to higher-level interaction-autonomy are problematic at best and incoherent at worst.

I would like to add that their second major insight – that social interaction and sense-making are inherently ethical – is also plausible. However, their reason for this is that these are inherently affective, or more specifically value-laden. This helps us to understand that social interaction and sense-making are invested by values, but values are not necessarily ethical in the sense that values can be individualistic. The Levinasian perspective can help us to understand why and how social interaction is inherently ethical, by pointing to the way in which the very encounter with an other calls our own freedom – including our own values – into question. That is, social interaction is inherently ethical because others call us to responsibility.

De Jaegher and Di Paolo do have resources for dealing with the objection that shunting responsibility away from individuals and onto the interaction is ethically incoherent, and considering their possible response can bring us to an up-shot of a Levinasian position. For De Jaegher and Di Paolo, although an interaction has a life of its own, i.e., its own autonomy, it nevertheless also constitutionally depends on the autonomy of the individual participants. They write:

… the autonomy of the individuals as interactors must also not be broken (even though the interaction may enhance or diminish the scope of individual autonomy). If this were not so, if

161

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. (Di Jaegher & Di Paolo, 2007, p. 492)

We could say, then, that in the case of the cold care-giver, since the elderly woman is particularly vulnerable (she is institutionalized and depends on care) her autonomy has been undermined. According to De Jaegher and Di Paolo, maintaining that individual autonomy is an *essential* condition of social interaction, the cold interaction ceases to be a genuine social interaction at all. This *seems* correct, as the caregiver does not even look at or speak with (or even to) the elderly woman, but instead lazily serves her food while gazing out a window, occupied in her world. This interaction, since it is not a genuinely social interaction, then, cannot appeal to a higher-level social structure in order to displace responsibility: the caregiver is squarely responsible for how she treats the elderly woman. In the tender interaction, however, we can understand the caregiver as genuinely engaging the elderly woman, respecting and inviting her agency, and accordingly initiating a social interaction that becomes greater than the sum of its parts. In this case, both parties are responsible since both are genuinely autonomous. It should be noted, though, that this solution somewhat deflates Colombetti and Torrance's conclusion about interaction-responsibility, since it is plausible that it is the former kind of "interaction" (which is not genuine interaction), and not the latter that will be most important when thinking about ethics and, of course, in the former we've now done away with interaction-responsibility.

Yet, it is incorrect to generalize the point and assert that a relation is not a genuinely social interaction simply because one of the relata's autonomy has been destroyed. Many of the great unjust, violent social institutions have done precisely this. On the one hand, we would not say that for this very fact they are not social: a social relation is one in which two or more

persons relate directly to one another. It is also not for this fact simply not an interaction (unless of course we only understand interaction as a kind of symmetric transaction, which I have suggested above is precisely the wrong move): this view overlooks the way in which one person's passivity itself can affect the interactional dynamics of a social relationship. That is, even if one's autonomy has been destroyed, this does not therefore make their presence, precisely as non-autonomous, non-active, that is, passive, irrelevant to the encounter.[11] To interpret the Hegelian master/slave dialectic loosely, though the master might dominate the slave this does not mean that the slave does not structure the relationship; it is by the slave's very passivity – by giving in to the master's will rather than his own – that the relation takes shape.[12] The point is that it is not only in acting that one interacts, but also in calling action into question; the more impotent the call to responsibility, that is, the less effective one is in actually holding another accountable, the more desperately forceful that call becomes. The very demand made by the presence of an other, even and perhaps especially a slave, the greater the effect on the overall ethical character of the interaction, regardless of any actual response to that call. Reconsidering the cold caregiver, can we really claim that ignoring a vulnerable someone's presence is in no way a form of social interaction? By her very presence, the elderly woman makes a claim on the caregiver, and this affects the overall character of the interaction whether or not the caregiver responds morally. Accordingly, two-way action is not strictly necessary for defining the social relation or even interaction (in the sense of affecting and being effected by another), as the other

---

[11] There may be some interesting connections between this claim and what George Kunz (1998) calls the "paradox of power and weakness," more specifically the paradox of weakness.

[12] One could also, rightly, say that the master is truly the passive member here, with the slave being active. But this doesn't change the point: one person's passivity still structures the overall interaction, and even if someone is dominated completely it cannot change the fact that their very presence as an other calls the activity of their interlocutor into question. This is one reason why so much ideology is necessary when attempting to justify oppressive dynamics.

affects us in our own passivity even if we try to reduce them to an object. The very vulnerability of the other thus reduced opens a dimension of sociality precisely as responsibility for her.

The difficulty with the position is that it privileges individual autonomy over the social relation. Though De Jaegher and Di Paolo want to move beyond methodological individualism, they nevertheless take as their starting point human agents as defined by autopoiesis (they endorse the idea, defended by Jonas, that "metabolism creates a perspective of *value* on the world" (2007, p. 488) (remember: this *value* might be hedonistic-solipsistic, i.e., not ethical) and extend the terminology of autopoiesis to describe genuinely social interactions as operationally closed and oriented towards sustaining themselves.

Primary Intersubjectivity as Address (Corrective)

Levinasian insights can help approaches that hope to move beyond individualism. Though Levinas affirms the individual at the level of sensibility (he does not deny the existence of something like an embodied self), he nevertheless makes even autonomy subservient to ethics. He could challenge the idea of autopoiesis thusly: we are *not* creatures that strive only to sustain ourselves by whatever means the environment affords, but instead we *enjoy* the world and enjoyment cannot be understood as a means-end relationship, not even of survival; ethics disrupts our enjoyment and in doing so invests our freedom. That is, our individual autonomy is the autonomy of a sensible being that enjoys the world, but as soon as a social relationship appears, autonomy is invested by the relationship to the other: relational-autonomy (the autonomy of the relationship itself) appears as a being called out of individual-autonomy, specifically as the questioning of one's naïve freedom to enjoy the world according to self-interest. The elderly woman, by her mere presence, autonomy aside, calls us to ethical

responsibility, and it is this call – below and beyond any normative context – that motivates our holding the care-giver accountable, motivates even outrage against norms themselves.

Crucially, from this perspective, individual-autonomy is a kind of myth. Without being called out of oneself, one's actions are *determined* by self-interest. Being called out of oneself in responsibility to others makes one's actions an open question – to enjoy or not; how and why; what to do? One can choose how one responds, and even ignoring someone is a response, but one cannot choose whether to respond at all. The neglectful caregiver responds poorly. Put another way, sociality constitutes us as individuals, in the sense that sociality calls us to account for ourselves, that is, to offer ourselves to the other. Relational-autonomy is not simply grafted onto individual autonomy but is, rather, a being-called-out of naïve individual autonomy that more closely resembles enjoyment-seeking automaticity than genuine freedom. The ethical relation is a primary relationship with another person beyond myself. My genuine freedom depends on being called into question by the other, as a way out of deterministic self-interest. This works itself out in the first instant of contact with an other, prior to any activity on my own part but instead simply in the presence of an other who expresses herself. The important point that has been missed is the asymmetric direction of this ethical dimension: the other calls me out of myself, whether I make any claims on the other or not, and before I respond in any way. Specifically, the other calls me to responsibility for the other, and whether the other is likewise called to responsibility by me, as Levinas says in his interviews with Philippe Nemo, "is *his* affair" (i.e., not my business, not my concern) (Levinas, 1985, p. 98). He elaborates: "the intersubjective relation is a non-symmetrical relation. In this sense, I am responsible for the Other without waiting for reciprocity, were I to die for it. Reciprocity is *his* affair" (Levinas, 1985, p. 98). Interaction *in the sense of reciprocal action* is not necessary for a concept of

165

relational autonomy, as my autonomy is already invested by the other (1) in passive terms, i.e., in being affected by the other independent of any *act* I go on to perform and (2) without waiting for any response on the part of the other. It is in an encounter with the other that my individual autonomy gets set-up, but an encounter that may not yet be interaction in the reciprocal sense.

Two important points. We should not start with autopoiesis and individual-autonomy and then try to sort out sociality and ethics at a higher level, but rather should begin with an ethical relationship and understand autonomy on the basis of sociality (in particular, in terms of disruption, investment, and responsibility), and we also must recognize that at times ethics might require that autonomy – even relational-autonomy – be compromised. Indeed, Levinas' ethics is one that makes possible self-sacrifice unto death, which is untenable from the perspective of autopoiesis or even an interaction that privileges my autonomy. Autopoiesis is a self-oriented (in the sense of referring to my own, organismic processes and interactions with the world) and exploitative (ethics as value, value as value-for-me) characterization of human experience. Calling autonomy into question is not the end of genuine social interaction but, instead, its beginning.

My proposal, which should already be anticipated, is this: construe primary intersubjectivity, our primary relationship with another person, as involving an asymmetric component of address. Building on the previous chapter as well as the concerns raised above by way of Levinas' criticism of Buber, we can better understand our experience of another person as distinct from our experience of objects, and can simultaneously better make sense of the positivity of both, if we understand the other as calling oneself out of oneself. In other words, our primary relationship to the other as an other is not simply a relationship to something similar to ourselves that enhances our own powers by referring them to a new kind of object or by teaching

us about ourselves. Instead, primary intersubjectivity involves the way in which the other

disrupts my powers, introducing a foreign element into consciousness (what Levinas calls

subjectivity as "the-other-in-me" or "inspiration") and thereby investing my freedom. The result

is a social relationship to an other not merely as an analogue to myself but rather as an

interlocutor to whom my actions – whatever they may be – constitute a response. As invested by

the other, which we can understand in terms of the way that a disruption of myself is

simultaneously a re-direction of attention and, thereby, of motivation, we can also understand the

social encounter as an overcoming of individual autonomy; it does not destroy my autonomy, but

calls me to accountability for it.

Importantly, the kind of asymmetry I am proposing here is an *ethical* asymmetry,

according to which by the very fact of expressing herself to me, I am responsible to and for the

other – my self-interest is disrupted and response is unavoidable. In his 2014 article, "In your

face: transcendence in embodied interaction," Gallagher is right in arguing that transcendence is

not *in* the other; rather, it is a relation to the other. But he too hastily dismisses asymmetry in

favor of reciprocal recognition: that the other and I meet each other's gaze (Gallagher, 2014;

2020, p. 206). On Gallagher's view, asymmetry is introduced to the intersubjective relationship

externally, by institutional arrangements for example. But on my own position, asymmetry is a

basic characterization of intersubjectivity as such: the relation to another subject just is a relation

to she who interrupts my content sensibility. He is perfectly correct that reciprocity plays an

important role in the establishment of mutual sense and, therefore, ethical *norms*, but I am

suggesting that we understand asymmetry not only as disrupting reciprocity as a higher-level,

institutional modification, but as motivating that very reciprocity from the ground up: the other

calls me away from myself toward the other, and in every moment of interaction, even when

167

reciprocal, we are oriented to an other who is reducible neither to myself nor that interaction. Crucially, it is the first moment of encounter with another – prior to any response on my own part – that we are already called to responsibility, regardless of reciprocity.

By understanding the asymmetry of primary intersubjectivity, by which I simply mean that the way that the other demands a response from me is the very basis of our relationship, we can better understand responsibility, ethical experience, altruism and self-sacrifice. In the social relationship, I am already directed to and by the other, and as such am already *for* the other; my responsibility to the other is not reducible to my own interests (called into question by the other) or to the interaction (i.e., transaction) itself constituted by this responsibility. There is a positive presence of the other, though it is not an ontological-perceptual presence (nor the in-principle impossibility thereof, which would be only a negative phenomenology), which is precisely the disruption-redirection-responsibility complex instituted in me by the other. Even an infant, who perhaps knows nothing of responsibility, is disrupted and redirected by the other, and as such responds. We can understand altruism and self-sacrifice, since my own freedom is already invested by the other from the start, and in any case its exercise is called into question and in need of justification: this account allows for the charge that the use of my freedom is *unjustified*, and the standard of justification lies not in my own self-organization but in the demands made on me by the other.

My proposal, then, allows for the enactivist to begin with autopoiesis for understanding an isolated organism, but provides the means by which self-interest can be overcome by the very start of sociality. The point remains that primary intersubjectivity is not a developmental stage to be overcome, and this implies that our higher-level systems through which we make-sense together always involve an asymmetric relationship where one is held accountable to the other as

the very ground on which such an interaction can be social in the first place. We can say that the interaction itself shapes the character of the social encounter, but we also can coherently hold individual participants responsible for that character even though it exceeds them; indeed, individual responsibility is the very motivation for understanding the ethical character of the interaction in the first place.

Enactivism, for its part, still plays a valuable role in understanding how we transition from ethics to understanding.

From Ethical Asymmetry to Doxic Symmetry

Totality and Infinity is shaped by a certain insecurity: Levinas wants to describe an alternative to ontology, yet is seemingly forced to do so in ontological terms. In *Otherwise than Being*, however, this is no longer seen as a bug but a feature; the anxiety of falling back into ontological terms is instead embraced as a kind of necessity. Though the other disrupts consciousness and thus challenges our familiar categories by affirming the priority of ethics, consciousness nevertheless consists in bringing-home even this disruption, casting it in terms of being, i.e., ontology. Though ethics is primary for Levinas, and is the very foundation of the social as such, it is nevertheless the case that we do (and must) make-sense of the other's expression in ontological terms. What I propose now is that, bearing in mind the new interpretation of primary intersubjectivity as involving a challenge to consciousness – a disruption of thought and perception – enactivism nevertheless shows how a transition is affected from ethical asymmetry (that the other calls me to responsibility prior to my response) to doxic symmetry (mutuality of sense, meaning, belief, etc., i.e., meaningful agreement), precisely through reciprocity.

Another way to pose the issue is in terms of particularism. Levinas appeals to the face to face encounter as a source of ethical significance, and the appropriateness of any response seems to be dependent on the particular demands being made in a given encounter. Yet, there is also in Levinas' work a concern for justice as universally applying to all. How could such a justice be based on a particularism that rejects universalizing principles? More broadly, it is sometimes argued (e.g., by Critchley & Bernasconi, 2002) that Levinas is not doing normative philosophy *at all* and has nothing to tell us about what we should actually do. Diane Perpich (2008) defends Levinas as articulating a "normativity without norms," yet we still need ask: whence norms, and more specifically, whence ethical norms that respect the face-to-face encounter? The claim here, then, is that ethical asymmetry explains an initial motivation for ethical sense-making, and that enactivism can then explain – through reciprocity – how norms are established in (and feed back on) interaction.[13]

Interaction theory and participatory sense-making prioritize reciprocity: making-sense is a give and take activity. And this is fair enough. Levinas, however, can help us to appreciate an underlying ethical asymmetry always operating beneath reciprocity. Interaction takes place in the face of an other who makes a demand on me in virtue of the very fact of interaction. That is, the other who expresses herself demands that I make-sense while simultaneously making sense-making possible, opening herself to sense-making as reciprocal. It is out of responsibility to an other with whom we make contact through vulnerability to expression that we take up projects of object-constitution in the first place. This underlying ethical asymmetry is broadly compatible

---

[13] The position developed here is similar to that developed by Dierckxsens, but with a few differences: (1) I make no claim about multiple senses of justice; (2) I emphasize the establishment of mutual sense and norms; (3) we express asymmetry differently, and I leave open whether our expressions are compatible; (4) I do not endorse "pre-ethical social interactions," which Dierckxsens claims enactivism reveals, as coherent; instead, I insist that enactivism must recognize the ethical dimension of social interaction generally. For all that, I highly regard Dierckxsens work on this topic.

with interaction theory and illustrates how an ethical dimension conditions further cognitive/affective dimensions and analyses. If primary intersubjectivity is interpreted as involving asymmetric address, then we can see multiple ways that common belief (and knowledge) depend on this prior ethical understanding. In the last chapter, we explored how it is precisely the ethical significance of interaction that directs our attention and underlies the learning of general categories. But we can also see how reciprocal sense-making, a give-and-take of propositions, things said, and expressions conceived as propositions or acts, rest on a responsibility one has to respond to the other. Or, on one's ability to respond, which already implies an asymmetric relation.

Rakoczy, Warneken & Tomasello (2008, 2009) and Schmidt & Rakoczy (2017) explore the ways that certain give-and-take interactions between children and adults, which can be seen as a development of natural pedagogy, result in the recognition of norms and normativity in terms of what one should do. The 2008 article shows that children can apply simple rules across contexts and will respond normatively to violations of those rules (they protest, critique, or teach). In 2009, the authors go on to show that when learning various acts and rules, young children engage in selective normative learning, that is, they prefer reliable teachers not only as providing better information but also normatively appropriate information; they will adhere to norms learned from a reliable agent and extend those norms to new contexts. In 2017, Schmidt and Rakoczy provided evidence that "very young children engage in rational and selective third-party norm enforcement, which suggests they understand some important features of normativity". This shows that "human social cognition is not only concerned with the prediction and explanation of others' behavior, but also with the prescription and evaluation of others' actions." This just demonstrates that normativity is bound up with cognition from a very young

age. But the relationships we are interested in here are those between normativity and ethics, and between teacher and student (in this case, children) such that children learn these norms in the first place.

In 2009, Rakoczy, Warneken & Tomasello raise the question: "*When, under what circumstances* and *how* do young children engage in selective normative learning? Which cues, for example, do they make use of (such as verbal, ostensive, or other potentially pedagogical ones)?*" (2009, p. 68). In the previous chapter, a connection was drawn between natural pedagogy and expression construed as asymmetric address. Since normative learning is a kind of learning, we can suggest here that the same conclusions apply: it is by being called to respond by the other that our attention is guided in such a way as to pick out general features of the situation beyond particular here-and-now qualities and thus institute normative learning. It is also intriguing that each of these studies employed a human teacher of norms, on the one hand, and either a puppet teacher or puppet transgressor, on the other. The findings seem to ignore the differences this might make. Overall, children preferred norms taught by human others, were able to apply those norms across different contexts more readily than with those norms learned by puppets, and children prescribed behavior to (as well as criticized and reprimanded) puppets much more so than humans. There is something in the response to a human that is relevant for the questions raised just above: it is in response to a human interlocutor that children learn these norms, preferring human-taught norms and discounting puppet ones. A consideration beyond our scope here is the question of power and authority, as it was shown that children would not typically reprimand adults in these studies, and would also trust adults as reliable sources of information over other children, choosing to enforce adult-learned norms over children-learned/co-constructed ones.

Our answer, then, to the kind of cues that actually motivate norm learning refers in large part to asymmetry: it is human expression that calls the child into question and directs their attention to the kinds of generalities necessary for entering into reciprocal interactions with other children, puppets, adults, and so on, out of which normative awareness – or better yet, normative activity – arises.

Enactivist accounts of participatory sense-making and interaction provide insights on the transition from a reciprocal interaction to *doxic symmetry*, i.e., to agreement of sense: through interaction, we come to understand *each other*, or come to understand the world in *largely the same way*. This is a kind of transaction in which we trade, adopt, modify, and challenge senses and understandings until they resemble one another closely enough to get on with business. The entire business of mind-reading and even direct perception is a part of this economy, whereby we understand the other by bringing something of our own to the table, and insofar as this process is externalized, that is, expressed, the other comes to better understand us as well. But the other is never completely known to us: they are always-more-than any category that we bring to them (Kunz 1998, p. 36) and this accounts for 'infinition' in Levinas, i.e., the production of the idea of infinity through contact with the other.

It is this excess, the escaping by the other from any conception such that a demand is always made by expression, that is at the heart of ethical responsibility (responsibility for someone other than myself), ethical experience (of others as irreducible to oneself and making necessary response, even if inadequate), and altruism and self-sacrifice (as action *despite* myself). This excess requires that our reciprocal interactions be sensitive to ethical asymmetry. That is, the other is a source of signification and can always disrupt our participation or interaction, can always disrupt the agreement of concepts and our give-and-take exchange in

173

order to call what we think we know and our activity of sense-making into question. The other, then, has a kind of authority even in a reciprocal interaction if the question is precisely the understanding of the other or her world. This applies to the discovery of ethical principles understood as rationalistic propositions – we might apply utilitarianism to determine how we treat the other, but the other can always protest our determination and demand otherwise – as well as to determining states of affairs.

Conclusion

Experience of others is distinct from experience of objects precisely insofar as the former and not the latter is constituted by an asymmetric relationship characterized by expression, address, or demand; it is our responsibility to others that distinguishes them from objects. Ethical responsibility is thus a fundamental dimension of intersubjectivity that cannot be derived from purportedly prior states, such as already established norms or particular orientations towards the other such as compassion or care; it is as already responsible that emotions and norms gain their significance, an ambiguous term signalling both sense and importance, and that we come not only to understand norms but also to create them. Insofar as enactivism attempts to make sense of ethics in terms of a reciprocal relationship that privileges autonomy it misses this ethical dimension and as such cannot satisfactorily account for responsibility, ethical experience, or the possibilities of altruism and self-sacrifice.

Ethical asymmetry founds learning, as argued in the previous chapter, but also satisfactorily accounts for the shortcomings of current enactivist thought mentioned above. Importantly, such a dimension can (and should) be embraced by enactivism. Most readily, ethical asymmetry can be incorporated into an enactive account of motivation as well as learning (including of norms and normativity); can serve as a standard for or check against sense-making,

that is, can make sense-making accountable to someone outside the ego (thus making sense of non-normative/counter-culture experience, including pathology phenomena for which a familiar analogue in one's own experience is lacking); and can help us better understand the establishment of mutual *doxic* sense, i.e., how rationality is based on a different kind of non-rational experience (recall *Chapter 2*, where it is shown that expression is largely a matter of affectivity, though an affectivity conceived as a radical passivity to the other), which is something enactivist approaches to sense-making and ethics alike are interested in pursuing.

In the final chapter, I go on to specifically explore the implications of these Levinasian insights for psychopathology.

Chapter 5

Levinasian Psychopathology:
Embodiment, Substitution, & Self

Introduction

In this chapter, I juxtapose phenomenological and 4e concepts of embodiment, self and intersubjectivity with Levinasian concepts of embodiment as sensitivity, substitution as a distinct form of perspective-taking, and self as produced in response to the other. I then propose three types of disturbances corresponding to what have been traditionally known as "self" disturbances or disorders, and apply these to depression, grandiose narcissism, and certain symptoms associated with schizophrenia. This typology has therapeutic implications, which I touch on only briefly. Finally, I consider a related, Levinas-inspired position: the Seattle School approach to psychotherapy. I argue against the approach's characterization of mental illness as ethical deficiency and emphasize that Levinas' challenge is specifically to the responsibility of his reader: it is the therapist's responsibility that is in question, not his patient's.

Embodiment, Self, and Intersubjectivity

Phenomenological and 4e approaches to psychopathology have recently begun to stress the importance of the body for understanding not just cognitive phenomena generally but also pathology phenomena. Many of the latter are accounted for in terms of disruptions to some form of embodiment – either one's tacit, embodied being in the world, i.e., living one's body as a set of motor dispositions and coordinations that allow one to competently navigate the environment without themselves becoming thematic, or in terms of one's body image or the body as an object, that is, one's perceptual, emotional, epistemic representations of one's own body.

The body is also closely connected to the idea of the self, and there are numerous debates surrounding this connection. First, it is contentious whether embodiment constitutes a form of

selfhood or, on the other hand, is pre-personal. Second, it is maintained that the body and the self alike – whatever their relation – are implicated in the question of intersubjectivity: we seem to encounter others *as* bodies with our own bodies. Miriam Kyselo (2014) raises what she calls the body-social problem, which is a question of resolving the tension between bodily and social aspects of individuating the self: does the self exist in some sort of individual form prior to social encounters with others – perhaps in virtue of mere embodiment – or does the coordination of bodies within social encounters constitute the identity of the self?

Aligning with these overlapping concerns, pathological disruptions of embodiment are typically taken to either be or to imply disruptions of self and/or intersubjectivity. Levinas' conceptions of embodiment and selfhood are helpful here. Recalling *Chapter 2*, embodiment constitutes a kind of ego, but there is also a simultaneously social and personal self achieved not by the body itself but by a relationship with the other.

Embodiment-Substitution-Self

Levinas' conception of embodiment, as explicated in *Chapter 2*, refers to one's vulnerable sensibility: the body is exposure to the elements, and therefore to force, hunger, ageing and death. Sensibility is not a strictly causal notion, since it is also characterized by an affective, phenomenal dimension: what is encountered by a sensible body matters to that body not only in terms of its own survival (and this is where the autopoietic notion of affective sense-making hits its limit) but also in terms of its enjoyment; the body's needy exposure to the world becomes a site for enjoying those very things it needs. It is even enjoyable to the body – and this is perhaps the fact that leads Plato and John Stuart Mill, for example, to distinguish higher from lower, i.e., bodily pleasures – to consume that which is antithetical to survival: we sometimes enjoy over-eating, unsafe sexual practices, and substance abuse. Importantly, there is a kind of

"for-me-ness" present here, insofar as what is given to the body in sensibility is given for the body to enjoy. The body constitutes a certain kind of ego or self, but this is not the same as the "minimal self," conceived of as a perspectival owner of experience. The sensitive body does not yet sharply distinguish that which is given to it from itself, since what is given to it is given precisely in terms of its own self-interest. Embodiment therefore does not yet get us out of subjectivism or apparition (again, see *Chapter 2* above), and bodily interest is not best characterized by a re-production or survival but, rather, enjoyment.

A full-fledged "I" or self, capable of disentangling itself from the world it inhabits, comes on the scene, for Levinas, "first in the accusative, then in the substantive." This means that the self is not accomplished by the body as a kind of subject of appearances, but rather exists first as the body called into question by the other: the body as disrupted. In our dealings with objects, the body is disrupted by *objective force*, meaning that we cannot readily appropriate worldly things and must sometimes adapt our methods or move on. The other, however, presents the body with *ethical resistance*: what is called into question is not my ability to overpower something in the world through my own action but, rather, to justify that action, to answer for myself to the other. Thus, called into question by an other (in the accusative rather than substantive), the self exists as a kind of accounting for the body's existence as a naïve enjoyment of its world. This "accounting for" somewhat resembles Varela's conception of the "I" as an "ongoing interpretative narrative" (Varela, 1999, p. 61), the crucial difference being that it is not I who call myself into question, but the other who establishes my identity by demanding an account of me.

Henceforth, I refer to the kind of self constituted as bodily sensibility as "ego" or "me," because it refers to a kind of egoist enjoyment and is the thing called into question by the other: an account is demanded of *me*. I refer to the self constituted by the other's demand, which

disrupts *me*, as the "I" or, since it is that to which our normal ascriptions to the self most readily

apply, the personal self or, simply, the self: an account is demanded of *me*; *I* respond.[1]

The moment at which the other who expresses herself calls me into question is the

moment in which the world, rather than simply appearing for the body, begins to appear to the

body as having appeared *for* the body and, simultaneously, as for-the-other that calls this body

into question. The world is no longer simply *for-me*; the for-me-ness of the world can no longer

be taken for granted. It is by occupying one's body as oriented by the expressions of the other,

that is, occupying one's body through the other's perspective despite one's own, that I become

constituted precisely as he who must respond to her expression. This displacing of my own

perspective in response to the other, according to which I take on their needs and desires as my

own, is called substitution.

Levinas formulates substitution as substitution "of one for the other", in association with

language, responsibility, "and the condition (or the uncondition) of being hostage" (Levinas,

1981, p. 6). He also writes that "substitution is signification" (ibid.), that is, putting oneself in the

place of the other is the first instance of significance in the sense of the intentional structure, X as

Y. As outlined in *Chapter 3*, it is in an encounter with the other that we come to account for

ourselves, the other, and the world by thematizing that which is given to sensibility as signifying

something beyond myself. In addition to the significatory sense of substitution, substitution also

has a material-responsive sense: to substitute oneself for the other is to respond according to their

needs, desires, expressions. If the other calls into question my right to food, substitution is not

only an understanding of what I have as food, but also an understanding of that food as *for-the-*

---

[1] These terms do not necessarily line up with Levinas' own. His uses of "ego," "me", "self," and "I" are somewhat slippery throughout his works. The terms I apply here are simply those that make sense in context and will allow the discussion to move forward.

other rather than *for-me*. Calling my right to food into question, then, involves not just thought but also the possibility that I actually give my food to the other.

When I refer to substitution here, I am referring to a kind of perspective-taking. I do not mean the acknowledgment that the other is a perspective on the world like me to be weighed into my epistemic judgments, and I do not mean to take on the other's perspective primarily in terms of understanding what they understand, or taking on their point of view. What I mean by perspective-taking, when I give it the name "substitution," is the orientation of my body towards the world not according to my own self-interest (which can still dominate any practice of seeing from the other's point of view), but instead according to the needs, desires, and – more generally – expressions of the other. Substitution is a taking up of the world as for-the-other, rather than as for-me. This does not primarily refer to understanding the world as the other does, but to putting oneself in a position to respond to or for the other's needs, desires, and so on. However, in order to materially respond to the other, the significatory role of substitution is indispensable.

Expression, as calling me to account for myself and, thus, to substitute myself for the other, is a principle of signification, the possibility of intentional experience characterized as being directed towards X *as* Y, that is, *as* something beyond my own sensible grasp on it. No amount of sensation may give to us a perfectly adequate understanding of the other's expression; it is by communication and co-operation, allowing any conception we form to be tested by the other and possibly disrupted – overturned, rejected, reinterpreted – that we come to understand.

In summary: Levinas' conception of embodiment refers to a pre-personal but egoic sensibility exposed to a world (and therefore to the possibility of an encounter with others). The self or I comes into being in the encounter with another who calls the ego's naïve existence into question and reorients one's attention by demanding the ego account for (or justify) itself and the

world. The self is embodiment thrown back on itself by responsibility. The Levinasian account therefore agrees that the body is pre-personal and perspectival but goes further by disambiguating for-me-ness and selfhood: the bodily ego experiences the world as for-me, but selfhood emerges from the disruption of that for-me-ness. Crucially, the self can and often does collapse back into egoism and an orientation towards the world as existing for-me, but this occurs precisely where one puts their own enjoyment above responsibility. That is to say that the ego and the self, mediated by substitution, exist in a dynamic relationship; one can overcome egoism by orienting themselves to the world according to the other's needs, but one can also fall back out of such an orientation into a hedonistic search for sensible pleasure.

This relationship between embodied ego, substitution, and personal selfhood also carves out an answer to the body-social problem by proposing the social mechanism that accomplishes the transition from individual embodied existence to a simultaneously bodily and social self. Kyselo's main concern in raising this question is to avoid choosing between embodiment conceived as a solipsistic ego separated from sociality, on the one hand, or as entirely lost in social immersion, on the other. This account navigates the dilemma by proposing a specific role for embodiment while requiring an ambiguous social relation for a full realization of the self. Substitution is the transition from a bodily ego to a social self, and it involves a social ambiguity or immersion in the sense that it involves disrupting (though never completely overcoming) my experience of the world as for myself by taking up the world as for-the-other.

Moving forward, except where noted, I use the following terms interchangeably: embodiment, ego, me; substitution and intersubjectivity; self, personal self, I.

Some Results: Disruption, Pathology, and Threat

Disturbances of self and intersubjectivity in pathology phenomena can be classified in two ways, according to Thomas Fuchs and Jann Schlimme. The first kind of disruption refers to the lived body or body schema – which for them includes an embodied sense of self and a tacit understanding of the world and inclusion in "common sense"; this is the kind of disruption posited in relation to schizophrenia and depression. The second kind of disruption refers to the object body or body image, one's explicit awareness of oneself; this kind is posited in relation to "body dysmorphic disorder, hypochondriasis, somatoform disorders or eating disorders such as anorexia nervosa" (Fuchs & Schlimme, 2009, p. 571).

The first important result of following a Levinasian approach is that disruption of embodiment is not necessarily pathological. Indeed, the constitution of the self presupposes a disruption of naïve embodiment by a social encounter with the other. According to Fuchs & Schlimme (2009, p. 571), the lived body becomes experienced as the object body – the schema becomes available as a set of images – precisely where the implicit functioning of the lived body breaks down. I agree, with two caveats. First, if this is so, we must ask what kinds of breakdowns distinguish the ordinary transition from lived body to object body from the kinds of breakdowns that characterize pathological phenomena. Second, we must include in our considerations not just interruptions of autonomous motor functionality, involving body schema, even those instituted by an encounter with the other, but also ethical interruptions, i.e., interruptions of the body's egoistic, hedonic naïve existence defined by content enjoyment of the elements.. It is not just when I trip over a rock or when my body becomes unhealthy that I conceptualize my own body in such a way as to differentiate myself from the world, but rather when I am called to respond to an other and thus thematize the world, myself, and the other as significant – epistemologically

and ethically – beyond myself. The point here is that disruptions to embodiment are not automatically disruptions to the self, and that such self-disruptions already include a non-motoric and non-perceptual salience: the I exists insofar as I am responsible to and for others.

The second major result is a new way of thinking about pathology phenomena. In particular, the distinction between ego and self, and the idea of substitution that accomplishes the transition from one to another, might help us think about a nexus of seemingly contrasting symptoms all associated with schizophrenia. Schizophrenia seemingly involves: (1) disruptions of self, such as alienation from one's own body or mind, separation between one's own body and mind, and the dissolution of one's personal identity associated with experiencing oneself as empty or void; (2) disruptions of intersubjectivity such as anomalous or inappropriate responses, disturbances in understanding the other's expressions, a sense that others are oppressive or otherwise threats to one's own identity, and divergence from commonly accepted meanings; (3) self/other ambiguity, fusion, or con-fusion, that is, ambiguity as to whether a given experience or action is one's own or, instead, attributable to an other. (2) and (3) represent certain forms of separation from others and a lack of separation respectively. Froese and Krueger (2021) focus on the latter, and suggest that a self/other ambiguity is actually inherent to intersubjectivity. They maintain that the self arises as a differentiation from this prior ambiguity. This jives with enactivism and extended cognition in general: if my own processes involve or depend on interactions with others, then some of my processes are genuinely ambiguous. To lose this ambiguity altogether is tantamount to being separated from intersubjectivity, and they call this a loss *of* the socially extended mind. On the other hand, to fail to differentiate oneself from this ambiguity is to be lost *in* the socially extended mind: to fail to constitute an identity for oneself out from the ambiguous processes that constitute my own processes. This distinction echoes

concerns raised by Kyselo (2014): solipsism (or "quasi-solipsism", in Parnas and Sass' (2001) terminology), on the one hand, and dissolution of the self on the other.

On the Levinasian account, we can understand this self/other ambiguity as the transitional stage between embodiment and the emergence of the self. Ambiguity, then, can be understood in terms of substitution, or taking the other's perspective. This already has an ethical significance: it is already a response to the presence of the other. *To lose the socially extended mind*, or self/other ambiguity, is not necessarily to suffer a disruption of embodiment, but could alternatively be to remain thoroughly within it, i.e., to maintain the ego against the others. This is one reason why a perceptual account of intersubjectivity cannot always make sense of intersubjective disruptions: one might perceive the other in all the ordinary, sensory ways, but without being disrupted by their address there may be no real contact at all. In this way, Levinas' conception of intersubjectivity diverges from Husserl's (see *Chapter 2*): for both it is an encounter with the other that motivates a sense of the world as objective, that is, as shared, but for Husserl this is a perceptual relation with an other taken as analogous to myself, whereas for Levinas it is an ethical relation through which the other calls me to take on their own perspective, and that can only be accomplished by a communication that disrupts my perception.

To put the point another way, the shared world comes about not only by a sensorimotor coordination between bodies; that very coordination must be taken as something produced between bodies in communication, where one is motivated to respond to the other. Just as a Levinasian position requires us to redefine primary intersubjectivity (see *Chapter 4*), it also requires that we rework the idea of *genuine* intersubjectivity. For Froese and Krueger (2021), genuine intersubjectivity consists in co-regulated experience. But there is already intersubjectivity prior to co-regulation, in the asymmetric encounter out of which co-regulation

arises. First there is disruption, and then only in response to disruption of our own embodied processes does co-regulation emerge. I thus agree with the enactivist claim that "reality is not something predetermined and external" but we must complicate the idea that it is instead "brought forth by a living being's *sensorimotor interaction* with its environment" (Fuchs, 2020, p. 63; italics his). The sensorimotor alone is inadequate for the constitution not only of a shared reality but of a significant reality at all, as well as the formation of a personal self; disruption in the social sense is necessary for the constitution of a non-body-centered "reality," even if not yet a reality of "common sense."

On the other hand, *to lose oneself in the socially extended mind* is already to be disrupted by others without that disruption resulting in the formation of a personal identity. This kind of disorder of self is not first and foremost disruption of embodiment nor even intersubjectivity (though self-disturbances could feedback on embodiment and intersubjectivity): a certain disruption to embodied existence is a necessary condition of the self on this account, and there is also already an intersubjective relationship entailed by the possibility of being lost in the socially extended mind. Instead, the matter is the transition *from* the embodied interaction *through* the intersubjective relationship *to* the constitution of the self. This will become especially important in the following discussion of the Seattle School of Psychology. For now, I note only that this account could help to explain reports by persons diagnosed with schizophrenia that the other manifests as a threatening presence, particularly that others threaten their identity (Lysaker, Johannesen, & Lysaker, 2005). The interaction with the other, beyond the ambiguity captured by the extended mind hypothesis in terms of co-regulated processes, is indeed a threat to one's pre-reflective mode of being, and even in ordinary cases is a threat to the self. What I am describing as disruption is not ended with the formation of a personal self, or with substitution being carried

out once and for all. When I can think of myself as a self and account for myself and the world in responsibility to others, that identity is nevertheless always called into question by the expressions of others: the call to responsibility can always disrupt my own existence. That is, the self is always a fluid navigation between my embodied existence and my responsibility for others. I propose, then, that not only is a self/other ambiguity central to intersubjectivity, as argued by Froese and Krueger, but that this ambiguity also involves a (sometimes felt but always operative) threat to one's identity: one's identity is established and re-established only by a disruption of ordinary being that institutes such an ambiguity. Experiences of the other as threatening can therefore be thought of in the same way as Froese and Krueger think of self-other fluidity, as "not indicative of hallucinations without any basis in reality, but of a heightened sensitivity and vulnerability to processes of interpersonal alignment and mutual incorporation that form the normal basis of life" (2021, p. 318). That is, the experience of the interpersonal relationship as threatening is just another sensitivity to a fact of normal interpersonal relations: the other does indeed threaten my identity. Some pathological cases consist of a pernicious thematization of this fact.

Self-Disturbances

To recap, the self is a product of an encounter between one body, enjoying a seemingly individual existence in the world, and another body that disrupts it. In the face of disruption, my body becomes substituted for the other, entering into an ambiguous relationship that blurs the distinction between us. Substitution is a simultaneous distinction and blurring of distinction: not only is my sensorimotor activity caught up with the sensorimotor activity of the other, but my attentional processes and the ethical significance of the world are all redirected by the other. It is in response to this redirection that I am constituted as a personal self, accounting for my role

186

played in the encounter and for my right to the things now apparently existing not only for-me but for-the-other. That is, it is in response to the other who disrupts my bodily existence that I first differentiate myself from the ambiguity that disruption institutes. This is not a developmental story that is completed once and for all: in any encounter with the other, responsibility to and for the other can disrupt my own existence (thus there is salience beyond my own survival, which accounts for the possibility of altruism; see *Chapter 4*). Specific disturbances to embodiment will translate to disturbances in intersubjectivity and selfhood; likewise, a disturbance to one of these latter dimensions will feedback on the others. Embodiment as sensitivity, substitution, and the constitution of the self stand in dynamic relations such that anything that affects one will affect the others.

Based on the preceding discussion, what are commonly known as disruptions of "self" may pertain to:

(1)   embodiment as vulnerable sensibility; (bearing in mind that what is often called the embodied self may simply refer to pre-personal embodiment)

(2)   the transition from embodiment to substitution, or self/other ambiguity

(3)   the transition from substitution to self, differentiated from ambiguity as response.

The question, then, is trying to understand for any given self-disturbance, what kind it is. I propose that certain phenomena, like those associated with major depression or Cotard delusion, involve something like type-1, while others, such as narcissistic personality disorder or, perhaps, psychopathy, are better understood in terms of type-2, and still others, such as certain symptoms of schizophrenia, as type-3. (Admittedly, schizophrenia is an especially difficult case.) Each type of disturbance will involve a nexus of characteristic phenomena relating to embodiment, intersubjectivity, and self. For example, a type-1 disturbance – "insensitivity", as I call it – will

have implications for intersubjective relations and the constitution of the self; type-2 disturbances, likewise, will have implications for both sensitivity and the self; likewise for type-3. The following is a cursory attempt to put this typology to use in relation to various pathology phenomena. I will accordingly use the terms "primary" and "secondary" disturbance, depending on which type seems to be at the core of each set of phenomena.

Type-1 Disturbances: Embodiment (Depression)

It is of course the case that all human selves (as far as we know) are indeed embodied.[2] But, this is not necessarily reflected in experience. It is possible for one's body to be given as insensitive, inert, or passive. This is possible in one's image of the body (body image) or their style of embodiment (body schema) but also in terms of sensitivity. This is sometimes reported by persons suffering major depression (Colombetti & Ratcliffe, 2012; Doerr-Zegers et al., 2017; Fuchs, 2005). In Cotard delusion the experience of disembodiment can go as far as feeling one's body to be dead. These phenomena would be type-1 disturbances, and I will also call them insensitivity phenomena.

We should expect type-1 disturbances to be associated with certain disruptions to intersubjectivity. If the body is experienced as inert, that is, as no longer sensitive, then it may become difficult for the other – whom we encounter with our body as a body – to impress upon one's sensibility. If the site of encountering the other is sensibility and one's body becomes, in various ways and degrees, insensitive, then there would also occur an insensitivity to the other. In a sense, if one's embodiment is already disrupted, the disruptive encounter with the other can become less impactful. Indeed, one might understand that they are faced by an other, even a demanding other, and perhaps further that they *ought* respond in some way, yet nevertheless not

---

[2] Husserl (1989) goes as far as to claim that even ghosts (considered as an empty, a priori possibility) have their own, ghostly bodies, that is, individuation in space.

be motivated to do so. This seems to correspond with depressive experiences in which guilt is still operative, implying that one is aware of one's own responsibility, yet nevertheless motivation to act – or even action itself – might seem impossible. Furthermore, if self/other ambiguity arises from a disruption of ordinary embodiment as a kind of taking on the other's perspective, an insensitivity to the other should correspond with diminished ambiguity. This would partly explain the close association between depression and loneliness: one may feel less a part of the world – even if one can see it, it might appear inaccessible – and detached from a sense of community; one may feel as if on their own. Varga and Krueger (2013) describe the felt loss of interpersonal connectedness as an important aspect of depression.

On the other hand, paralysis and uncoordinated movement would likewise correspond to an inability to co-regulate in the normal ways associated with self-other ambiguity. It is here where body schema and the present sense of embodiment come apart: disruption of body schema does not imply insensitivity. So long as the body remains sensitive, such circumstances might not hinder one at all from substituting oneself for the other, that is, from nevertheless being disrupted by others in such a way as to have one's attention redirected according to their perspective (even if in abnormal ways). That is to say, although it is still useful in these cases to talk in terms of body schema and image, it is not sufficient; along with sensitivity they cannot be doubted as legitimate aspects of embodiment. I cannot pretend to speak expertly on the intersubjective implications of these latter conditions – e.g., paralysis or uncoordinated movement – though I suspect they will differ from insensitivity disruptions.

It is often difficult if not impossible to tell if someone is experiencing depression, sometimes even in severe cases. This is one reason why laymen, at least, are often shocked to discover the severity of a depressed person's condition. Communication, rather than perceptual

189

observation, is crucial for determining whether someone is undergoing depression. Granted, the person herself may not realize that what she is experiencing is what we would call depression, but nevertheless it through a communicative expression of her experience and the dialogue that foments around that experience that we come to make such an ascription. According to Fuchs (2005), melancholic depression involves corporealization, in which the body feels weighty and rigid, an obstacle to the world rather than an exposure. This also corresponds with weakened sense perception. Motor function is disrupted but disfunction can be overcome (except in severe cases) by an "effort of will." Nezlek, Imbrie, and Shean (1994) and Nezlek, Hamptom, and Shean (2000) find that depressed persons in general have the same quantity of interactions as do non-depressed persons, but that these interactions are less enjoyable and that depressed persons perceive themselves as less responsive, influential, and confident when interacting with others. On my account, depression involves a disruption of sensibility understood not only as body schema and image but also as exposure to the environment and enjoyment. This is one reason why it is useful to think in terms of sensibility, as well as body schema and image: it helps us to capture this element of enjoyment.

Furthermore, even if insensitive to the other, one remains open *viz* coupling to intersubjective sense-making – even if this requires an effort of will – and, as seen in *Chapter 4*, this means that one can nevertheless learn common morality and *understand* responsibility (although schematic and image issues may result in abnormal sense, too). Understanding has its own affective force: knowing one is responsible does matter. But insensitivity of the body can flatten this affect as much as any other. This understanding of morality but general insensitivity might feed into the guilt that is often associated with depression, of which I say more below.

Standing in a dynamic relation with the intersubjective disturbances associated with insensitivity, type-1 disturbances will also correspond to certain forms of self-disturbances proper. If the ordinary self/other ambiguity that constitutes intersubjectivity is sufficiently lessened, as in insensitivity phenomena, this will also affect the kind of self constituted out from such an ambiguity. One might experience oneself as guilty, even if not for any particular reason. This corresponds to what Ratcliffe (2010) calls "deep guilt" in depression: rather than feeling guilty about this or that particular thing (though, this can also occur), guilt becomes a way of experiencing the world. This makes sense if insensitivity can coincide with an awareness of responsibility to others without a sensitivity to those others for whom one is supposed to respond, or even a sensitivity to that very responsibility of which one is aware. One could feel inadequately responsive to the other regardless of their actual responses (again, Nezlek et al. 1994 and Nezlek et al. 2000 suggest this is indeed the case), and so the self that is constituted is experienced as a *guilty self*. We could also expect that this kind of disruption to embodiment and the corresponding disruption to intersubjectivity would result in a certain kind of self-centered-ness. This should not be confused with caring only about oneself, but rather I mean that one might emphasize or focus on themselves, since self/other ambiguity plays a decreased role in delimiting or displacing the self. Again, this is reflected in the kind of guilt experienced by persons with depression, which Ratcliffe describes as involving "a more profound sense of estrangement from all other people," thus foreclosing "our guilt" in favor of a deeply personal guilt. Where I disagree with Ratcliffe, however, concerns whether this kind of personal guilt implies a relation to others; I think that it does, and that the relationship is one that makes me aware of my responsibility while simultaneously thematizing the inadequacy of any possible response.

It might be thought that this kind of disruption to embodiment can also occur in schizophrenia, but this is not obvious. Afterall, persons diagnosed with schizophrenia *also* report self-other fluidity or ambiguity, which has been described by Froese and Krueger (2021) as a *heightened*, not diminished, sensitivity or vulnerability to others. To con-fuse oneself with the other presupposes a certain degree of sensitivity. This approach can also account for the commonly held position that what is really disrupted in such cases is *common sense*, or the common world: The person with schizophrenia experiencing such con-fusion remains sensitive and vulnerable to the environment and makes *some sense or other* of that environment. Perhaps that sense is rightly construed as idiosyncratic, even akin to the "*ídios kósmos* of the dreamer" (Fuchs, 2020, citing Heraclitus), but is sense and sensitivity nonetheless, and to such a degree that the very vulnerability of the body may become thematized in experiences of threatening relationality (I'll say more about this when considering type-3 disturbances below).

One might suspect that the well-documented "anomalous perceptions" associated with schizophrenia, including abnormal perception of space and objects (Silverstein, Demmin, & Škodlar, 2017), visual and auditory distortions (Silverstein, 2016), and object recognition and naming (Gabrovska et al., 2002), are another kind of disruption to embodiment. On my account, however, such perceptual phenomena are not necessarily indicative of disruptions to embodiment. It is in relation to the other that we thematize objects for themselves, allowing X to become salient in our attention, recognizing X *as* Y and naming it. These symptoms might then be better classified as type-2 or type-3. Schizophrenia, however, involves such a diversity of symptoms – some maintain that it really names no one single condition at all – that we should not rule out type-1 disturbances altogether. I can here conclude only that the phenomena presently under consideration do not seem to fit the bill.

Type 2 Disturbances: From Embodiment to Substitution (Narcissism)

While type-1 disturbances were characterized in terms of embodiment – one's basic, sensitive exposure to the world, setting up an affective phenomenal field– types-2 and -3 refer to transitions. Type-2 deals with substitution, which refers to the way that, faced by an other who expresses herself, one's embodied enjoyment is called into question, and attention is thereby redirected towards the other and the world, so as to allow the body to take on a perspective other than one's own. Substitution is not only a perceptual perspective taking, though it is also that; it is through the modulation of attention that one comes to tend to a shared world thus thematized. It is also a displacing of the interests of bodily enjoyment in response to the demands made by the other.

Ordinarily, substitution achieves a kind of self/other ambiguity as well as a differentiation out of that ambiguity in virtue of having to account for one's own embodied being. The other directs my attention to herself, myself, and the world, in such a way as to bestow what was previously mere sensible appearance with a significance beyond myself. I account for myself by thematizing myself, and the world that appears to me, and allowing these thematizations – guided by the other's demand – to orient my response.[3] (This is not entirely distinct from Varela's conception of a narrative account, though rather than such an account being occasioned by myself and only accomplished in language, it is occasioned by language and accomplished intentionally; intentionality flows from an encounter with the other). Expression establishes ambiguity by substituting my own perspective with that of the other, and accounting for myself

---

[3] Diane Perpich (2008, p. 93) writes of account: "The face invites the ego or commands it – and the difference is difficult to determine on Levinas's account – to enter into a rational discourse, meaning a discourse that puts the world in common between us and demands of me an account recognizable by you."

differentiates me from that ambiguity.[4] In a type-2 disturbance, however, that ambiguity is not established. That is to say, one's own bodily enjoyment of the world is not displaced in the encounter with the other. This might loosely correspond to Froese and Krueger's 'loss of the socially extended mind', but with a significance beyond sensorimotor co-regulation, understanding, or knowledge. Indeed, remaining a sensitive body, it is entirely possible for a type-2 disturbance to nevertheless coexist with not only sensorimotor co-regulation but also common sense (or "doxic symmetry", in the terminology of *Chapter 4*). The loss of the socially extended mind, here, refers to a lack of the displacement of my own bodily interest in favor of the other. Remaining in egoism, I remain a sensitive body that nevertheless does not substitute itself for the other. The other may appear as an object, or even as another perspective, but, crucially, not as one that usurps my right to enjoy the world. Thus it is primarily the affective dimension of sense-making (recognizing that this is not sharply distinguishable from cognition in general) that is called into question here, specifically value and emotional connection.

Type-2 disturbances seem to be a good fit for characterizing grandiose narcissistic personality disorder. If self/other ambiguity is not even pre-reflectively recognized, then one's concern would primarily remain with oneself and the world will be experienced as for the ego.[5] In a sense, narcissists can certainly recognize that the other is a person besides them, but they do not recognize them as valuable despite oneself. Though a grandiose narcissist will be able to

---

[4] If I am dealing only with myself and there is, in consciousness, no awareness of the other at all, then there is no differentiation to be made between myself and the other; it is only in communion with the other that I can be differentiated *as I*. In that sense, the ego is not yet differentiated from the other: it is a differentiation of me from the world that appears for me, but only a differentiation in terms of the activity required to appropriate that world; it is not a differentiation of me from the other since, in egoism, I am "deaf" to the other (rather than "against" them).

[5] I am here assuming that self/other ambiguity is closely connected to the idea of the socially extended mind in general. Self/other ambiguity refers to the fact that our own agency, identity, and also what matters – even the value of the other mattering to me *despite* my own self-interest – is bound up with the agency, identities, and concerns of others. The socially extended mind hypothesis is roughly this idea, though it has emphasized agency and identity, as well as my own concern, over the possibility of usurping my own concern not for the sake of an interaction but for the other as such. *Chapter 4* explored these themes.

inhabit or understand a common world, in virtue of the broadly epistemic functions of sensorimotor coupling, in terms of value one will experience oneself as a center of the world. A lack of substitution helps to explain the "interpersonal dominance" (Zajenkowski et al., 2018) and interpersonal exploitation (American Psychiatric Association, 2013, pp. 669-70) characteristic of grandiose narcissists.

The dimensions of embodiment and self corresponding to a narcissistic type-2 disruption are closely connected. The kind of self constituted in such a case will be one given (mistakenly) as independent of others. As Otto Kernberg notes, "pathological narcissism is always characterized by the crystallization of a pathological grandiose self" (2014, p. 866). Rather than a responsive self which might be described as being-for-the-other, the narcissistic self will see others as being-for-the-self. This results not only in delusions of grandeur and self-importance (American Psychiatric Association, 2013, pp. 669-70), but also a lack of felt responsibility for others; the grandiose narcissist is primarily self-serving and resentful even of being associated with people seen as beneath them (which, as it so happens, is most people). In terms of embodiment, there is never a departure from experiencing the world in terms of enjoyment. The narcissist learns that there are other perspectives on the world, but not that these perspectives *matter as much or more than* his own. The type (2) disruption corresponds with what Levinas, citing Pascal, describes as the "usurpation of the whole world": one sees the world as "*my* place in the sun" (Levinas, 1981, p. viii, citing Pascal's Pensées, p. 404). The narcissist does not allow his place in the sun to be called into question but is, instead, content in his own enjoyment of the world.

Claire Katz, in "Education east of Eden: Levinas, the psychopath, and the paradox of responsibility" (2010) suggests that something like a type (2) disturbance also occurs in

psychopathy: the psychopath does not "see" the face of the other, which in one sense means they do not feel themselves to be responsible for the other and in another sense means that the face of the other does not call the psychopathic person out of oneself. This amounts to a claim that such a person does not substitute themselves for others. She uses the example of Cain and Abel, naming Cain the first psychopath, who does not put aside his own interests in service or responsibility to and for Abel. Persons diagnosed as psychopathic, according to Katz, do know right from wrong, but are simply not motivated by the distinction. This suggests that they can indeed entertain perspectives other than their own in the ordinary epistemic sense, but do not recognize the irreducible value of those perspectives. (The psychopath is perhaps a prototype for intersubjectivity without ethics. When discussing the Seattle School below, we ought keep in mind the implications of suggesting that someone does not stand in an ethical relation to the other.)

This description of psychopathy is remarkably similar to the description of grandiose narcissism I offered above, and there are reasons to think the former involves the same type (2) disturbances as does the latter. Psychopathy, like narcissism, involves a manipulative or exploitative orientation towards others, egocentricity or self-indulgence, and grandiosity (Widiger, 2006; he also provides a useful overview of comorbidities of and other connections between psychopathy and other pathological experience, including narcissism). Stone (1993, p. 292) maintains that psychopathy entails narcissism, with Hart and Hare (1998, p. 429) calling the former a higher-order construct involving narcissism as one aspect. It may be, then, that type (2) disturbances can explain the narcissistic dimensions of psychopathy, though psychopathy might involve other aspects as well.

On the other hand, the self/other ambiguity reported by some persons diagnosed with schizophrenia – the con-fusion between oneself and the other such that boundaries become fluid and the two seem to fuse or merge within experience – seems to resist a type (2) analysis. Not only is naïve embodiment disrupted by an encounter with the other, but such encounters can also be experienced as a threat to one's own identity. Furthermore, some persons who experience schizophrenia report the intersubjective encounter as the end of the world, which I can interpret as an end to the world of naïve embodiment, as well as a feeling of supporting the world or being responsible for the world in a cosmic sense (more on these reports later). It is thus not the case that persons who experience schizophrenia remain in a world unperturbed by the other nor that they necessarily feel irresponsible to others.

Type-3 Disturbances: Transition from Substitution to Self (Schizophrenia)

Type-3 disturbances pertain to the constitution of the self, which occurs in response to the other. Ordinarily, the self is a way of accounting for my embodiment and activity; the thematization originating in an encounter with the other applies to my body as much as it does to the world and others. In this sense, the self is embodiment thrown back on itself in responsibility. A disruption to the constitution of the self from out of a self-other ambiguity instituted as substitution correlates loosely with Froese and Krueger's notion of being 'lost in the socially extended mind,' but again the Levinasian perspective allows us to articulate another dimension of significance to such phenomena. A type-3 disturbance refers to an undifferentiated or otherwise unstable self, that is, difficulties accounting for oneself and determining one's own boundaries, limitations, powers, etc. These disturbances will thus carry with them disruptions to the sense of ownership and agency well explored in relationship to certain schizophrenia phenomena, such as thought insertion (see, for example, Gallagher, 2015): indeed, schizophrenia

197

is often considered a self-disorder because it undermines one or more such senses commonly associated with the concept of a self. If one cannot account for oneself in relation to others and the world, then it will be unclear what belongs to or is properly myself, and which actions are my own agentive productions. If 4e approaches and participatory sense-making can teach us anything, it is that our own agency is tied up with material and social realities, such that my own agency is not to be located exclusively "in me" (given some notion of "in"). Yet, we draw certain lines in order to account for ourselves and it is this accounting for oneself that is called into question in type-3 disturbances.

Type-3 disturbances will of course relate back to substitution and embodiment. At the level of self/other ambiguity, a type-3 disturbance will be associated with salient con-fusion: without differentiating oneself from others, I will experience myself as fused, merged, or confused with others. It is only self-differentiation that allow self and other to be intentionally distinguished. Otherwise, there is merely a sensitive body wrapped up in its relations with the world, and not yet a self constituted as an account of that body and not others. The threat that others pose to my identity – on my reading, to my naïve embodiment – would be heightened if not mitigated by the formation of a self that would limit that ambiguity. Responding to others would become difficult, that is, action or executive function becomes difficult (Orellana & Slachevsky, 2013), which would feed-back to further disrupt the sensorimotor coupling necessary for making sense of the world and others, exacerbating these difficulties. Thus, we would expect cognitive/affective impairments in social encounters in general. Importantly, then, a type-3 disorder, unlike a type-2 disturbance, will make the inhabitation of a joint world much more difficult, since in type-2 (and even, sometimes, type-1) disturbances, sensorimotor coupling and therefore sense-making might remain relatively in-tact. Lastly, if one is unable to account for

their own identity – that is, for their own embodiment, activity, thoughts and emotions – and if, furthermore, one's perspective remains submerged (or "lost in") self/other ambiguity, we should expect for the demands made by others to be experienced as overwhelming. It is normally the sense of self that can serve not only as a basis for planning and initiating a response, but also as a basis for re-asserting to some extent one's own interests as a countermeasure to the demands made by others.

In terms of embodiment, type-3 disturbances are associated with alienation. There is some debate concerning whether schizophrenia is best understood as primarily a disruption of the social self (see, for instance, Kyselo, 2015) or a disruption of embodiment (e.g., Fuchs & Schlimme, 2009). While I here side with the former, there are undeniably significant aspects of schizophrenia corresponding to embodiment. Fuchs and Schlimme describe schizophrenia as involving a "weakening of the basic sense of self, a disruption of implicit bodily functioning and a disconnection from the intercorporeality with others" (2009, p. 571). If there were some disturbance to the differentiation of the self out from self/other ambiguity, we might expect all of this: if the self is not adequately differentiated from others, then we should expect a disruption to sense of self as well as a disruption of implicit bodily functioning, since the body has been disrupted in substitution but has not been accounted for in terms of oneself. There will also be a disruption to intercorporeality, since there will be a disruption to one's bodily responses. Stanghellini (2009) describes schizophrenic persons as "living and behaving like a soulless body or a disembodied spirit." Of course, the schizophrenic person literally *has* a body: they remain exposed to the world and to others. The point is that they do not identify with the body in the normal way, which is to say that they are not differentiated as a stable identity among the myriad sensations delivered to their body. The term "soulless body," while problematic, perhaps is more

accurate, if we understand the soul to correspond to personal identity. But, again, the schizophrenic person does embody the world sensitively; they are just alienated from that sensitivity as it is not necessarily incorporated into a sense of self.

Type-3 accounts could easily be seen as consistent with approaches that focus on the disruption of ownership and agency in schizophrenic thought insertion and delusions of control, since in both cases what is in question are aspects of one's own identity. Delusions of control pertain to both bodily movements as well as mental states (see Martin Riemer, 2018), and describe an experience, associated with schizophrenia, where one's own acts or states are experienced as controlled not by oneself but, rather, by another. It is highly contentious whether the difficulty here is a matter of sense of agency – these are my movements and mental states, but I did not bring them about – or sense of ownership – these movements are not my own. Type-3 disruptions could be spun either way: what is called into question is the self as an account of bodily activity and experience alike; the self as responsive to the world around it. One lesson learned from dynamical systems theory and the extended mind hypothesis is that our actions and identities are entangled with forces external to us; failing to differentiate oneself from those sources could easily lead to an experience of alien control. Thought insertion often follows a similar analysis, including the debate between agency and ownership (see Gallagher, 2015).

Summary of Types 1-3

I have proposed that Levinas' conceptions of embodiment as vulnerable sensibility, substitution as a kind of perspective taking that institutes an experiential self/other ambiguity, and the self as constituted in response to the other, help us to think of certain pathology phenomena in new ways. A type-1 disturbance refers to bodily insensitivity, that is, a

modification of the *enjoyment* relation characterizing bodily existence in the world. As articulated in *Chapter 2* above, enjoyment does not only refer to pleasure but rather to an entire affectivity. If one *enjoys the world* abnormally, we will expect a different relationship to self as well as others. The Levinasian conception of embodiment as sensitivity (or vulnerable exposure, characterized in terms of enjoyment) is irreducible to either body image or body schema, as it is through being open to a world that we enjoy that we come to develop any sort of bodily style, habitus, or self-conception at all. As such, this conception constitutes another tool to be used alongside those formulations of image and schema, and I have proposed that we can understand some important symptoms of depression as type-1 phenomena, including not only insensitivity but also intersubjective alienation, guilty self-consciousness, and self-narrative.

A type-2 disturbance refers to substitution, or the way that my own enjoyment of the world is called into question by the presence of the other, such that I become for-the-other rather than for-myself. This helps us to understand the constitution of a sense of objectivity, in a way partly analogous to Husserl's position, but also to understand how we come to stand in relation to an other who matters independently of oneself. To talk about a *disturbance* regarding substitution is to invoke the idea of one who does not exist for-the-other, but for-the-self; ordinary bodily enjoyment remains a privileged mode of existence. As such, I propose that we can understand grandiose narcissism and psychopathy (to the extent that it involves narcissism) as involving type-2 disturbances, which are related to delusions of grandeur, a sense of entitlement, exploitation of others, and embodying the world in terms of one's own enjoyment.

Type-3 disturbances are the most difficult to formulate, and for this reason I rely on Froese and Krueger's formulations of being lost in and losing the socially extended mind. Substitution enacts a kind of ambiguity into experience, whereby I relate to something outside of

me through responsibility. Yet, ordinary existence in a society consisting of many others nevertheless requires that I construct my own identity and, to some extent, protect my own interest as well. One way to formulate this idea is to refer to the transition from ethics to justice, characterized by a violent generalization that constitutes me as one among the others. To be overwhelmed by one's relation to others is to be lost in the socially extended mind, so to say. A type-3 disturbance refers to this being lost, that is, to a non-differentiation of self out from the ethical self-other ambiguity constituted by substitution. What I add to Froese and Krueger's account is an articulation of the ethical aspects of such a relation. I propose that certain phenomena associated with schizophrenia – admittedly a very difficult case – can be in part understood in terms of type-3 disturbances, relating to the constitution of the self; without emerging from the self/other ambiguity instituted by substitution, others will be experienced as threatening both one's world and identity, response becomes difficult, and we should expect self/other con-fusion as well as phenomena like delusions of control and thought insertion, which can be construed as kinds of self-alienation. Importantly, the idea of being a "support for the world," and the idea of the other as profoundly threatening my own agency and existence, can easily be understood in terms of the overwhelming position represented by a type-3 disturbance.

I must stress that, on this account, while embodiment is meant to refer to a basic fact of human existence – we exist as sensitive bodies (even if this sensitivity can be diminished) – substitution and the constitution of the self are not individualistic achievements or accomplishments. Rather, they are inherently social achievements that refer to but are not produced exclusively by one's own capacities: it is the other who directs our attention, and it is in response to the other that the self is constituted.

Therapeutic Implications

I can only briefly cover some implications for therapeutic intervention here. We should expect type-1 disturbances to respond well to interventions that foster bodily sensitivity and sensitivity to others. This is born out by the efficacy of dance therapy, for example, in depression. According to Michelle Maiese, dance/movement therapy (or DMT) "centers on the use of movement to further physical and emotional integrity and looks to body movement as a primary means of communication" (2018, p. 14). She goes on to describe that

> Typical treatment sessions might include rhythmic dance, spontaneous and creative
> movement sequences, thematic movement improvisations, unconscious symbolic body
> movement, group dance, and a range of relaxation exercises. (2018, pp. 14-5)

On our account, DMT works because it engages the body in an activity that requires attunement to sensitivity (Maiese might describe this as "revitalization"). This involves moving one's body in a sensitive way and, especially in group dance, responding sensitively to others. The emphasis on the body on communication highlights the relevance of the Levinasian framework, where rather than serving a primarily perceptual function, bodily movements serve as address and response; they make salient demands made on me and my responses to those demands. Koch, Kunz, Lykou and Cruz (2014), Krueger (2018), Martin, Koch, Hiriak, and Fuchs (2016), Röhricht & Priebe (2006), and others, have suggested that DMT might be useful in treating certain forms of schizophrenia as well. Due to the embodied aspects of some schizophrenic experience, that might be so, and connecting with one's body might plausibly condition the constitution of the self. But I emphasize that DMT would only be a partial intervention, as it addresses type-1 disturbances which are secondary to the type-3 disturbance I take as primary in schizophrenia.

Type-2 disturbances would require intervention that would decenter the self in relation to the other, and impose on the self the independent value of others. Freudian approaches to narcissism naturally focus on the subject's relationship with their parents and the question of having the narcissist's needs met. One proposed therapy, for instance, is called "reparenting," and involves "identifying core unmet needs" and "working to help the patient get those needs met" (Behary & Dieckmann, 2013). I have suggested, however, that an important aspect of narcissism is not the relationship to any particular other but to others in general: the grandiose narcissist (noting that "reparenting" might be ideal for vulnerable narcissism) sees themselves as unique and independently valuable, while denying these traits in others. Something like the proposal offered by the Seattle School of Psychology (explored in detail below), i.e., to restore responsibility for others to the patient, might be apt here.

Type-3 disturbances would require intervention that helps someone to account for themselves, that is, to differentiate oneself – including one's agency, thoughts, etc. – in response to others; to responsively differentiate oneself from others and the world, and to present this account to others. The constitution of the self is a complex matter mediated not only by one's embodiment but also one's social positioning. Indeed, very different selves are appropriate in very different circumstances, and the difficulty of accounting for oneself – what to account for in what ways, what to attribute to myself at what time, and what to exclude – is demonstrated by the idea of the socially extended mind: I am bound up with the world and others. On the participatory sense-making framework (and I agree), one is not even individuated as a self except through an encounter with others. Furthermore, persons diagnosed with schizophrenia vary widely in their symptomology; any accounting of oneself must be accountable to these differences. Thus, the constitution of the self cannot be seen as an individualistic

accomplishment, cannot be determined in character ahead of time, and should be focused only on allowing one to strive in their social and material relations. It may turn out that one instance of such a self would be quite minimal indeed, while others more robust.

The Seattle School: Ethical Deficiency

I now turn to consider the Seattle School approach, which I have alluded to above. The Seattle School takes Levinas to be a central figure, and focuses on the responsibility of the person who enters into therapy. In *Chapter 1*, we saw that the deficit method, by which we understand others across difference by supposing that an otherwise ever-present feature of consciousness is, for those persons considered, absent, is epistemologically and ethically problematic: it can only conceive what phenomena are *not* rather than what they are, reduce people to *less than* normative consciousness, and, in the case of structural phenomenology, illicitly apply individual structures that can only be verified intersubjectively. If the deficit method considered there involves *ontological* deficiency – a deficit in the being of one's conscious experience (a lack in some fundamental structure constitutive of ordinary experience), then some thinkers in the Seattle School of Psychology proffer, instead, *ethical* deficiency. The proposed type-1 through -3 disturbances above might be seen as lending themselves to this kind of interpretation.

I must note that I here distance myself from the deficit method in general because what I am exploring are not individualistic (or "transcendental subjective") structures, but rather relational accomplishments the absence of which do not necessarily imply deficiency on the part of any one individual.

If ontological deficiency was problematic, it is even easier to see the trouble with ethical deficiency. The claim is, effectively, that persons suffering mental illness are irresponsible or

unethical, and that the job of the therapist is to return such persons to responsible relations with others. This position follows from a particular supposition, made explicit by George Sayre: "we are assuming that it is psychologically healthy to be ethically good" (Sayre, 2015, p. 161). Ethical responsibility is thus taken as the standard of mental health, with the consequence that mental illness involves a departure from ethical responsibility. Kunz draws this consequence, writing that

> Pathology is isolation from ethical responsibility. Health is getting outside oneself in responsibility to others. Therapy comes when the client becomes responsibly responsible toward others in his life. (Kunz, 2015, p. 230)

This sentiment is shared by Sayre, for whom "therapy [is] not about their [the client's] needs but about the needs of another person – but not myself" (Sayre, 2015, p. 167), against "the traditional ethic of therapy [as] an ego-logical one in which the client is self-centered" (Sayre, 2015, p. 164). Likewise, David Goodman and Brian Becker agree with C. Fred Alford that "the therapeutic goal is not to refound or reground or integrate the self. The goal is to find productive – that is, involved with other humans – ways to give oneself away" (Alford, 2002, pp. 73-74). They go further, however, in questioning the client's preparedness for such practice and thus "posit that psychotherapy has as its goal an awakening of the self to the Other" (Goodman & Becker, 2015, p. 156). This view is also shared by Richard Williams who, noting that the client comes to therapy in "neediness and the quest for satisfaction," implies that psychotherapy has to move beyond such a quest, i.e., must respond instead return the client to responsibility for others (Williams, 2015, p. 137).

The idea that mental illness involves a departure from responsibility – the return to which is the goal of therapy – is problematic for at least three reasons. First, it is dangerous and

206

stigmatizing to conceive of persons who suffer mental illness as irresponsible or unethical. I take this point to be so obvious as to not warrant further attention here. Second, it is apparently false: at least some – probably most – persons who experience pathology phenomena *do* stand in ethical relationships to others, even if those relationships are not without their difficulties. (Inversely, some people we might view as mentally healthy are unethical. One could consider the case of *Eichmann in Jerusalem*, in which the eponymous Nazi leader and one engineer of their atrocities was deemed mentally healthy and even to possess many desirable traits. That such a diagnosis might only be possible in a sick society is, today, of little comfort.) Third, proponents of a Levinasian position, such as that of the Seattle School, must hold in tension the seemingly contrary insights that we are responsible even for others' responsibility, and yet whether the other is responsible to me is "his affair," i.e., none of my business. The second and third points require elaboration.

Pathology and the Ethical Relationship

For Levinas, recall that the ethical relationship is one in which the other makes a demand on me through her expression, which de-centers or displaces my own naïve existence. Anyone experiencing self/other con-fusion already stands in a relationship to the other that can properly be called ethical, even if it does not yet rise to the level of constituting a responsive self. In the case of self/other ambiguity, it seems that the issue is precisely the differentiation of an I from out of the ambiguity – one is lost in the socially extended mind, to use Froese and Kruger's terminology, rather than emerging as a self, (or "I") from it.

There are other reports of pathology phenomena that speak against the formulation of mental illness as irresponsibility. In schizophrenia, for instance, much has been made of reports that the person affected experiences herself as the center of the world or universe, supporting it

all (Fuchs, 2020; Payne, 2012; Sass, 2001). On the one hand, Thomas Fuchs argues that such reports indicate a break with the common world and enactment of a solipsistic world in which all sense originates subjectively rather than intersubjectively. There is an inability, according to Fuchs, to take up an "excentric" perspective, i.e., to decenter oneself relative to the world or, put another way, to relativize one's own experience as just one perspective among others (Fuchs, 2020). On the other hand, Sass argues that such reports do indicate a kind of excentricity, specifically the ability to step outside of a social "common sense" and be open to new, unanticipated possibilities (Sass, 2020).

I have implied that even the "solipsistic" world of schizophrenia is one already disrupted by others; others are not entirely absent from such a world at all (and this is one more way that such a solipsism must be qualified as "quasi-solipsism"). Self/other con-fusion demonstrates the point, but other examples are possible. In schizophrenia, others are often reported as threatening figures, figures that threaten one's own identity (Lysaker, Johannesen, & Lysaker, 2005). This threatening atmosphere is not localized to experience of others but even to the whole of one's world. Jaspers, on the topic of "cosmic experiences", cites self-reports and analysis by A. Wetzel to the effect that the world is sometimes experienced as *the end of the world*, as collapse, annhilation, or catastrophe (Jaspers, 1997, pp. 294-5).

The resonances to Levinasian thought are significant. The encounter with the other – the ethical relation – must bear with it precisely the possibility to break with common sense. That is, our responsibility to the other is not exhausted by any established norms, practices, or even the common meanings and significances we impose onto the world. The break with common sense found in certain pathology phenomena, then, might be a prolonged, intense, even problematic version of something crucial to ethics: the possibility of acting outside of established

208

possibilities. Insofar as this break is established within the face-to-face encounter, no matter how

disturbing the break may be, that relationship may for that very reason be characterized as

"ethical," even if it is combined with suffering. This lends itself to a further resonance: the face-

to-face encounter with the other is not just described by Levinas as peace ("Peace and

Proximity" in Levinas, 1996) but also as violence, persecution, and trauma (Levinas, 1969, p. 25;

Levinas, 1981, p. xii, throughout). The other, in a sense, heralds the end of my world.

Intersubjectivity is the end of the subjective world, the overcoming of hedonistic, naïve being.

The other threatens my own identity – I am substituted for the other in an inescapable

responsibility – and this threat seems reflected in some reports of schizophrenic insecurity in the

face of the other (Lysaker, Johannesen, & Lysaker, 2005). I support the world by being he who is

called to thematize the world not for myself but for all the others. Levinas reminds us of the

ethical significance of this support with his favorite quote from Dostoevsky's *The Brothers

Karamazov*: "We are all responsible for everyone else – but I am more responsible than all the

others." On this account, some delusions of persecution may not be delusions at all, but rather

reflect an aspect of the intersubjective relationship as such.

I am not suggesting here that there are not intersubjective difficulties experienced by

persons diagnosed with schizophrenia (as detailed above) but, rather, that the enactment of any

thematizable world whatsoever, no matter how solipsistic it may appear, already refers to others

who call us out of a naïve embodiment in which significance – experiencing X *as* Y,

experiencing X as something beyond its immediate appearance for me – is not yet possible. The

very fact that the person diagnosed with schizophrenia engages in speech implies or *is* a relation

between that person and he to whom she speaks, him for whom she formulates her world and

expresses it. In *Totality and Infinity*, Levinas considers a silent world, one in which others do not

respond to my inquiries and argues that this world nevertheless refers to an other conceived of as an evil genius: the very presence of signs, the significance of sensible things precisely as signifying something *behind* their mere appearances *for me,* refers to an other lurking behind the signs; the appearances must signify for some other beyond me for them to signify beyond their own appearances at all. But if the other does not respond to my formulations, the world is transformed into a kind of apparition and the presence of the other becomes menacing and ridiculous; the other becomes a threatening presence that threatens the world and my own identity (Levinas, 1969, p. 94). The silent world is a world of calamity, a world of unstable or idiosyncratic appearances, but a world set-up in relation to an other nonetheless.

That is all to say, there may be an issue when it comes to response and to constituting a stable, differentiated identity out from an ambiguous relationship with others, but an ambiguous relationship with an other is still an ethical relationship in the sense that it is still a calling out of oneself towards what one is not. That one experiences the world differently than do others, and that one knows this, already presupposes a relationship to others. But it is not the epistemological relationship of confirmation and correction taken to characterize intersubjectivity since Husserl.

To elaborate the issue with another example, consider depression. In depression, there is often reported a perceived inability to do or change anything at all (Ratcliffe, 2014). In such a case, one is not avoiding responsibility but reporting a foreclosure of certain possibilities of response to the other. Furthermore, for many people who experience major depression, their incapacities are associated with guilt and shame (Ratcliffe 2010). This further demonstrates that, in at least some forms of psychopathology, there is already an ethical relationship to others, already a relationship even of responsibility, but also an inability (not avoidance) to act on that responsibility well.

Dual Insights: Responsibility for Responsibility, and Asymmetry

Let us return to the third point, i.e., that a Levinasian position must account for the seemingly contrasting insights of responsibility for the responsibility of the other and the asymmetry insight, that whether the other is responsible to me is his affair. In some ways, the Seattle School does a good job here, in respect to both insights but especially the first. They take seriously the responsibility for the other's responsibility, taking therapy to be a fostering of responsibility in the other. Sayre is most explicit about these simultaneous insights when he writes that "therapy [is] not about their [the client's] needs but about the needs of another person – but not myself" (Sayre, 2015, p. 167). However, the motivating claim is problematic, namely that in describing ethical subjectivity Levinas is providing a model for psychological health. Levinas is writing precisely for his audience, that is, for those who read Levinas. It is the reader's own self-interest and freedom he wishes to challenge, for the sake of the claims made against the reader by others. Levinas is challenging the reader, not those others with whom the reader interacts. He is not arguing that ethical responsibility is equivalent to or necessary for psychological health (at times he even describes responsibility as obsession, a term we would not necessarily associate with health), and therefore that psychological illness amounts to ethical irresponsibility. Williams (2015) goes astray, then, when he argues that instead of directing ourselves to the needs of the other we address their responsibility for still other others. Reminding us, by quoting Rabbi Israel Salanter, that "the Other's material needs are my spiritual needs" (Cohen, 1986, p. 24), Levinas calls us to respond to both the other's responsibility *and* their needs. Accordingly, the task of psychotherapy is not only to foster responsibility but also to act out of one's own responsibility for the client's needs, as varied they may be. A corollary of this idea is that a truly Levinasian psychiatry requires the ability to change the socio-material

situation of those who come for help; this may be impossible without radical changes to existing social organization.

Edwin Gantt has produced excellent work on this topic (Gantt, 2000). For Gantt, the purpose of therapy is to suffer-with and -for the other. That is, the other comes to therapy in suffering, and the therapist's job is not first and foremost to understand, nor to return the other to responsible relationships, but rather to take on the suffering of the patient as one's own. This means to allow that suffering to matter to the therapist, to allow their suffering-for-the-other to disrupt their preconceived theoretical constructs and to orient therapy towards what actually matters to the suffering person. Gantt thus takes seriously the Levinasian challenge to self-interest by accepting that Levinas' challenge is leveled at *him*, the therapist himself, rather than universalizing the challenge by conceiving of ethical responsibility as a universal structure of (healthy) consciousness and imposing the challenge on she who enters therapy looking for help. The demand that is relevant to (incumbent upon) the therapist is precisely the demand made on the therapist.

Conclusion: From the face-to-face to community

My goal here has been twofold: use Levinas to better understand psychopathology, and use psychopathology to better understand Levinas. I begun by juxtaposing some phenomenological and 4e conceptions of embodiment, self, and intersubjectivity with Levinas' notions of embodiment (as sensitivity), substitution (instituting self/other ambiguity through a certain form of perspective-taking), and self (constituted in response to the other). I then moved to consider some implications of the latter: disruption is not inherently pathological, and these concepts allow us to rethink a nexus of seemingly contrasting symptoms associated with schizophrenia. I also point out that sensitivity and intersubjectivity are not exhausted by the idea

of sensorimotor coupling, alluding to the possibility that the latter remain in-tact while the former be disrupted. I then provided a 3 types of disruption that Levinasian terminology allows us to articulate. That list is not meant to exhaust pathology phenomena or displace body image/schema oriented discussions wholesale, but rather to complement and complicate those analyses. My focus in that section was on depression, grandiose narcissism (and psychopathy), and certain symptoms associated with schizophrenia. The 3-part characterization of disturbances has some natural implications for therapeutic intervention, which I was only able to consider briefly. I then considered and argued against the Seattle School position that mental illness involves ethical irresponsibility, i.e., an ethical deficit, on three grounds. First, this a stigmatizing and dangerous position to hold; second, it is false in the sense that persons who experience pathology are not, simply *qua* mental illness, irresponsible; third, the view results in a therapeutic orientation away from the material needs of the other in favor of what could be described as proselytizing.

I now want to make two, final points. The first is to note that, for Levinas, substitution may already *be* responsibility – it is certainly at least an affective response to substitute oneself for the other. The transitionary phases I have articulated between embodiment and substitution, on the one hand, and substitution and the self, on the other, are therefore perhaps artificial. There are a few ways that we can interpret this. One is that psychopathology shows us that these elements can indeed come apart – that the transition from embodiment to self is not necessarily a smooth one. This is an appealing possibility, since it demonstrates a give-and-take relationship between Levinas and psychopathology out of which may emerge new, useful sense. Another possibility is to consider (which we can now only do most briefly) Levinas' thoughts on symmetry, principle, justice, or community. For Levinas, the face of the other is also the face of

all others. That is, that the other appears to me as other – as entirely beyond me – means that the other also signifies the relation with alterity in general. This is one reason why thought always collapses into systematic ontology.

As Richard Cohen expresses in his foreword to *Otherwise than Being*, "To give all to one, is to leave all others destitute" (Levinas, 1981, p. xvi). That is, I cannot be indifferent to the relationships between those who manifest themselves as my others. As soon as the "third" – another other – appears, I must take a stance on relations between multiple others. I cannot wash my hands of injustices between them. This generalizing move whereby I must weigh the interests not only of one other against myself but of each other against all others is a move that ultimately, if it is to consider *all* others, must result in universal (though never a priori) principles that, as such, apply to me as well. To quote at length:

> In proximity the other obsesses me according to the absolute asymmetry of signification, of the-one-for-the-other: I substitute myself for him, whereas no one can replace me, and the substitution of the one for the other does not signify the substitution of the other for the one. The relationship with the third party is an incessant correction of the asymmetry of proximity in which the face is looked at. There is a weighing, thought, objectification, and thus a decree in which my anarchic relationship with illeity is betrayed, but in which it is conveyed before us. There is betrayal of my anarchic relation with illeity, but also a new relationship with it: it is only thanks to God that, as a subject incomparable with the other, I am approached as an other by the others, that is, "for myself." "Thanks to God" I am another for the others. (Levinas, 1981, p. 158)

Translating the theological terminology into a quasi-theology, he immediately continues:

God is not involved as an alleged interlocutor: the reciprocal relationship binds me to the other man in the trace of transcendence, in illeity. The passing of God, of whom I can speak only by reference to this aid or grace, is precisely the reverting of the incomparable subject into a member of society. (Levinas, 1981, p. 158)

For those suspicious of theology, God is spoken of only as absolute alterity, as an idea of infinity to which all concepts are inadequate, revealed to thought only by the ethical relationship of face-to-face encounters. But as a general concept of otherness, asymmetry is betrayed and I myself am considered an other to others. The responsive I is incomparable, irreplaceable; this just is its identity. But in reciprocity – which, as noted in the previous chapter, does indeed characterize participatory sense-making or coming to understand one another – the responsive I is further transformed into a member of society.

I suggest, then, that in addition to types-1 through -3 above, we might also locate self-disturbances in terms of situating oneself with respect to one's community.[6] If one does not feel oneself supported by one's community, if one is legitimately excluded from that community, then this not only affects the sense of who one is and the way that one embodies their relationships with others in the world, but it correspondingly disrupts (or at least significantly conditions) participation in communal processes of sense-making. Therapy and theory alike, then, must account for one's community-situation, and thus not only must therapy respond to the concrete needs expressed by she who enters therapy, but a successful therapy would also have an eye to community level intervention, e.g., interventions that directly target one's social, political conditions.

---

[6] This kind of consideration might fit into type-2 and -3, but those types were considered here primarily with face-to-face encounters with individuals in mind.

Second, a major objection must be admitted: am I not just offering a new structural phenomenology, a new ontology by which to think these concepts? Am I not betraying the spirit of Levinas? Am I not totalizing the other by offering general concepts by which to understand the experience of individuals? Yes and no.

Yes, I am offering general concepts and I am constructing a kind of system through which to interpret others. However, as Levinas confesses himself, this is unavoidable. No, however, insofar as what matters is that whatever our systems might be, we base them first on the asymmetric ethical relationship with the other, are open to their rupture by others, and in the name of ethics disrupt such systems ourselves. In critiquing deficient methodology of every kind and in recognizing that what I provide here is in no way a total system (and certainly not a structure of transcendental subjectivity, understood in an individualistic sense), nor should it ever be privileged above the concrete expressions made by others, I believe I am doing the best that can be done if we nevertheless want our concepts to be useful for concrete human endeavors. We must generalize to meet the demands of justice, even if this betrays our anarchic face-to-face relationship with the other. The latter must be and remain she for whom we generalize: all others, for each of whom we seek justice. Furthermore, my concern is first and foremost with the demands made on *me*, and the demands made on interested parties, particularly those in a position of power over those making demands. That is, what I offer here are some tools to aid the therapist and the theorist in responding to the demands made on them by those who suffer. It is an incomplete system, open to rupture, oriented by responsibility expressed in demands made across seemingly radical difference.

Conclusion

Throughout this dissertation, I have pointed out limitations in contemporary phenomenological and enactivist approaches to psychopathology, and have offered correctives based primarily on Levinas' ethical conception of expression. The position offered here is broadly complementary to 4e and phenomenological approaches to cognitive science and psychopathology, though it is only partly compatible with the latter. The major addition is the articulation of an asymmetric, ethical significance to human interaction that is not exhausted by the epistemic, perceptual, or ontological, or participation-based conceptions – including direct perception theory, interaction theory, and participatory sense-making – that characterize phenomenological cognitive science and enactivism today. Recognizing this distinct form of significance allows us to conceptualize intersubjective understanding on the basis of communication with others across differences that may initially be – and perhaps even threaten to remain – incomprehensible. First we communicate, and only on that basis can we understand.

For this reason, the position articulated here can make sense of understanding in the context of psychopathology and can also provide us with a number of methodological and conceptual tools for theory and practice in that domain. Psychopathology is a field that thematizes seemingly radical differences: we try to understand others who experience the world very differently than ourselves. Rather than presupposing similarities taken to characterize human experience in general, or even some specific subset of human experience, our approach begins with taking seriously the demand placed on us by people who suffer. Response to suffering is the only adequate test of a psychopathological theory or practice, and thus should orient the field.

I have tried to distance the claims of this work from considerations of transcendental subjectivity, instead endorsing the position that identity and experience in general, as well as understanding, are produced in the encounter between persons. I have thus proffered some general insights into the ways that we seem to respond to one another, while reminding the reader to bare in mind the asymmetric nature of those relations. The theory of natural pedagogy and 4e approaches to sense-making, especially interaction theory and participatory sense-making, help us to bridge the gap between intersubjectivity as a primarily ethical relationship and understanding one another in terms of concepts and categories held in common.

Our concepts can be applied for understanding specific pathology phenomena, but it must be remembered that these insights can always be ruptured by the other; they refer to generalities and not universalities, and they must be responsive to individual differences. It must be noted that embodied existence, the expression of the other that interrupts that existence, and the self as constituted in response to that expression stand in dynamic relations to one another; the concept of naïve embodiment does not refer to a state temporally or developmentally prior to contact with others, but rather refers to the mode of being we would occupy without such contact. It is also a common and familiar occurrence to slip back into naïve embodiment, in which to some extent or other we orient ourselves to the world in terms of our own enjoyment rather than being-for-others. Even the therapist or pathology theorist can easily content himself with pre-fabricated concepts that do not serve the concrete demands made on them by others.

Throughout, I have taken a distinctive Levinasian inspiration, and have – especially early on – rendered Levinas' concepts and insights as faithfully as I could. However, when applying these concepts to any concrete human endeavor that relates itself to a plurality of persons, certain concessions must be made. I have at times, therefore, remained Levinasian only in my most basic

218

concepts and inspirations, and have bent derivative concepts to my purpose. This is especially clear in *Chapter 5*.

In *Chapter 1*, I explicated the foundational concepts of interpersonal understanding in phenomenological psychology by reviewing the work of Karl Jaspers, Edmund Husserl, and Edith Stein. For them, empathy is the primary intentional act directed towards the other, and so is either necessary for or identical with interpersonal understanding. Empathy, however, relies on similarities and familiarities that cannot be presupposed in general, let alone in the context of psychopathology. This conception of understanding leaves open the possibility of incomprehensibility and recommends, as a limited countermeasure, what I have called deficit methodology. That method was shown to be inadequate to contemporary demands. While transcendental phenomenology can enhance the method, it does not overcome all its shortcomings, and also carries out an illicit inference by applying structures of one's own consciousness to another person for whom those structures are precisely what is called into question; it presupposed understanding where understanding is precisely at issue.

In *Chapter 2*, I articulate Levinas' ethical conception of expression. Rather than defining expression in terms of recognizable bodily movements constituted as articulating co-existing, personally familiar mental states – as do Jaspers and the phenomenologists – expression is defined in terms of a disruption to one's embodied sensitivity. As a body, one is exposed to a world and this exposure sets up an affective, phenomenal field with its own morality: the body is content to enjoy the world without seeking a deeper significance outside this enjoyment. But the expression of the other acts as a disruptive force, calling me out of contentment by demanding a response. As placing a demand on me for which I am responsible, expression is an inherently ethical phenomenon.

In *Chapter 3*, I demonstrate how the dynamic of demand-and-response contributes to the production of understanding in the traditional sense. I co-opt the theory of natural pedagogy, interpreting pedagogical cues specifically as *address* in the disruptive sense, to connect the ethical conception of expression with understanding general features of the world, oneself, and others. I also suggest that dynamical systems theory approaches in cognitive science should emphasize the affective responses that one undergoes in the presence of the other so as to allow our lived and attentive sensitivity to the other to direct research. In this chapter, I rethink the notions of incomprehensibility and understanding introduced in *Chapter 1*. Understanding is not a default relationship that collapses into incomprehensibility only when something goes wrong. Instead, understanding must be produced as a response to the demands placed on one by the other; incomprehensibility, then, is a ground for understanding, even though we often find ourselves in familiar situations that already go a long way toward covering up that ground.

In *Chapter 4*, I pursue the production of understanding in relation to enactivist conceptions sense-making. This discussion is oriented by a concern with ethical experience: I want to account not only for responsibility but also the experience of responsibility and the extreme possibilities of altruism and self-sacrifice. That is, I want to understand interest despite oneself and, relatedly, sense despite self-interest. I begin with autopoietic accounts of the organism and move to interaction theory and participatory sense-making in order to set up a discussion of the enactivist literature on ethics. I argue that current attempts at ethical theory in enactivism have only considered ethics from the perspectives of normativity and affectivity, and in problematic ways that have either not sufficiently called into question the freedom or activity of individuals, or otherwise cover up ethical asymmetry and therefore unduly deflate the idea of responsibility. Understanding expression as disrupting organismic, autonomous, or symmetric

modes of sense-making, however, complements these approaches by providing a ground on which to make sense of ethical experience while nevertheless benefitting from those approaches' ability to account for higher-order, ethically relevant phenomena, e.g., the intersubjective production of norms and emotions.

In *Chapter 5*, I attempt a Levinasian contribution to phenomenological psychopathology. I begun with a brief explication of the connections between embodiment, intersubjectivity, and self in phenomenological and 4e approaches, before introducing my preferred versions of these concepts: embodiment as sensitivity, substitution as a certain form of perspective-taking (embodying the world according to the other's demands), and self as constituted in response to the other. I propose three "disruptions" that correspond to what have traditionally been known as "self-disruptions," noting that none of these three dimensions of experience are solipsistic accomplishments. I then apply these three disruptions to consider depression, grandiose narcissism (and, to some extent, psychopathy), and certain symptoms associated with schizophrenia, e.g., alienation from one's own body and experiencing the other as a threatening presence. I then consider a Levinas inspired approach known as the Seattle School of Psychotherapy. According to the Seattle School, mental health is characterized as a kind of ethical deficiency and the task of the therapist is to return patients to responsibility. Though there is much that is right about the approach in general, I reject the characterization of mental illness as ethical deficiency and note that Levinas, far from proposing a model of psychological health to be applied to oneself and others alike, is instead reminding his readers of their own responsibility.

Putting Levinas in contact with the literature on psychopathology is a challenge. There have been only a few attempts to do so. The challenge is specifically a challenge to established

psychopathology. I consider this project to be successful if any theorist or therapist learns to better respond to those who come to them in suffering because of it, or if any reader of Levinas can make better sense of certain of his concepts, even if only partially.

References

Adams, F. & Aizawa, K. (2010). *The bounds of cognition (2ⁿᵈ Edition)*. Oxford: Blackwell.

Agamben, G. (2003). *The open: Man and animal* (K. Attell, Trans.). Stanford University Press.

Alford, C. F. (2002). *Levinas, the Frankfurt school, and psychoanalysis*. Middletown, CT: Wesleyan University Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Atterton, P. (2011). Levinas and our moral responsibility toward other animals. *Inquiry, 54*(6).

Behary, W. T. & Dieckmann, E. (2013). Schema therapy for pathological narcissism: The art of adaptive reparenting. In J. S. Ogrodniczuk (Ed.), *Understanding and treating pathological narcissism* (pp. 285-300). American Psychiatric Press.

Bortolotti, L. & Broome, M. (2008). A role for ownership and authorship in the analysis of thought insertion. *Phenomenology and the Cognitive Sciences, 8*, 205-224.

Carpendale, J. I. M. & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27*, 79-151.

Carr, D. (1973). The fifth meditation and Husserl's Cartesianism. *Philosophy and Phenomenological Research, 34*, 14-35.

Carr, D. (1974). *Phenomenology and the problem of history: A study of Husserl's transcendental philosophy*. Evanston: Northwestern University Press.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: The MIT Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, action, and cognitive extension*. Oxford: Oxford University Press.

Clark, A. & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7-19.

Cohen, R. A. (1986). *Face to face with Levinas*. Albany: State University of New York Press.

Colombetti, G. (2017). *The feeling body: Affective science meets the enactive mind*. Massachusetts: MIT Press.

Colombetti, G. & Ratcliffe, M. (2012). Bodily feeling in depersonalization: a phenomenological account. *Emotion Review*, 4(2), 145-150.

Colombetti, G. & Roberts, T. (2014). Extending the extended mind: the case for extended affectivity. *Philosophical Studies*, 172, 1243-1263.

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

Corris, A. & Chemero, A. (2019). The broad scope of enactivism. *Adaptive Behavior*, 1-2.

Costelloe, T. M. (1998). Husserl's fifth meditation and the phenomenological sociology of Alfred Schutz. *Journal of British Society for Phenomenology, 29*(1), 23-46.

Critchley, S. & Bernasconi, R. (2002). *The Cambridge companion to Levinas*. Cambridge: Cambridge University Press.

Crowe, J. (2008). Levinasian ethics and animal rights. *Windsor Yearbook of Access to Justice, 26*(2), 313-328.

Csibra, G. & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B Biological Sciences, 366*(1567), 1149-57.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y. & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience, 9*(1), 28. https://doi.org/10.1038/nn1611

Davy, Barbara Jane (2007). An other face of ethics in Levinas. *Ethics & The Environment, 12*(1), 39-65

De Haan, S. & Fuchs, T. (2010). The ghost in the machine: Disembodiment in schizophrenia – Two case studies. *Psychopathology, 43*, 327-333.

De Jaegher, H. (2019). Loving and knowing: Reflections for an engaged epistemology. *Phenomenology and the Cognitive Sciences*.

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

De Jaegher, H., Di Paolo, E. & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences, 14*(10), 441-447.

De Jaegher, H. & Froese, T. (2009). On the role of social interaction in individual agency. *Adaptive Behavior, 17*(5), 444-460.

Derrida, J. (1978). Violence and metaphysics: An essay on the thought of Emmanuel Levinas. In A. Bass (Trans.), *Writing and difference*. Chicago: The University of Chicago Press.

DeSouza, N. (2013). Pre-reflective ethical know-how. *Ethical Theory and Moral Practice*, 16, 279-294.

Devereux, G. (1967). *From anxiety to method in the behaviorial sciences*. The Hague: Mouton and Co.

Dierckxsens, G. (2020). Enactive cognition and the other: Enactivism and Levinas meet halfway. *Journal of French and Francophone Philosophy*, XXVIII(1), 100-120.

Doerr-Zegers, O., Irarrázaval, L., Mundt, A. & Palette, V. (2017). Disturbances of embodiment as core phenomena of depression in clinical practice. *Psychopathology, 50*(4).

Donohoe, J. (2004). *Husserl on ethics and intersubjectivity: From static to genetic phenomenology*. Amherst, New York: Humanity Books.

Falck-Ytter, T., Bölte, S. & Gredebäck, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders, 5*(1), 28. https://doi.org/10.1186/1866-1955-5-28

Frith, C.D. (1992). *The cognitive neuropsychology of schizophrenia*. New York, NY: Psychology Press.

Froese, T. & Krueger, J. (2021). Lost in the socially extended mind: genuine intersubjectivity and disturbed self-other demarcation in schizophrenia. In C. Tewes & G. Stanghellini (Eds.), *Time and body: Phenomenological and psychopathological approaches* (pp. 318-340). Cambridge, UK: Cambridge University Press.

Fuchs, T. (2005). Corporealized and disembodied minds: a phenomenological view of the body in melancholia and schizophrenia. *Philosophy, Psychiatry, & Psychology*, 12(2), 95-107.

Fuchs, T. (2013). Depression, intercorporeality, and interaffectivity. *Journal of Consciousness Studies, 20*(7-8), 219-238.

Fuchs, T. (2020). Delusion, reality, and intersubjectivity: a phenomenological and enactive analysis. *Philosophy, Psychiatry, & Psychology*, 27(1), 61-79.

Fuchs, T. & Koch, S. C. (2014). Embodied affectivity: on moving and being moved. *Frontiers in Psychology, 5*, 508.

Fuchs, T. & Schlimme, J. E. (2009). Embodiment and psychopathology: a phenomenological perspective. *Current Opinion in Psychiatry, 22*(6), 570-575.

Gabrovska, V. S., Laws, K. R., Sinclair, J. & McKenna, P. J. (2002). Visual object processing in schizophrenia: evidence for an associative agnostic deficit. *Schizophrenia Research, 59*(2-3), 277-286.

Gallagher, S. (2005). *How the body shapes the mind*. Oxford, UK: Oxford University Press.

Gallagher, S. (2009). Two problems of intersubjectivity. *Journal of Consciousness Studies*, 16(6-7), 289-308.

Gallagher, S. (2012). Interactive coordination in joint attention. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience*. The MIT Press.

Gallagher, S. (2014). In your face: transcendence in embodied interaction. *Frontiers in Human Neuroscience*, 8.

Gallagher, S. (2015). Relations between agency and ownership in the case of schizophrenic thought insertion and delusions of control. *Review of Philosophy and Psychology, 6*, 865-879.

Gallagher, S. (2020). *Action and interaction*. Oxford University Press.

Gallagher, S. & Crisafi, A. (2008). Mental institutions. *Topoi*, 28, 45-51.

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493–501.

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*, 593–609.

Gantt, E. E. (2000). Levinas, psychotherapy, and the ethics of suffering. *Journal of Humanistic Psychology*, 40(3), 9-28.

Gantt, E. E. (2002). Utopia, psychotherapy and the place of suffering. In E. E. Gantt & R. N. Williams (Eds.), *Psychology for the other: Levinas, ethics, and the practice of psychology*. Duquesne University Press.

Gantt, E. E. & Williams, R. N. (2002). *Psychology for the other: Levinas, ethics and the practice of psychology*. Duquesne University Press.

Gergely, G. & Csibra, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148-153.

Gergely, G. & Csibra, G. (2013). Natural pedagogy. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: what infants, children, and other species can teach us*. Oxford University Press.

Gillet, G. & Harré, R. (2013). Discourse and diseases of the psyche. In K. W. M. Fulford, M. Davies, R. G. T. Gipps, G. Graham, J. Z. Sadler, G. Stanghellini & T. Thornton (Eds.), *The Oxford handbook of philosophy and psychiatry*. Oxford, UK: Oxford University Press.

Goldman, A. I. (1995). Interpretation psychologized. In M. Davies & T. Stone (Eds.) *Folk psychology: The theory of mind debate*. Blackwell Publishing.

Goodman, D. M. & Becker, B. W. (2015). Trauma as violent awakening. In K. C. Krycka, G. Kunz & G. G. Sayre (Eds.), *Psychotherapy for the other: Levinas and the face-to-face relationship*. Pittsburgh, Pennsylvania: Duquesne University Press.

Gordon, R. M. (1995a). Folk psychology as simulation. In M. Davies, M. & T. Stone (Eds.), *Folk psychology: The theory of mind debate*. Blackwell Publishing.

Gordon, R. M. (1995b). Simulation without Interpretation or Inference from Me to You. In M. Davies & T. Stone (Eds.), *Folk psychology: the theory of mind debate*. Blackwell Publishing.

Gordon, R. M. (1995c). The Simulation Theory: Objections and Misconceptions. In M. Davies, & T. Stone (Eds.), *Folk psychology: The theory of mind debate*. Blackwell Publishing.

Gordon, R. M. (1996). "Radical" Simulation. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind*. Cambridge University Press.

Hart, S. D. & Hare, R. D. (1998). Association between psychopathy and narcissism: Theoretical views and empirical evidence. In E. F. Ronningstam (Ed.), *Disorders of narcissism: Diagnostic, clinical, and empirical implications*. Washington, DC: American Psychiatric Press.

Heal, J. (1995a). Replication and Functionalism. In M. Davies & T. Stone (Eds.) *Folk psychology: The theory of mind debate*. Blackwell Publishing.

Heal, J. (1995b). How to Think about Thinking. In M. Davies & T. Stone (Eds.) *Folk psychology: The theory of mind debate*. Blackwell Publishing.

Heal, J. (1996). Simulation, Theory and Content. In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind*. Cambridge University Press.

Heal, J. (1998). Co-cognition and off-line simulation: Two ways of understanding the simulation approach'. *Mind & Language, 13*, 477–498.

Heal, J. (2005). Review of Nichols and Stich (2003). Mindreading: An integrated account of pretence, self-awareness and understanding other minds. *Mind, 114*, 181–184.

Hegel, G. W. F. (1977). *Phenomenology of spirit* (A. V. Miller, Trans.). Oxford University Press.

Heidegger, M. (2008). *Being and time*. (J. Macquarrie & E. Robinson, Trans.). Harper Perennial Modern Classics.

Hoerl, C. (2013). Jaspers on explaining and understanding in psychiatry. In G. Stangellini & T. Fuchs (Eds.), *One Century of Karl Jaspers' General Psychopathology* (pp. 107-120). Oxford, UK: Oxford University Press.

Husserl, E. (1859-1938). *Husserliana: gesammelte Werke*. Den Haag: M. Nijhoff.

Husserl, E. (1927-1997). *Phenomenology* (The Encyclopedia Britannica article) (T. Sheehan, Trans.). Kluwer Academic Publishers.

Husserl, E. (1931). *Cartesian meditations*. (D. Cairns, Trans.). The Hague, The Netherlands: Martinus Nijhoff Publishers.

Husserl, E. (1960). *Cartesian meditations: An introduction to phenomenology* (D. Cairns, Trans.). The Hague/Boston/London: Martinus Nijhoff.

Husserl, E. (1970). The origin of geometry. In E. Husserl, *The crisis of European sciences and transcendental phenomenology* (D. Carr, Trans.). Northwestern University Press.

Husserl, E. (1970). *The crisis of European sciences and transcendental phenomenology* (D. Carr, Trans.). Northwestern University Press.

Husserl, E. (1989). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy, second book* (R. Rojcewicz & A. Schuwer, Trans.). Dordrecht: Kluwer Academic Publisher.

Husserl E. (1999). *The idea of phenomenology: A translation of Die Idee Der Phänomenologie Husserliana II* (L. Hardy, Ed. & Trans.). Dordrecht: Kluwer Academic Publishers.

Husserl, E. (2019). *First philosophy: Lectures 1923/24 and related texts from the manuscripts (1920-1925)*. S. Luft & T. M. Naberhaus (Eds.). Springer Netherlands.

Janicaud, D. (1991). *Le tournant théologique de la phénoménologie française.* Combas: Editions de l'Éclat. (English translation in Dominique Janicaud et al., *Phenomenology and the Theological Turn: The French Debate* [New York: Fordham University Press, 2001]).

Jaspers, K. (1920). *Allgemeine Psychopathologie, für Studierende, Ärzte und Psychologen*. Berlin Heidelberg: Springer-Verlag.

Jaspers, K. (1963). *General Psychopathology* (7th ed.). Chicago: University of Chicago Press.

Jaspers, K. (1968). The phenomenological approach in psychopathology. *The British Journal of Psychiatry, 114*(516), 1313-1323.

Jaspers, K. (1997). *General psychopathology, Volumes I & II*. (J. Hoenig & M. W. Hamilton, Trans.). Baltimore, MD, US: Johns Hopkins University Press.

Jones, W. & Klin, A. (2013). Attention to eyes is present but in decline in 2–6-month-old infants later diagnosed with autism. *Nature, 504*(7480), 427–431. https://doi.org/10.1038/nature12715

Jordan, R. W. (1987). On not living in the primordial world. Presented 1987 at Washington University and the University of Notre Dame.

Katz, C. (2010). Education east of Eden: Levinas, the psychopath and the paradox of responsibility. In P. Atterton & M. Calarco (Eds.), *Radicalizing Levinas* (pp. 171-183). Albany, New York: State University of New York Press.

Kernberg, O. F. (2014). An overview of the treatment of severe narcissistic pathology. *The International Journal of Psychoanalysis, 95*(5), 865–888.

Klin, A., Jones, W., Schultz, R., Volkmar, F. & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*(9), 809–816. https://doi.org/10.1001/archpsyc.59.9.809

Koch, S., Kunz, T., Lykou, S. & Cruz, R. (2014). Effects of dance movement therapy and dance on health-related psychological outcomes: A meta-analysis. *The Arts in Psychotherapy, 41*(1), 46-64.

Koster, E. H. W. & Fang, L. (2015). Eye movement tracking. In R. L. Cautin & S. O. Lillienfeld (Eds.), *The encyclopedia of clinical psychology*. Wiley-Blackwell.

Krueger, J. (2010). Extended cognition and the space of social interaction. *Consciousness and Cognition, 20*, 643-657.

Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology, 4*.

Krueger, J. (2018). Direct social perception. In A. Newen, L. de Bruin & S. Gallagher (Eds.), *The Oxford Handbook of 4e Cognition*. Oxford: Oxford University Press.

Krueger, J. (2018). Schizophrenia and the scaffolded self. *Topoi*, 39, 597-609.

Krueger, J. (2020). Watsuji, intentionality, and psychopathology. *Philosophy East and West, 70*(3), 757-780.

Kugelmann, R. (2002). Pain, exposure, and responsibility: Suffering pain and suffering for the other. In E. E. Gantt & R. N. Williams (Eds.), *Psychology for the other: Levinas, ethics, and the practice of psychology*. Duquesne University Press.

Kunz, G. (1998). *The paradox of power and weakness: Levinas and an alternative paradigm for psychology*. Albany, New York: State University of New York.

Kunz, G. (2015). Weak enough. In K. C. Krycka, G. Kunz & G. G. Sayre (Eds.), *Psychotherapy for the other: Levinas and the face-to-face relationship*. Pittsburgh, Pennsylvania: Duquesne University.

Kyselo, M. (2014). The body social: an enactive approach to the self. *Frontiers in Psychology, 12*.

Kyselo, M. (2015). The enactive approach and disorders of the self – the case of schizophrenia. *Phenomenology and the Cognitive Sciences, 15*, 591-616.

Laing, R. D. (1969). *The Divided self: An existential study in sanity and madness*. London, England: Tavistock Publications.

Levinas, E. (1969). *Totality and infinity: An essay on exteriority* (A. Lingus, Trans.). Pittsburgh, Pennsylvania: Duquesne University Press.

Levinas, E. (1981). *Otherwise than being, or beyond essence* (A. Lingus, Trans.). Pittsburgh, Pennsylvania: Duquesne University Press.

Levinas, E. (1985). *Ethics and infinity: Conversations with Philippe Nemo* (R. Cohen, Trans.). Pittsburgh: Duquesne University Press.

Levinas, E. (1989). *Ethics as first philosophy (Éthique comme philosophie première)* (A. T. P eperzak, Ed.). Paris, France: Editions Payot & Rivages. (Original work published 1982.)

Levinas, E. (1996). *Levinas: Basic philosophical writings*. (A. T. Peperzak, S. Critchley & R. Bernasconi, Eds.). Bloomington and Indianapolis: Indiana University Press.

Levinas, E. (1996b). Peace and proximity. In A. T. Peperzak, S. Critchley & R. Bernasconi (Eds.), *Emmanuel Levinas: Basic philosophical writings*. Indiana University Press.

Levinas, E. (1996b). Transcendence and intelligibility. In A. T. Peperzak, S. Critchley & R. Bernasconi (Eds.), *Emmanuel Levinas: Basic philosophical writings*. Indiana University Press.

Lewkowicz, D. J. & Hansen-Tift, A. M. (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(5), 1431-1436.

Loaiza, J. M. (2019). From enactive concern to care in social life: towards an enactive anthropology of caring. *Adaptive Behavior*, 27(1), 17-30.

Loewenthal, D. (2011). *Post-existentialism and the psychological therapies: towards a therapy without foundations.* Routledge.

Lysaker, P. H., Johannesen, J. K. & Lysaker, J. T. (2005). Schizophrenia and the experience of intersubjectivity as threat. *Phenomenology and the Cognitive Sciences*, 4, 335-352.

Maiese, M. (2018). An enactivist approach to treating depression: cultivating online intelligence through dance and music. *Phenomenology and the Cognitive Sciences*, 19, 523-547.

Martin, J. R. & Pacherie, E. (2013). Out of nowhere: Thought insertion, ownership and context-integration. *Consciousness and Cognition*, *22*(1), 111–122. https://doi.org/10.1016/j.concog.2012.11.012

Martin, L. A., Koch, S. C., Hirjak, D. & Fuchs, T. (2016). Overcoming disembodiment: The effect of movement therapy on negative symptoms in schizophrenia – A multicenter randomized controlled trial. *Frontiers in Psychology, 7*(483).

Maturana, H. R. & Varela, F. J. (1973). *De máquinas y seres vivos: Una teoría de la organizacion biológica.* Santiago, Chile: Editorial Universitaria.

Maturana, H. R. & Varela, F. J. (1980). Autopoiesis and cognition: The realization of the living. *Boston Studies in the Philosophy of Science, 42*. Dordrecht: D Reidel.

Métais, F. & Villalobos, M. (2020). Embodied ethics: Levinas' gift for enactivism. *Phenomenology and the Cognitive Sciences, 20*(3), 1-22.

Metzinger, T. (2003) *Being no One: The self-model theory of subjectivity*. Cambridge, MA: MIT Press.

Mikropoulos, T., Delimitros, M., Gaintatzis, P., Iatraki, G., Stergiouli, A., Tsiara, A. & Kalyvioti, K. (2020). Acceptance and user experience of an augmented reality system for the simulation of sensory overload in children with autism. *6th International conference of the immersive learning research network (iLRN)*, 86-92.

Moran, D. & Cohen, J. (2012). *The Husserl dictionary*. London/New York: Continuum International Publishing Group.

Muse, K. R. (1981). Edmund Husserl's impact on Max Weber. *Sociological Inquiry, 51*, 99-104. doi:10.1111/j.1475-682X.1981.tb01032.x

Nezlek, J. B., Hampton, C. P. & Shean, G. D. (2000). Clinical depression and day-to-day social interaction in a community sample. *Journal of Abnormal Psychology, 109*(1), 11–19.

Nezlek, J. B., Imbrie, M. & Shean, G. D. (1994). Depression and everyday social interaction. *Journal of Personality and Social Psychology, 67*(6), 1101–1111.

Nortvedt, P. (2008). Sensibility and clinical understanding. *Medical Health Care and Philosophy, 11*, 209-119.

Overgaard, S. (2003). On Levinas' critique of Husserl. In D. Zahavi, S. Heinämaa, & H. Ruin (Eds.), *Metaphysics, facticity, interpretation: Phenomenology in the Nordic countries* (pp. 115-138). The Netherlands: Kluwer Academic Publishers.

Overgaard, S. (2005). Rethinking other minds: Wittgenstein and Levinas on expression. *Inquiry, 48*(3), 249-274.

Orellana, G. & Slachevsky, A. (2013). Executive functioning in schizophrenia. *Frontiers in Psychology, 4*(35).

Parnas, J. & Sass, L. A. (2001). Self, solipsism, and schizophrenic delusions. *Philosophy, Psychiatry, & Psychology, 8*(2/3), 101-120.

Payne, R. (2012). Night's end. *Schizophrenia Bulletin*, *38*(5), 899–901

Peperzak, A. (1993). *To the other: An introduction to the philosophy of Emmanuel Levinas*. West Lafayette, Indiana: Purdue Research Foundation.

Perpich, D. (2008). *The ethics of Emmanuel Levinas*. Stanford University Press: Stanford, California.

Piaget, J. (1995). *Sociological studies*. Routledge. (Original work published 1977.)

Piredda, G. (2017). The mark of the cognitive and the coupling-constitution fallacy: A defense of the extended mind hypothesis. *Frontiers in Psychology, 28*.

Qin, S., Nagai, Y., Kumagaya, S., Ayaya, S. & Asada, M. (2014). Autism simulator employing augmented reality: A prototype. *4th joint IEEE international conference on development and learning and on epigenetic robotics*, 155-156.

Rakoczy, H., Warneken, F. & Tomasello, M. (2008). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology*, 44(3), 875-881.

Rakoczy, H., Warneken, F. & Tomasello, M. (2009). Young children's selective learning of rule games from reliable and unreliable models. *Cognitive Development, 24*, 61-69.

Ratcliffe, M. (2010). Depression, guilt and emotional depth. *Inquiry, 53*(6), 602-626.

Ratcliffe, M. (2013). The World of Depression. https://www.academia.edu/6054108/The_World_of_Depression

Ratcliffe, M. (2014). The phenomenology of depression and the nature of empathy. *Medicine, Health Care and Philosophy, 17*, 269-280.

Ratcliffe, M. (2015). *Experiences of Depression: A Study in Phenomenology.* Oxford, UK: Oxford University Press.

Ratcliffe, M. & Wilkinson, S. (2015). Thought insertion clarified. *Journal of Consciousness Studies, 22*(11), 246-269.

Riby, D. M. & Hancock, P. J. B. (2009). Do faces capture the attention of individuals with Williams syndrome or autism? evidence from tracking eye movements. *Journal of Autism and Developmental Disorders, 39*(3), 421–431. https://doi.org/10.1007/s10803-008-0641-z

Riemer, M. (2018). Delusions of control in schizophrenia: Resistant to the mind's best trick? *Schizophrenia Research, 197*, 98-103.

Röhricht, F. & Priebe, S. (2006). Effect of body-oriented psychological therapy on negative symptoms in schizophrenia: a randomized controlled trial. *Psychological Medicine*, *36*(5), 669-78.

Rossiter, A. (2011). Unsettled social work: The challenge of Levinas' ethics. *British Journal of Social Work, 41*(5), 980-995.

Rümke, H. C. (1958). Der klinische Differenzierung innerhalb der Gruppe der Schizophrenien. *Nervenarzt*, 29, 40-53.

Rümke, H. C. (1990). The nuclear symptom of schizophrenia and the praecox feeling. *History of Psychiatry, 1*, 331-341.

Sass, L. A. (2001). Self and world in schizophrenia: three classic approaches. *Philosophy, Psychiatry, & Psychology*, 8(4), 251-270.

Sass, L. A. (2014). Delusion and double book-keeping. In T. Fuchs, T. Breyer & C. Mundt (Eds.), *Karl Jaspers' Philosophy and Psychopathology* (pp. 125–147). New York: Springer-Verlag

Sass, L. A. (2019). Three dangers: Phenomenological reflections on the psychotherapy of psychosis. *Psychopathology, 52*, 126-134.

Sass, L. A. (2020). Delusion, reality, and excentricity: Comment on Thomas Fuchs. *Philosophy, Psychiatry, and Psychology, 27*(1), 81-83.

Sayre, G. G. (2015). Toward a therapy for the other. In K. C. Krycka, G. Kunz & G. G. Sayre (Eds.), *Psychotherapy for the other: Levinas and the face-to-face relationship*. Pittsburgh, Pennsylvania: Duquesne University Press.

Schmidt, M. F. H & Rakoczy, H. (2017). Developing an understanding of normativity. In A. Newen, L. de Bruin & S. Gallagher (Eds.), *Oxford handbook of cognition: Embodied, embedded, enactive and extended*. Oxford, UK: Oxford University Press.

Schutz, A. (1970). *Collected papers III: Studies in phenomenological philosophy*. The Hague: Martinus Nijhoff.

Shah, P., Livingston, L.A., Callan, M.J. & Player, L. (2019). Trait autism is a better predictor of empathy than alexithymia. *Journal of Autism and Developmental Disorders*. doi: 10.1007/s10803-019-04080-3

Shimada, S. & Hiraki, K. (2006). Infant's brain response to live and televised action. *Neuroimage, 32*(2), 930-9.

Siles i Borrás, J. (2010). *The ethics of Husserl's phenomenology: Responsibility and ethical life*. London: Continuum.

Silverstein, S. M. (2016). Visual perception disturbances in schizophrenia: A unified model. In M. Li & W. D. Spaulding (Eds.), *Nebraska symposium on motivation: Vol. 63. The neuropsychopathology of schizophrenia: Molecules, brain systems, motivation, and cognition* (pp. 77–132). Springer International Publishing AG.

Silverstein, S., Demmin, D. & Škodlar, B. (2017). Space and objects: On the phenomenology and cognitive neuroscience of anomalous perception in schizophrenia (ancillary article to EAWE domain 1). *Psychopathology, 50*(1).

Simmons, W. P. (1999). The third: Levinas' theoretical move from an-archical ethics to the realm of justice and politics. *Philosophy and Social Criticism, 25*(6), 83-104.

Slaby, J. (2014). Emotions and the extended mind. In C. von Scheve & M. Salmela (Eds.), *Collective emotions*. Oxford: Oxford University Press.

Smith, J. (2010). Seeing other people. *Philosophy and Phenomenological Research*, 81(3), 731-48.

Stanghellini, G. (2000). Vulnerability to schizophrenia and lack of common sense. *Schizophrenia Bulletin, 26*(4), 775-787.

Stanghellini, G. (2001). Psychopathology of common sense. *Philosophy, Psychiatry, & Psychology, 8*(2/3), 201-218.

Stanghellini, G. (2004). *Disembodied spirits and deanimated bodies: The psychopathology of common sense*. Oxford Medical Publications.

Stanghellini, G. (2009). Embodiment and schizophrenia. *World Psychiatry, 8*(1), 56-9.

Stanghellini, G. (2019). The PHD method for psychotherapy: Integrating phenomenology, hermeneutics, and psychodynamics. *Psychopathology, 52*(2), 75-84.

Stein, E. (1989). *On the problem of empathy* (W. Stein, Trans.). Washington, D.C.: ICS Publications.

Sterling, L., Dawson, G., Webb, S., Murias, M., Munson, J., Panagiotides, H. & Aylward, E. (2008). The role of face familiarity in eye tracking of faces by individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(9), 1666–1675. https://doi.org/10.1007/s10803-008-0550-1

Stone, M . (1993). *Abnormalities of personality: Within and beyond the realm of treatment*. New York: Norton Professional Books.

Takarae, Y., Minshew, N. J., Luna, B., Krisky, C. M. & Sweeney, J. A. (2004). Pursuit eye movement deficits in autism. *Brain, 127*(12), 2584–2594. https://doi.org/10.1093/brain/awh307

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind.* Cambridge, Massachusetts: Harvard University Press.

Trevarthan, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before speech: The beginning of human communication* (pp. 321-347). Cambridge: Cambridge University Press.

Trevarthan, C. & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding, and acts of meaning in the first year. In J. Lock (Ed.), *Action, gesture and symbol* (pp. 183-229). London: Academic Press.

Tsang, N. M. (2017). Otherness and empathy – implications of Levinas' ethics for social work education. *Social Work Eductation, 36*(3), 312-322.

Tsang, T., Atagi, N. & Johnson, S. P. (2018). Mouth-looking supports expressive language skills in monolingual and bilingual infants. *Journal of Experimental Child Psychology, 169*, 93-109.

Urban, P. (2014). Toward an expansion of an enactive ethics with the help of care ethics. *Frontiers in Psychology, 5*.

Urban, P. (2015). Enactivism and care ethics: Merging perspectives. *Filozofia*, 70, 119-129.

Varela, F. (1999). *Ethical know-how: Action, wisdom and cognition* (L. Gius & R.B. Figli, Trans.). Stanford, California: Stanford University Press.

Varga, S. & Krueger, J. (2013). Background emotions, proximity and disturbed emotion regulation. *Review of Philosophy and Psychology*, *4*, 271-292.

Wheelwright, P. (1959). *Heraclitus*. Princeton, New Jersey: Princeton University Press.

Widiger, T. A. (2006). Psychopathy and DSM-IV psychopathology. In C. J. Patrick (Ed.), *Handbook of psychopathy* (p. 156–171). The Guilford Press.

Wiggins, O.P. & Schwartz, M. (2013). Phenomenology and psychopathology: in search of a method. In G. Stangellini & T. Fuchs (Eds.), *One Century of Karl Jaspers' General Psychopathology* (16-26). Oxford, UK: Oxford University Press.

Williams, R. N. (2015). Levinas and psychoanalysis. In K. C. Krycka, G. Kunz & G. G. Sayre (Eds.), *Psychotherapy for the other: Levinas and the face-to-face relationship*. Pittsburgh, Pennsylvania: Duquesne University Press.

Wittgenstein, L. (2009). *Philosophical investigations*. (G.E.M Anscombe, P.M.S. Hacker & J. Schulte, Trans.). Blackwell Publishing Ltd.

Zahavi, D. (2005). Being Someone. *Psyche, 11*(5), 1-20.

Zahavi, D. (2010). Empathy, embodiment and interpersonal understanding: From Lipps to Schutz. *Inquiry, 53*(3), 285-306.

Zahavi, D. (2014). *Self and other: Exploring subjectivity, empathy, and shame*. Oxford University Press.

Zajenkowski, M., Maciantowicz, O., Szymaniak, K. & Urban, P. (2018). Vulnerable and grandiose narcissism are differentially associated with ability and trait emotional intelligence. *Frontiers in Psychology, 9*.