Altered Self and Altered Self Experience

Book · August 2014

2 authors, including:

Some of the authors of this publication are also working on these related projects:

Image in Science and Art View project

Cognitive Foundation of the Self View project

Altered Self and Altered Self-Experience
Alexander Gerner & Jorge Gonçalves (eds)

ASASE

Altered Self and Altered Self-Experience

.

Alexander Gerner & Jorge Gonçalves (eds)

Contents

II
Altered Selves

Eu sou uma anthologia

I am an anthology

.

[61A-27ʳ]

Fernando Pessoa 17-12-1932

Introduction

From a theoretical point of view the experiences in which the feeling of self is temporarily or permanently altered, pose opportunities to apply, critically verify or even renew theories of the self. The altered self, from a human point of view, supposedly gains its most relevance when it is related to psychological sufferings, their comprehension, relief or treatment. However, not all altered self-experiences imply suffering or respective treatment. In eastern (and some western) philosophical traditions a „No-Self" is considered. Techniques of a Self/ No- self include states of meditation, mindfulness or effortless attention. Moreover, for some, self-alterations are desired as in technically achieved, or drug induced suspensions of a "fixed", „regular", "normalized" or "orthodox" experience of oneself. If those normative approaches to the self in psychopathologic description actually help to elucidate the self or just risk affirming the concept of self-pathology instead of seeing the self as a vulnerable mode of self-preservation in its varieties of expressions, has to be considered. Different conceptual and clinical notions of the altered self and different modes of altered self-experience will be explored as for instance the general question of what a mental disorder could be as conceptual description of the altered self and how these different approaches to altered selves and altered self experiences help to clarify the notion of self.

This book is the result of a selection of research papers of the project "Cognitive Foundation of the Self"[1], with the generous support of the Portuguese foundation FCT[2] and the IFILNOVA and with contributions of international scholars who mainly presented and discussed their work at the international conference called *TheAltered Self and Altered Self-Experience*[3] held at the Universidade Nova de Lisboa (FCSH-UNL) at the institute of philosophy IFILNOVA on the 30th and 31st of May 2013, and with a special contribution of one of the consultants of this project, Georg Northoff.

If the self is a necessary condition for the possible constitution of experience and consciousness, how then does an altered self influence experience or consciousness? Is the self actually a stable or a mixed stable/changeable experience condition? How many different states of the self do we experience during our life? If we take the famous Portuguese writer Fernando Pessoa (1888-1935)- "I am an anthology"- and his 136 fictious characters (Pessoa 2013) under which fictious names he wrote, serious, we could even ask: *How many (altered) selves are we?* Does each one of us have only and exclusively one permanent and constantly identical (minimal) self or

do we have to deal with experiential variations, functional and formal development and dynamic modifications and change of self experiences, in which for instance the existential feeling of self is altered or in which we can take perspective of the other, instead of holding on to a clear unity and continuity of being a unique and unchangeable self over time? What affects our self temporarily and what changes it permanently? Do we have to struggle on the phenomenal level with a disunity of a multiplicity of phenomenal "aspects" and therefore as well with a multitude with epistemic conceptions of the self? Is the self "simply" a "complex" but problematic pattern (Gallagher 2013)? Can a self be altered voluntarily -and thus, can we experience altered states of selves? What do permanently altered selves tell us about the problem of normativity in relation to designating selves as "ill" or "disordered"? Is there an instituted "altered" self by the knowledge discipline of psychiatry? And can we speak of altered selves for example in self-impairments and disruption of the self as described in *schizophrenia* or *autism*?

This book deals with the above mentioned questions from a decidedly interdisciplinary point of view including multiple and foremost divergent conceptual and empirically informed perspectives from philosophy, developmental psychology, mindfulness praxis, and different approaches to "neuro"-philosophies, philosophy of psychiatry, psychoanalysis, phenomenology as well as religious studies, cinema and literature studies.

The first part of ASASE is on Altered Self-Experience and the second part deals with the problems of altered self, specifically in relation to schizophrenia and autism.

I
Altered Self-Experience

Emotional Aspects of Altered Self-Experience

Recently in relation to neurobiological theories of the self Catarine Malabou (2013) outlined in her "Go Wonder: Subjectivity and affects in neurobiological times" in chapter 7 "On Neural Plasticity, trauma and the loss of affects" two meanings of brain *plasticity*, a positive one "in the formation process of neural connections and the fact that these connections may be transformed or altered during our lifetimes under the influence of experience and the kind of life we are leading" (Malabou 2013, 56), while a second mode of brain plasticity lies in a negative plastic power as in brain damage and self disorders. What *altered affects* for our Self mean, and the possibility of alteration of self-experience is given in "existential feelings of being"(Rattcliffe 2008) in which certain affect disappearance in the neurological realm make part. One could even argue that it makes part of the self to be a temporal

operator of appearance and disappearance of affect and emotions. Even so that everything we used to be has drift away into infinity by time, forgetting, pain, the Self might only exists in a permanent auto-apparition or a resistance of disappearance into mere episodic occurrences of certain affects or emotions, therefore regaining a sometimes lost self-relation. How is the self then altered by basic or complex emotional patterns such as anger or guilt?

1 Dina Mendonça - Emotions and Altered Self
(New University of Lisbon, Portugal)

Dina Mendonça´s paper examines how emotions alter the self in various ways and explores the importance of recognizing different types of deepness of emotions in order to establish different ways in which emotions alter the self. Besides looking at the double notion of emotion as episode and as disposition (Goldie 2007), Mendonça concentrates on the change and the mastering of the emotional control: What does it mean to master *anger*: how to avoid becoming angry in the first place, how to cease being angry, how to deal with anger in others (either in episodic occurrences as well as dispositional anger). She develops the hypothesis of "*Being angry as a sonar device"* and that from all emotions anger is the one best suited for this to test "boundaries" of the self's social world and of the self's ability and power to change the social world. Thus anger is seen as a socially important emotion and as an dynamic test-ground of the self in testing and reestablishing boundaries of the self on its own right and in relation to others.

2 Amber Griffioen - Regaining the "Lost Self": A Philosophical Analysis of Survivor's Guilt
(University of Konstanz, Germany)

In "Regaining the "Lost Self": A Philosophical Analysis of Survivor's Guilt" Amber L. Griffioen analyzes the feeling of guilt that survivors of traumatic events (war, rape etc.) feel. Traditionally this guilt is interpreted as resulting from a sense of guilt (identification with the causing agent of the trauma) or out of a sense of shame. The author believes that we should not focus on any of them separately but rather jointly. She characterizes these two kinds of guilt feelings concluding that the guilt lies more connected to the action while the focus of shame is the "self". She also analyzes the known varieties of guilt: guilt by substitution (feeling that their survival involved the death of another), guilt for transgression (for having committed certain act or being to blame for a certain omission of acting right) and guilt for its involving role (feeling that the survivor somehow caused the trauma). What Griffioen puts into perspective is a "self" that feels hurt in a) its *fundamental values*, b) in *its autonomy* as well as c) in *its integrity*. This self thus feels ashamed to

continue to exist in a world that is no longer what it had been before the trauma. The author's assumption then is that the *feeling of guilt rehabilitates the agency of the self*, restoring its independence and regaining its integrity. If the subject thinks it had some guilt in the trauma, then is becomes an agent. In this sense Griffeen's thesis is crystal clear: it is preferable to be an agent that is guilty, a self to be blamed, than not being an agent at all, and thus being lost or having lost oneself.

Actor's Self

3 Sara F. Bizzaro - Acting and the Self
(Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal)

With Douglas Hoffstaedter (2007) and his book "I am a strange loop" Bizzaro conceives the self in self rehearsals as a strange loop explaining how different types of acting and their respective theories work in changing the actor's self for a performance. In order to understand how an actor forms the character's personality, Bizzaro analyses the acting techniques proposed by Strasberg, Stanislavski, Adler and Meisner. These authors' present techniques of acting that go from a heavily *internal emotion based acting* to an *external action based acting.* Bizzaro argues that the efficacy of all these acting methods can be understood within the framework created by Hofstadter to explain the way "selves" are represented in the brain. Analyzing how actors create other "selves" can be useful for our understanding of what the "self" is. Our abilities to play-act as children, to full fledge act as professional actors, or socially act in fulfilling our different social roles and our ability to view others as "selves", all seem to indicate that the "self" is indeed less firm and stable than we like to think. Healthy "self" deviations can be just normal different uses of the same "self" mechanism and in the sense of Bizarro Hofstadter 's theory is adequate to easily explain how different types of "acting" work.

Gender and Altered Self Experience

4 Iwona Janicka - Cutting the Cords - Rethinking Judith Butler's Position on Transgender with René Girard's Theory of Mimesis
(University of Cambridge, UK)

Situated aspects play some role in shaping who we are. They include the kind of family structure and environment where we grew up; cultural and normative practices and narratives define our way of living. The constitution of a wide or complex self is no exception in this. An extended self includes self-concepts and self-narratives as well as gender accounts to be looked at when referring to an *extended autobiographical self* in which self-image and identity and thus the rational

intentional make-up of selves is in question. This chapter puts the extended self theoretically shaped by psychoanalytical gender theory to test. Psychoanalytical theories of the self often make use of cultural founding narratives to explain how selves are altered and may deviate towards "strange chambers of our soul". Freud's *Oedipus* is such an cultural founding narrative of the self and Butlers approach towards the Oedipus myth and the desiring self in Girard (1986) -in a dialogue with two psychiatrists (Jean-Michel Oughourlian and Guy Lefort)- in his *mimesis* account in "Things hidden since the foundation of the world", explores the self as other. Janicka theorizes about gender identity in subjective alteration. For the author the alteration is relative to social normality, traditionally considered. Starting from the theory of Judith Butler (1990, 2004), she tries to show the limits of this theory that are visible, when it comes to interpreting *transgender*. The limits of Judith Butler, according to the author- result from Butler's attachment to a too narrow psychoanalytic explanatory model. This critically viewed model gives an account of heterosexual and homosexual desire but not- according to Janicka- to transgender desire. The assumptions of Butler's model that are due to a *male/female dichotomy* in which no other modalities are possible is based on the assumption that desire is primary and that gender identity is formed mainly from sexual desire and that consequently in this dual proposal of gender is seated in the prohibition of homosexuality, which becomes necessary for the formation of two orthodox poles of sexual identity (male/female). Janicka proposes the model of René Girard as an alternative instead. Girard allows to understand transgender desire because it does not necessarily link sexual desire to gender identity. Girard's starting point is not sexual desire, but imitation: There are models that are to imitate and other to not imitate, i.e., they are rivals. By imitating a model, sexuality is also imitated, for example, the boy imitating his father desiring his mother as his father does. Janicka in the following analyzes several possible combinations of gender and sexuality, trying to demonstrate that Girard's model solves the difficulties of psychoanalytic model best.

Disembodiment of Self-experience: Out-of-Body Experience, Full-Body Illusion and Cinematic Experience

When reflecting on the what the altered state of self in Out-of Body experience can actually tell us about the self we should distinguish different fields of research that should be considered in relation to a philosophical notion of the self: Important in researching the global aspects of the self when they are experienced as altered are: (a) *embodiment* and the bodily foundations of self-consciousness (b) perspective change and *perspective-taking* (2nd, 1st or 3rd PP) in altered selves and altered self experiences: biological and ecological aspects (Fuchs 2012) are important which allow the self to distinguish between self and what is not-self or other, including

aspects that define the *egocentric (body-centered) spatial frame of reference-oposed to allocentric reference frames*, considering a first-person perspective for possible actions in peri-personal space and its altered conditions as for instance in autoscopic experiences such as Out-of-Body experiences (Brugger, Blanke; Lenggenhagger etc.) (see: Gerner & Guerra, Zippel, Fonseca in this volume). Perspective taking can as well be seen as an essential component in the mechanisms that accounts for intersubjectivity and agency in relation to the self.

According to Mohr&Blanke (2005) autoscopic phenomena (AP) are "rare", illusory visual experiences during which the subject e.g. has the impression of seeing a second own body in extra-personal space. AP - in their view- consist in "out-of-body experience, autoscopic hallucination, and heautoscopy". For Brugger et al. (2006) Heautoscopy can be described as the encounter with one's double (the reduplication of a single body and self and thus a breakdown of integrative multimodal processes that let me identify with my body), in the sense of a multimodal illusory reduplication of one's own body and self. The phenomenon of polyopic Heautoscopy (a multiplication of body and self) according to Brugger et al (2006) "points to the multiple mappings of the body, whose disintegration may give rise to the illusory experience of multiple selves." Moreover, recent research has not only described phenomenologically these strange doubling, mirroring or shadowing phenomena of an altered or "disrupted" self (Mishara 2010) but has shown as well that invasive technical manipulation of the brain can even induce an "illusory shadow person" (Arzy et al 2006) for example by artificial brain stimulation of the temporal parietal junction. How these strange autoscopic phenomena could be made fruitful in understanding not only pathologic or deviated self-experience but as well the working of the self in for example cinematic experience shall interest us here. How can the REM sleep state show itself as a proper state of conscious experience in which OBE could find a foundational explanation. Is it then possible to say that we have at least three different conscious states of self (unconscious, REM conscious and full conscious)? Is the self explained well by Northoff 2013b inside a triadic concept of consciousness that takes into consideration the brain´s intrinsic spatio-temporal activities in "resting states", either slow loops (Northoff 2013b) in which the resting state's slow wave fluctuations supposedly in the frequency ranges between 0.001 and 4 Hz are seen as central in consciousness (He et al 2008; He and Raichle, 2009; Raichle, 2009; Northoff 2013b), or in the distinction of resting states in REM sleep to fully conscious states (40Hz), in which the neural reactivity of the resting state to external stimuli distinguishes the awake state from the REM state (Llinás et al. 1998 and Llinás 2002). Accordingly, we could propose three alternating states of the self: the "unconscious" self(?), the properly REM self and the fully conscious self. OBE´s

could be explained to fall into an *in- between territory of overlapping fully awake Self and an REM self*.

5 Alexander Gerner & Michele Guerra - The Cinematic Self- Cinematic Experience, perspective taking and Out-of Body-Experience
(CFCUL, University of Lisbon, Portugal & University of Parma, Italy)

Theoretically and practically the self is something every one of us experiences and as such for Gerner & Guerra is subjected by alterations, mutations, and developments and in some cases even transformations. This chapter asks how can the experience of the self be altered temporarily (Altered Self-Experience) by immersion in cinematic experience including cinema, but as well VR. Gerner & Guerra focus on why the perspective (Petkova et al 2011) and its alteration matters as changes in the 1st PP realized by a) alienation (OBE) or b) appropriation (avatar identification) (see: Ganesh et al 2011) as well as the switch in between 2nd , 1st and 3rd PP in relation to the full body image and full body hallucinations, in which issues of self-identification and *bi-localization* and other doubling phenomena become important to understand altered self-experience. The authors ask: Why OBE' s could be interesting to be treated as cinematic experience in itself in the sense of "cinematic experience without a screen"? They recall as well more recent phenomenological theories of cinematic subjectivity in order to wonder how and whether the movie can be endowed with subjectivity, at least a simulated or enacted form of intersubjectivity assured by the degree of immersion and self-loss of the viewer into the movie and its "extended empathy" (Fuchs 2014) towards the characters represented on the screen (without their real bodies being present), looking at the definition of empathy towards the "virtual other" in Fuchs (2014). Altered self-experiences seem to show a switch of a) The ego-centered and the b) allo-centered reference frame or simply a possible dynamics of different "centeredness" perspectives of (somatic, exosomatic or I-You somatic) self experience. As cinematic experience the authors define switches in between the body centeredness to world-centeredness in relation to what happens on the screen in the moment of suspension of disbelief. But also the extension of the body image is made possible in cinematic experience as well as *empathetic perspective taking* with certain characters or actions on the screen. The question "What do I experience as my body?" (see: Blanke and Metzinger 2009) leads us as well to the problem of the proper *body as a screen* between sensed and projected body of (identity, body image etc) unfolding layers and repetitive projections of a film that is my "own" or my "entire" or my "narrative" body, besides or superimposed on my somatic body schemata. The authors follow the question in relation to immersive experience: *How can we be distracted from our somatic body and immersed into a fictional or cinematic body- the body I feel and am affected by as long as the cinematic experience lasts?*

6 Niccola Zippel - The Altered Self in Dreaming State related to the Out-Of-Body Experiences. A Phenomenological Account interfacing with Neuroscience
(Sapienza University of Rome, Italy)

Zippel presents a *neurophenomenological* account of Out-of-Body experience related to sleep and dreaming states of consciousness. The author proposes a phenomenological reading of the alteration of the self in the out-of-body experiences during dreaming, which, based on the notion of the self as an embodied self, understands the self and the body as an indissoluble whole. The dream state represents a very peculiar condition of bodily experience: while during the NREM sleep one experiences a muscular-skeletal relaxation, in the REM sleep one lives the paralysis of her own body, to which a vivid mental process triggered by the ponto-geniculo-occipital (PGO) waves corresponds. During this phase of the sleep, dreaming takes the shape of a bizarre narration that, despite its incongruity, uncertainty, and discontinuity, inserts itself in the network of neural connections already in place, and for this reason it is another way of performing personal emotions, thoughts, and associations. Precisely by the inability to move, the embodied self experiences such a rich mental performance as an alienating dimension of his/herself, since the body becomes an inactive though often invoked part of the dream (e.g. one tries to escape or to jump in dream without succeeding in doing it). On the other hand, being the motionless counterpart of the frantic activity of the dreaming mind, the body, as it were, is outside the dream events. This complex neurophysiologic situation is interesting from a phenomenological perspective, for it offers the opportunity to consider the condition of the embodied subject as regards the unaware mental process of dreaming. As a general goal, the paper aims at showing the basic features of the relationship between the dreaming self and his/her own body in the light of the most significant empirical dream's researches findings.

But still we have to understand what notion of the self we are talking about: For example Blanke and Metzinger propose a *representationalist* account of *Minimal Phenomenal Self* in which a minimal phenomenal self would be a necessary condition for a strong and cognitive 1PP (see: Blanke & Metzinger 2009, 7) In this part the dynamics of **self-location** (*from where from do I perceive, speak, feel)* become important as well as the question how the **relation between my experiential point of view of my self and the point of view of the somatic body** expresses a dynamicity of the natural media of perspective and perspective-taking with my body (eye position, attentional gaze, attentional mode –focused; distracted; etc.), arms, legs, trunk, head and their positionality: *From which vantage point do I experience the world?*

7 João Fonseca – Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self: some Functional, Computational and Neurological Considerations
(New University of Lisbon, Portugal)

In his *neuro-computational* approach João Fonseca uses autoscopic phenomena such as Full-Body-Illusions (FBIs) and Out-of-Body-Experiences (OBEs) to test his thesis for a taxonomical distinction between two kinds of Core-Self: Core-Self *Simpliciter* (CSS) and Cognitive Core-Self (CCS). He draws differences between a cortical proposal (Motor, Pre-Motor and Prefrontal areas) of a Core-Self and Björn Merker's (2005, 2007) sub-cortical proposal **(**Upper-brainstem complex: *superior culliculus – Pariequeductal Grey Area/hypothalamus - Substancia Nigra/Basal Ganglia*) to develop the notion of Core-Self in this neurophilosophical model would correspond contrastively to the notion of a narrative or auto-biographical Self in the sense of Damasio. Fonseca suggests the fragmentation of the notion of Core-Self into two sub-concepts: CCS and CSS which would correspond to the most "basic sense of selfhood", non-extended in time whereas CCS adds cognitive complexity and a short-time dimension to the former. Fonseca supports his distinction by invoking distinct neural and computational properties to each notion in which CCS would be build up in CSS's neural structures and computational capacities.

Altered Self-Experience in Religious Experience and Intimacy

8 Gabriel Levy - The Implications of Anomalous Monism for Intimate Selves
(Norwegian University of Science and Technology Trondheim, Norway)

The religious study scholar Gabriel Levy looks at the intimate relation of man and woman as in the narrative model of Adam and Eve and brings this religious topic together with Donald Davidson's concept of anomalous monism in interdependence of self and other (Davidson 1970; 2001). Levy's paper focuses on the relationship between *intimacy and the self*. By using one of the paradigmatic narratives about self-altering encounters of (sexual) intimacy, Adam and Eve, Levy explores the concept of intimacy in relation to intimate selves. Davidson's *anomalous monism* argues that, although the first-person has a certain kind of authority in ascribing mental states to him or herself, the concept of the self is irreducible because it is grounded in intersubjectivity. Since intimacy requires a sense of ownership of the subjective self, this can lead to problems within a Davidsonian framework, in which the inter-subjective level is often prioritized. Taking the subjective level seriously from a Davidsonian perspective and from enacted mind (Kyselo & Di Paolo 2013) requires- according to Levy- a *radical subjectivity* that is exemplified by a personal experience of altered intimate self of the author that

happened at the proper Lisbon ASASE meeting in 2013 and that he reflects upon in an narrated interlude. In this way, this chapter pursues the subject of intimate selves not just from a scientific objective but as well from a subjective and intersubjective level in which the study methodologically switches perspective from a historical case study to self-observation and back.

9 Bernardo Palmerim - Attention and Redemption of the Self in Weil and Sartre
(University of Lisbon, Portugal)

In his paper Bernardo Palmeirim addresses a classical issue of religious "conversion" (cf. greek *metanoia*) with Heidegger– which he names within the religious tradition with the term as "attention". Conversion assumes a suffering, sick, divided self, and its supposed (re-) unification in religious or transcendental experience. The author compares two types of conversion. The first is given in the character Antoine Roquentin from the novel *Nausea* by Jean Paul Sartre and that can be seen as a model of "aesthetic conversion." The second is the "religious conversion" in Simone Weil. They have in common that the self in both conversion models turns its attention from itself to something outside by "loosing" its proper self in a religious or sublime "other" by contemplating "God" or an "artwork". For Palmerim in religious conversion we are confronted with a moral transformation in which the world takes on a new meaning in God, while in aesthetic conversion there is no universal structure that explains its existence.

Altered Self-Experience in Drug-Experiments and Mindfulness Meditation

10 Michaela Hulstyn - The Mutative Self and the Language of Insight in Henri Micheaux's Mescaline Experiments
(University of Stanford, USA)

Hulstyn examines the concept of a "mutative Self" in the texts of the Belgian-born poet Henri Micheaux that account for his Mescaline self-experiments in the 1950's and 60's.
She asks: *Can narratives on drug experimentation shed light on philosophic and psychological accounts of the self?* The questionable status of the self during altered states of consciousness includes the moments of individual's existence during which he oneself is not "in its right mind," including the time one spends asleep, drunk, or on the drug mescaline, as in the case of Henri Michaux. While terms like "mind-altering" are commonplace in discussions of drug culture, especially as related to the psychedelic 1960s, few studies of altered states have considered a connection with philosophical discussions of the self, and what self-

experimentation can tell us about the nature of the self during ordinary circumstances. In this paper Hulstyn examines Michaux's take on the process of >unselfing< - a concept developed by Joshua Landy (2012) in *How to Do Things with Fictions* that covers both *self-transcendence* and *self-rupture*, as described during the self-induced drug experiments that began in 1956 and were carried out through the 1960s. Michaux's conception of the self during so-called normal states is considered using terminology from current research in psychology, philosophy and neuroscience to help understand the self-experience that Michaux describes in his work. Michaux's take on altered states of selfhood in a synchronic sense, as described by himself in the experience of disunity in the moment, as well as the effect of the long-term experience, or diachronic sense of the self is analyzed. While doing so Hulstyn examines the implications of Michaux's work on the development of a metaphysical vocabulary for otherworldly experience in the absence of faith in an omnipotent third party witness.

11 Vera Pereira - Mindfulness and Self-Experience: No-Self Consciousness or Empty-Self Consciousness?
(New University of Lisbon, Portugal)

Mindfulness training has been used for centuries in eastern traditions such as Buddism (Albahari 2011) and Yoga with the goal of reaching altered states of consciousness capable of changing the experience and perception of self for spiritual purposes for instance by open monitoring techniques and focused attention (Travis & Shear 2010). Recently studies have shown that even traditional *focused attention* techniques also become effortless with practice (Lutz et al 2008). Thus in western cultures in recent years mindfulness psychotherapeutic programs have become popular. The developmental psychologist Vera Pereira explores in her chapter phenomenological and neurophysiological data on effects of these techniques. She reviews within the field of neuroscientific studies *how mindfulness meditation affects self-referential processing and how meditation influences brain plasticity*. She proposes that continuous regular practice of meditation techniques can lead to the progressive de-identification with the *narrative self* and an identification with the *minimal-self*, that is experienced as "empty" or a "non-personal psychological entity", thus proposing a certain suspension of the narrative (Schechtman 2013) or egoic (Husserl) as well as extended autobiographic self (Damasio 2010) approaches by means of the technique of mindfulness meditation.

II
Altered Selves

The second part of ASASE is on Altered self in states of permanent self alteration in general and schizophrenia, depression and autism in specific. First we will ask with Gonçalves (12) how we can actually define a mental disorder. Problems then addressed are (13) embodied aspects of the self and the self-presenting character of abnormal conscious experience with López-Silva (14) the lack of unity of the self and self-recognition (Pio Abreu) (15) the neurological "midlinestructures" explanatory hypothesis of altered selves (Northoff) related to depression and schizophrenia (16) the interaction of self and other (autism and interaction theories of Cognition(Hipólito) and the enactive approach to self-other while asking (17) Is autism a relational self disorder? (Ciaunica)

What is a Mental Illness?

12 Jorge Gonçalves - How to define a Mental Disorder?
(New University of Lisbon, Portugal)

Gonçalves wonders why certain changes of the self are classified as mental illness. He is critically reviewing, not the search for causal explanations of mental illnesses, but the mere fact of *epistemically knowing how to define these altered states of self as diseases* and according to this definition how disorders come to exist objectively. His idea is that mental illness is defined based on a background of social, ethical and epistemic norms (Broome & Bortolloti, 2009). We may consider them as objective only if we assume that there is progress in human reason in the sense of overcoming cultural relativism in relation to mental illness. Gonçalves responds to possible alternatives to the idea that mental illness is defined as based on social, ethical and epistemic norms. Gonçalves refers to George Graham's (2010) conventionalist thesis that causal-mechanist factors will be always present if there is a mental disorder. Graham argues that if it is true that mental disorders depend on social values, this is not a sufficient reason for us to reject them as "illnesses" in the realistic sense of the term. An alternative is that mental illness would be a malfunction of a natural function. The prevailing thought then would be that it is possible to determine certain natural functions of the brain and injuries that impair these functions. A further alternative is considered by Gonçalves to define mental illness as the subjective experience of suffering. The author, however, tries to show that these last two alternative perspectives are insufficient to define mental illness suggesting the aforementioned solution for reconciling the context dependency and the need for objectivity in Psychiatry.

Altered Self in Schizophrenia and Depression

13 Pablo López-Silva - Self Awareness and the Self-Presenting Character of Abnormal Conscious Experience
(University of Manchester, UK)

From a first-person perspective a self/non-self distinction is given in its various sensory-motor modalities, contributing to an embodied feelings of being (Ratcliffe 2008) a self (a) *sense of ownership* and a (b) *sense of agency* for one's actions (Gallagher 2000, 2012; Rochat 2011). For Zahavi there is a "first personal givenness"(Zahavi 2003) of experience in the sense of Nagel´s (1974) famous "what does it feel like" (to be a bat) to be a self or what it feels like. The feeling of being present and if a self can be directly given or presented, and how this idea of self "presence" is conceived is another field of controversy (c) "*identification*" with a body and its actions as a whole (phenomenally experienced as belonging to me thus also called 'mineness').

The paper of López-Silva critically reviews phenomenological views- especially of Zahavi- that assume that minimal forms of self-awareness are an integral element inherent in all experiences. What is often called the sense of "mineness" of experience or sometimes is refered to in the debate on the "Immunity of Error to Misidentification" - IEM (Shoemaker 1984) - relative to the first person pronoun is called the *self-presenting claim (SPC)* of all experiences in Lopés-Silva's paper. The author uses arguments from a disrupted sense of mineness, such as in schizophrenia, that is specifically cases of "Thought Insertion" but as well reflects cases of somatoparaphrenic delusion and alien hand syndrome as arguments against Zahavi´s notion of mineness and the "I" as the subject of self-awareness in the sense of a necessary feature of givenness of all experience.

Here the question arises how for example Zahavi explains (or fails this task) how the conceptual distinction between sensed and sensing grounds the distinction between the sense of agency and mineness. Altered self experiences such as autoscopic experiences, mindfulness or buddist no-self models pose general questions to a coherent concept of the self and deeply question *the coherence of the body as one global minimal bodily self*.

14 José Luís Pio Abreu - Schizophrenia is the Collapse of Self-Recognition. What Self?
(CFCUL, Portugal)

The professor of psychiatry Pio Abreu (Abreu 2012) addresses questions of the self in schizophrenia from a standpoint of an *intentional self* and the problem of *self-recognition*. After introducing an account of neurobiology of schizophrenia, Abreu shows difficulties that patients with schizophrenia are confronted with as having problems on the neuronal level with brain connectivity by way of "long fibers" which link distant zones of the *neo-telencephalon* (Fornito et al., 2011). These difficulties for Abreu are due to several genetic variations and early environmental aggressions that disturb the development and organization of long fibers in the telencephalon, which are disrupted in pre-schizophrenic and schizophrenic subjects. He asks why we need such a large telencephalon. His thesis is that we need it (1) to handle signs – words, drawings and manufactured objects - which permits us to extend our territory, since signs are essential for all cognitive activities; (2) to recognize other people and ourselves, since human territory is composed of other people, (3) to evaluate reality while shifting between our intentions and the intentions of others. Schizophrenic patients have specific difficulties in every one of these three domains. Abreu defines the main problem of the schizophrenic self as the alternating between its and others' intentions, and assuming alternative meanings. All through life, there are situations where self-recognition and recognition of self-agency is decisive but laborious. These -for Abreu- are the situations that might precipitate a schizophrenic episode.

The *lack of unity of the self in schizophrenia* is expressed by Abreu in the following way: In psychopathology, we usually say that schizophrenic symptoms disrupt the unity of the self as opposed to self identity which is disturbed in dissociative pathologies and is concerned with the temporal continuity of the autobiographical self. In his point of view we should look for the self, which is disturbed in schizophrenia in upper areas of the brain. His take on the intentional self is shown by the example of why we can't tickle ourselves: if I have the intention of tickling myself, I do not feel ticklish because the efferent copy inhibits my sensorial perception. In this case, normal people feel ticklish, but schizophrenic patients have some difficulty in feeling the difference, perhaps because they cannot organize the efferent copies through long fibers in the telencephalon very well. In Pio Abreu's concept of the unity of self-, which is challenged when schizophrenic symptoms are experienced- we have to consider an upper *intentional self*, constructed by learning, and processed in the more recent areas of the brain, above the autobiographical (or narrative) self, the nuclear self and the primordial self.

15 Georg Northoff - How is our Self altered in Psychiatric Disorders – Neuropsychopathological Hypotheses
(University of Ottawa, Canada)

The Neurophilosopher and Neuroethicist Georg Northoff in his chapter and special issue of "Psychopathology of the Self in Psychiatric Disorders – Neuropsychopathological Hypotheses" conceives the self as central to our experience that has often been assumed to be necessary for any kind of consciousness in philosophy. For Northoff to understand the working and failures in the self we have to focus empirical attention on the midline structures in the brain. Recent investigations in neuroscience demonstrate- according to Northoff- a particular set of regions like the cortical midline regions to be associated with the processing of stimuli, specifically related to the self as distinguished from those remaining unrelated to the self. Furthermore, findings show close overlap between self-related activity and high levels of resting state activity in especially anterior midline regions. Recent findings in psychiatric disorders like depression and schizophrenia show both resting state abnormalities in exactly these regions that is, the cortical midline structures. Based on the psychopathology of the self and the recent neural findings of the resting state, Northoff develops a specific *neuropsychopathological hypotheses about the self in depression and schizophrenia* that are amenable to future experimental testing.

Altered Social Selves? Autism and Intersubjectivity

16 Inês Hipólito - On Autism and Interaction Theories of the Self
(New University of Lisbon, Portugal)

Hipólito reminds us that persons with Autistic Spectrum Disorder show various difficulties in social skills, cognitive processing and other co-occurring behavioral and physical problems. As people with autism demonstrate *a lack of ability in dealing with others*, the study of autism, with its specific constellation of behavioral and cognitive deficiencies this may highlight the structure, development and nature of social cognition in general and Self/Other awareness in particular. Hipólito challenges *the failure in >Theory of Mind< in Autism view* that characterizes social cognition as an individual achievement that happens within a particular person's brain and body and treats social phenomena as external events that require interpretation, namely, mind-reading. Alternatively, she defends Gallagher's "Interaction Theory" (Gallagher 2009) which proposes that the ability to "read" others is primarily a form of *body reading* rather than mind reading. Hereby the capability to comprehend other persons ultimately rests in a form of embodied practice that is emotional, sensory-motor, perceptual and nonconceptual.

Interaction Theory comprehends primary intersubjectivity and secondary intersubjectivity. *Primary Intersubjectivity* is seen hereby as the innate or early developing capacity to interact with others manifested at the level of perceptual experience— we see or more generally perceive in the other person's bodily movements, facial gestures, eye direction, what they intend and what they feel. On this view, in *second-person interactions*, the mind of the other is not entirely hidden or private, but is given and manifest in the other person's embodied behavior. Interaction theory- according to Hipólito- contends that these embodied practices constitute our primary access for understanding others, and continue to do so even after we attain theory of mind abilities, that in autism in this view is supposed to fail.

17 Anna Ciaunica - Autism: a Relational Self Impairment?
(University of Fribourg, Switzerland)

Ciaunica treats the controversy over the basis for young children's experience of themselves and others as separate yet related individuals, each with a mental perspective on the world – and over the nature of corresponding deficits in autism. On the dominant view, it has been argued that the human aptitude at inferring mental states is one of the crucial preconditions for the evolution of cooperative social structure in human societies. Consequently, *limitations in psychological perspective-taking among autistic children* have been regarded as signs of a "Mindblindness" deficit in having metarepresentational mental states (Baron-Cohen et al 1985; Leslie 1994). Lombardo and Baron-Cohen (2010) recently acknowledged that Mindreading researches have left a gap in terms of devices that may be responsible not only for atypical self-referential processes in autistic children but also in their integration into the bigger picture of *how individuals navigate and interact with the social world*. This paper defends the idea that the *intersubjective relatedness of the self is foundational for the development of higher order mentalizing skills and not the other way around*. Hence, autistic individuals might suffer from a *specific Relational-Self Impairment* which triggers metarepresentational deficits.

Ciaunica argues for a developmental account of the self in which the Self is build through Others by the centrality of what has come to be termed as "primary intersubjectivity" characterizes the mental experience of infants during infant–caregiver interactions from the earliest phases of life. Given this attunement between a baby's subjective experience of her own as well as of the caregiver's experience during the organized patterns of mother–infant interactions from birth (Rochat and Striano, 1999), it has been hypothesized that there is a basic human-specific drive to share psychological states with others such as in joint attention (see:

Seeman 2011; Metcalfe & Terrace 2013) and joint intentionality (Tomasello et al 2005, 2014). Moreover, attention has been paid to the intricate organization of the early bi-directional affective and imitative interaction sequences and their characteristic contingent 'protoconversational' turn-taking structure (Beebe et al 1997; Jaffe et al 2001). Later, towards the end of the first year, infants relate to others' actions and attitudes with reference to a shared world, for example by showing things to others and making requests, imitating others' actions on objects, and engaging in social referencing (Bretherton 1992). A further stage that occurs between the middle and the end of the second year is when children show conceptual understanding of self and other, for example in adjusting their actions to the needs and feelings of others, referring to themselves as 'I' and addressing others as 'you', and talking about their own and others' mental states. This latter stage has been typically taken to reflect a child's 'theory of mind', but also the child's increasingly sophisticated concept of 'selves', i.e., of individual persons who have distinctive psychological relations with the social and physical world. These findings according to Ciaunica suggest that infants' capacity to understand other minds is rooted in their capacity to actively *engage* in intersubjective scenarios. The self in this view is not understandable without an irreducible embodied aspect to such interactions (De Jaegher and Di Paolo 2007). The "interaction component" thesis stipulates that interaction *constitutes*, not just causally contributes, to intersubjective understanding.

Ciaunica's chapter opens up towards a second part of the Lisbon research on the *social* (cognitive) foundations of self, an important follow up result of the *Altered Self and Altered Self-Experience*.

Lisbon, July 2014
Alexander Gerner & Jorge Gonçalves

Endnotes

References

lbahari, M. (2011). "Nirvana and ownerless consciousness" In: M. Siderits, E. Thompson, & D. Zahavi (eds). *Self, No Self? Perspectives from Analytical, Phenomenological and Indian traditions*. Oxford: Oxford University Press, 79 – 113

Arzy, S., et al. (2006). Induction of an illusory shadow person. *Nature* 443: 287

Baron-Cohen, S. et al. (1985). "Does the autistic child have a ''theory of mind''?" *Cognition 21*, 37–46

Butler, J. (1990). *Gender trouble: feminism and the subversion of identity.* New York: Routledge

___ (1997). *The psychic life of power: theories in subjection*. Stanford, Calif.: Stanford University Press

Chateau, D. (ed.) (2011). *Subjectivity: Filmic Representation and the Spectator's Experience*. Amsterdam: Amsterdam University Press

Beebe B et al (1997). "Mother– infant interaction structures and presymbolic self and object representations" *Psychoanalytic Dialogues 7*, 113–182

Blanke, O. et al (2004). „Out-of-body experience and autoscopy of neurological origin." *Brain 127*: 243-258

Blanke, O. et al (2002). "Stimulating illusory own-body perceptions." *Nature* 419: 269-270

Blanke, O. (2012). „Multisensory brain mechanisms of bodily self-consciousness." *Nature Reviews Neuroscience 13*, 556-571

Blanke, O., Metzinger, T. (2009). "Full-body illusions and minimal phenomenal selfhood." *Trends in Cognitive Science 13*, 7-13.

Bretherton, I. (1992). "Social referencing, intentional communication, and the interfacing of minds in infancy." In: S. Feinman (ed.). *Social referencing and the social construction of reality in infancy*. New York: Plenum, 57–77

Broome, M. & Bortolotti, L. (2009). "Mental illness as mental: a defence of psychological realism."*Humana.Mente 11*, 25-44

Brugger, P. et al. (1997). "Illusory reduplication of one's own body: phenomenology and classification of autoscopic phenomena." *Cognitive Neuropsychiatry 2*, 19–38

Damásio, A. (2010). *O Livro da Consciência: A construção do cérebro consciente.* Lisboa: Círculo de Leitores

Davidson D (2001). *Subjective, Intersubjective, Objective.* Oxford: Clarendon Press

___ (1970). "Mental Events." In: Actions and Events. Oxford: Clarendon Press, 1980.

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

Dieguez, S., et al (2009). „Numbness for somebody else's finger." *Current Biology 19*:, R1108-1109

Ducommun, C. et al (2004). "Cortical motion deafness." *Neuron 43*, 765-777

Fornito, F. et al (2011). "General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance." *Biol Psychiatry*. *70(1)*, 64-72 doi: 10.1016/j.biopsych.2011.02.019

Fuchs, T. (2014). "The Virtual Other. Empathy in the Age of Virtuality." *Journal of Consciousness Studies 21, 5-6*, 152-173

___(2012). "Selbst und Schizophrenie". *DZPhil 60 (6)*, 887-901

Gallagher, S. (2013). "A Pattern theory of Self". Front Hum Neurosci. 7, 443; online: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3730118/

___(2005). How the Body Shapes the Mind. Oxford: Oxford University Press

___(2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci. (Regul. Ed.) 4*

___(2011). (Ed.) The Oxford Handbook of the Self. Oxford: Oxford University Press

__ (2012). "First-person perspective and immunity to error through misidentification." In: S. Miguens, G. Preyer (eds.). *Consciousness and Subjectivity.* Frankfurt: Philosophical Analysis Ontos Publisher, 187–214

___(2009) "Inference or Interaction: Social Cognition." In: *Philosophical Explorations. An International Journal for the Philosophy of Mind and Action 11 (3)*, 163-174

Ganesh, S. et al (2011). "How the Human Brain Goes Virtual: Distinct Cortical Regions of the Person-Processing Network Are Involved in Self-Identification with Virtual Agents", *Cerebral Cortex, Sept. 2011* doi: 10.1093/cercor/bhr227

Girard, R. (1987). *Things hidden since the foundation of the world*. Stanford: Stanford University Press

Goldie, P. (2007). "Emotion." *Philosophy Compass 2*, 928–938

Graham, G. (2010). *The Disordered Mind*. New York: Routledge

Ionta, S. et al. (2011). „Temporo-parietal cortex encodes self-location and first-person perspective." *Neuron 70*, 363-374

Jaffe, J. (2001). "Rhythms of dialogue in infancy: coordinated timing in development." *Monographs of the Society for Research in Child Development* 66 (2), i-viii, 1-132

He, B. & Raichle, M. (2009). „The fMRI signal, slow cortical potential and consciousness. *Trends Cognitive Science 13 (7)*, 302-309

He, B. et al (2008). „Electrophysiological correlates of the brain´s intrinsic large scale functional architecture". *Proceedings of the Natural Academy of Sciences of the United States of America 105 (41)*, 16039-44 doi:10.1073/pnas0807010105

Hofstadter D (2007). *I am a Strange Loop.* Basic Books

Hutto, D., Myin, E. (2013). *Radicalizing Enactivism. Basic Minds without Content.* Cambridge Mass.: MIT Press

Kyselo, M. and Di Paolo, E. (2013). "Locked-in Syndrome: a Challenge for Embodied Cognitive Science." *Phenomenology and the Cognitive Sciences*, 1-26

Landy, J. (2012). *How to Do Things with Fictions*. Oxford: Oxford University Press

Legrand, D., Ruby, P. (2009). "What is self-specific? Theoretical investigation and critical review of neuroimaging results." *Psychol. Rev. 116*, 252-282

Lenggenhager, B., et al (2007). „Video ergo sum. Manipulating bodily self-consciousness." *Science 317*, 1096-1099

Llinás, R., et al. (1998). "The neuronal basis for consciousness." Philos Trans R soc Lond B Biol Sci, 353 (1377), 1841-9

___(2002). I of the vortex: From neurons to self. Cambridge, Mass. London:

MIT Press

Lombardo, M. and Baron-Cohen, S. (2010), "The role of the self in Mindblindness in autism". *Consciousness and Cognition*, doi:10.1016/j.concog.2010.09.006

Lutz, A. et al (2008). "Attention regulation and monitoring in meditation." *Trends in Cognitive Sciences*, 12(4), 163-169

Malabou, C. (2013). "Go Wonder: Subjectivity and affects in Neurobiological Times" Part I In: A Johnston, C Malabou: *Self and emotional Life. Philosophy, Psychoanalysis, and Neuroscience*. Columbia University Press, 1-72

Merker, B. (2007). "Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. Target article, commentaries and author's response." *Behav. Brain Sci. 30*, 63–134

___(2005). "The liabilities of mobility: a selection pressure for the transition to consciousness in animal evolution." *Conscious. Cogn. 14*, 89–114

Metcalfe, J., Terrace, H. (eds.) (2013). *Agency and Joint Attention*. Oxford: Oxford University Press

Metzinger, T. (2003). *Being no one*. Cambridge: MIT Press

Mishara, A. (2010). "Autoscopy: Disrupted self in neuropsychiatric disorders and anomalous conscious states. In: S. Gallagher and D. Schmicking (eds). *Handbook of Phenomenology and Cognitive Science*, Berlin: Springer, 591- 634

Mohr, O., Blanke, C. (2005). "Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin. Implications for neurocognitive mechanisms of corporeal awareness and self consciousness." *Brain Research Review 50*, 148-199

Nagel, T. (1974). "What is it like to Be a Bat?" *Philosophical Review*, 435-50

Northoff, G. (2013). „Brain and self- a neurophilosophical account." *Child and Adolescent Psychiatry and Mental Health, 7*, 28

___ (2013 b). *Unlocking the Brain. Vol. 2 Consciousness*. Oxford: Oxford University Press

___ (2011)."Self and brain: What is self-related processing?" *Trends Cogn Sci 15*, 186-187

Northoff G et al (2006)."Self-referential processing in our brain—A meta-analysis of imaging studies on the self".*NeuroImage 31*, 440-457

Pessoa, F. (2013). *Eu sou uma antologia. 136 autores fictícios*. Ed. J. Pizzaro and P Ferrari. Lisbon: Tinta da China

Petkova, V., Ehrsson, H. (2008). "If I Were You: Perceptual Illusion of Body Swapping." *PLoS ONE 3(12)*, e3832. doi:10.1371/journal.pone.0003832

Pio Abreu, J. (2012). *Elementos de Psicopatologia Evolutiva*. Lisboa: Fundação Calouste Gulbenkian

Raichle, M. (2009). "A brief history of human brain mapping"*Trends in Neurosciences 32 (2)*, 118-126 doi:10.1016/j.tins.2008.11.001

Ratcliffe, M. (2008). *Feelings of Being: Phenomenology, Psychiatry, and the Sense of Reality*. Oxford: Oxford University Press

Rochat, P. (2011). "What is it like to be a newborn?" In: S. Gallagher (ed.). *The Oxford Handbook of the Self*. Oxford: Oxford University Press, 57–79

Rochat, P. and Striano, T. (1999). "Social-cognitive development in the first year." In: P Rochat (Ed.), *Early social cognition*. Mahwah, NJ: Lawrence Erlbaum, 3–34

Schechtman, M. (2013). "The narrative self". In: Gallagher S (ed.). *The Oxford Handbook of the Self*. Oxford: Oxford University Press, 392-415

Seemann, A. (ed.) (2011). *Joint Attention. New Developments in Psychology, Philosophy of Mind and Social Neuroscience*. Cambridge Mass: MIT

Shoemaker, S. (1984). *Identity, Cause and Mind*. Cambridge: Cambridge University Press

Slater, M. et al. (2009), "Inducing Illusory Ownership of a Virtual Body." *Frontiers in Neuroscience 3*, 214–220

Tomasello, M. (2014). *A Natural History of Human Thinking*. Harvard: Harvard University Press

Tomasello, M. et al (2005). "Understanding and sharing intentions: The origins of cultural cognition." *Behavioral and Brain Sciences 28*, 675–691; discussion 691– 735

Travis, F. & Shear, J. (2010). "Focused attention, open monitoring and automatic self-transcending: Categories to organize meditations from Vedic, Buddhist and Chinese traditions." *Consciousness and Cognition 19 (4)*, 1110-1118

Zahavi D (2003). "Phenomenology of the Self." In: T Kirchner, A David (eds.). The Self in Neuroscience and Psychiatry. Cambridge: Cambridge University Press, 2003, 56-75

I
Altered Self-Experience

...

1

Emotions and Altered Self

.

Dina Mendonça

Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

.

The paper[1] examines how emotions alter the self in various ways and explores the importance of recognizing different types of deepness of emotions in order to establish different ways in which emotions alter the self and proposing that anger works as a sonar device to find out information about the world and ourselves.

Why should we begin with anger?[2] In a very interesting article entitled "Anger, Present Injustice, and Future Revenge in Seneca's De Ira" Katja Maria Vogt writes that Seneca said that anger was the most hideous and frenzied of all the emotions, and that the general assumption about anger was that understanding its nature was a way to avoid its occurrence. As Vogt writes, "[t]he idea that we should aim at getting rid of anger is tied to the theoretical discussion of anger: once we understand what we actively do when we experience anger, we can stop short of getting angry" (Vogt 2006, 57) further on pointing out that, for the most part, scholarly debate assumes that Seneca's analysis of anger to be an exemplary analysis of an emotion and it can be expanded to a general theory of all emotions (Vogt 2006, 60). Vogt offer a thorough analysis of the definition and treatment of anger for Seneca raising several questions about its nature and the response given by the Stoic philosophical position. She explains how anger is, in spite of the particular attention that is devoted to it, particularly and surprisingly difficult to understand within the Stoic framework. One of the difficulties she indicates is the fact that emotions do not present a specific course of action (Vogt 2006, 61, 69, 73). The exemplary case of anger identified by the ancient tradition is perhaps due to the fact that anger is said to be one of the most frequently experienced emotions (Averill 1982), having consequences for health and illness for denying to feel anger, controlling its outbursts and general suppression of its experience is linked to a myriad of health problems such as asthma, ulcers, headaches, coronary heart disease and cancer

(Tucker and Friedman 1996; Siegman 1994; Tavris 1982: cited in Schieman 1999, 274).

In this paper I will begin by looking at the specific case of anger to unfold different levels of emotion's deepness, and similarly to Seneca's insight, think that some of the issues we observe with anger can be taken to be part of the overall conception of the emotions. One of the questions I want to keep in the background of the analysis is this questions about the connection of emotion to a course of action and the fact that no determinate course of action seems, at least at first, to be attached to emotions and return at the end to this pressing query.

It seems that when a person gets angry one alters the self because it makes a person perceive the world through the lenses of anger. The tone of subjectivity introduced by the feeling of anger can be identified as showing that anger makes people think more optimistically because dangers seem smaller, actions seem less risky, ventures seem more likely to succeed, and unfortunate events seem less likely (Lerner & Keltner 2001, 150). Consequently, angry people are more likely to make risky decisions, and make more optimistic risk assessments because anger increases the perception of what is familiar and decreases the perception of what is risky (Gambetti & Giusberti 2009, 14).

In addition, because emotions trigger certain changes in cognition, physiology and action the alterations caused by anger manifests itself in retrospective thinking as well, and therefore anger makes people less trusting, and slower to attribute good qualities to outsiders when they reflect upon events (Lerner & Keltner 2001, 151). An the list goes on as an angry person tends to anticipate other events that might cause anger, and tends to place more blame on another person for the misery, and anger can make a person more desiring of an object to which anger is tied, etc. The changes given by anger also seem to be sensitive to proximity of other people, that is, the status of anger changes with proximity and with the type of intimacy we have with people such that people who are close to us can offend us, wrong us, or deny us and overall make us more angry than distant people and thus anger is a far more difficult issue with those same close relationships.

What the previous description tells us is that when we consider anger solely by the episodic moments of anger we get a very poor image of what anger is making it be bound to simply identify physical events that happen with anger, or specific types of injustice that may cause anger. It is within this critical position about the nature of anger that we can understand Nichola Dent article entitled "Anger is a Short Madness: dealing with Anger in Émile's Education". Dent writes that we clearly "misunderstand anger, its depth and pervasiveness, if we think of it principally in terms of episodes or bouts, rather than in terms of a mode of understanding or a

cluster of such modes, with accompanying estimates of value, imperatives for action, patterns of reaction and so on which can endure and shape an entire character or culture" (Dent 2000, 314). Dent's paper indicates how we must conclude that, in order to understand emotion in all its complexity, it is necessary to think how anger modifies the person who feels it, how it modifies the persons' relationships to others and to the overall reality (Dent 2000, 314).

Thus, there are a wide variety of reasons why we can identify that people get angry and though we may be able to identify a specific episode for the manifestation of anger it is important to list all possible causes of anger to properly understand these episodes, as well as its consequences. First, people become angry when they think they, or those mattering to them, have been treated unjustly, contrary to their expectations of good treatment, when assumed social and interpersonal norms are violated. Secondly, people become angry when they are injured or insulted (*hubris*) by another such as when the person doing the action wants to show their superiority to the other person (bullying is a good example). Thirdly, people can get angry with another person not for what they in fact *do or have done*, but for what one assumes they *are likely to or will do.* Fourthly, people get angry when another person prevents, hinders, or even fails to help them with something they want and it feels to be due. As a consequence, even someone bothering a person in a state of unsatisfied desire can make them angry. Fifthly, people get angry when events go contrary to their expectations. And finally, we can say that people become angry when what they value, or what take very seriously, is denigrated or treated contemptuously by others[3]. To complexify matters even further the nature of anger its nature is also complemented by the cultural and literary inheritance of the meaning of anger with its manifold of meanings. In "A Brief History of Anger" Potegal and Novaco explain that thought when we look at scholarly work on anger there seems to be an overall suggestion that social tendencies work for moderating anger, there also seems to be the case that these tendencies are disturbed by the way in which anger plays an important role in maintaining the social hierarchy (Potegal & Novaco 2010, 19). So, thought there are several cultural restrictions on expression and experience of anger in many cultures around the world, as Potegal and Novaco identify, the human inheritance about anger must include both the prohibitions about anger as well as the images of sustenance of anger. That is, next to the fear of social and physical condemnation of anger, the belief that anger invites ill-health and misfortune and that expressing it is inappropriate for adults for it is amoral, bestial and childish, and finally that it is an emotion to be shameful of (Ibid, 12), we must also place the fact that in warrior type cultures, anger is cultivated as a way to produce indifference to wounds and pain and promote fearlessness in the battles (Ibid, 14). This is easily testified by the Viking fighter and also represented by Homer's Illiad which starts with the phrase "wrath of Achilles" and can be

considered as to contain a complete and complex description of anger and its rise from various social causes and its various manifestations both facially and by verbal expression and its subsequent action (Ibid, 13).

Since it has been shown that -"*The capacity to control expression of emotion, particularly negative emotions, develops over the first years of life and has particular importance for the unfolding of appropriate and adaptive social behavior (Eisenberg et al., 1996; Eisenberg, 1995; Thompson, 1994)" (Fox & Calkins 2003, 7)*- it must be the case that we learn anger within this complex and mixed information cultural heritage and that looking into how we learn about emotions may be an interesting way to tackle the complexity of emotion.

So we learn with others from birth how to feel, how to react to our feelings and emotions, how to integrate it into our emotional landscape. One of the strange outcomes of emotional education is that, unintentionally, the education of emotion aims to provide those growing up with a type of temporal neutrality, which can be stated as the general command:

"Feel this way (that I am telling you or showing you) truly and you will not be feeling this other emotions which you do because you happen to be four and get easily frustrated". (Brink 2003, 241)

This temporal neutrality has several consequences: one is to assume that education of emotion is a domestication of emotion such as to produce the proper way to behave and feel; second, it leaves to chance the development of emotions for it assumes a mistaken simplicity; and third it maintains the current ambiguous stance to emotions, as is seen in the case of anger. This ambiguity about emotions impregnates the overall look upon emotions making them seem to both justify certain choices and attitudes while, at the same time, they are taken as ways that impair and block the best choices and attitudes. Consequently, we tend to look at our own emotional lives through these same incoherent lenses and we seldom take into account the complex surroundings of an episode of an emotion. For example, we are not educated to interpret ourselves (or others) to take seriously into account the amount of tiredness of lack of sleep impact of emotional experience. So, though it is reasonably to find oneself cranky and generally unsatisfied with everything on a Friday after a whole of week of resting only 5 hours at night, the maxim we allow ourselves is the lack of control on displaying crankiness. Instead we tend to take the issues that promote the crankiness seriously while, at the same time, we clearly know that being tired tends to change our perception of the world so as to make people more prone to being cranky. So we can use this information in many ways depending on how it suits us: sometimes to diminish the impact and importance of others people's emotional reactions (if it so pleases us) by pointing out that a cranky

tired person does not deserve our attention because they are just exhausted; other times by validating his crankiness saying that though it is reasonable to be cranky due to tiredness it is also really unfortunate that this has happen to this person and she is right to be cranky.

What is interesting about the previous description of the education of emotion is that it allows us to understand some of the underlying assumptions about emotion. First, we assume that emotions happen in a specific moment of time thus reinforcing the conclusion generally accepted in psychology, neuroscience and anthropology that emotions last for brief seconds, and consequently the reaction to them (either to dominate them or follow them) is also a matter of reaction within a short period of time. If I am angry and I think I shouldn't be all I have to do is endure anger for a few seconds while it happens. Second, education of emotion assumes that there is a proper way to feel it (it should be clear independently of the fact of one following it or not). If I perceive an injustice towards me I should feel angry and if I don't it is either because I do not interpret it properly as an injustice or I have some inability to defend myself. Thirdly, education of emotion assumes that the value of an emotion is set thus dividing emotional life into positive or negative valence. In a very thorough paper on emotional valence, entitled "Appraising Valence" by Colombetti this assumption is clearly identified and analyzed as in urgent need of revision. And finally, education of emotion worries much more about the properness of dealing with an emotion than about the insights of alteration it causes. That is, an angry person does not think: good this alteration will make me have a sharper eye about this situation, nor it will tell her ways to release anger so as to overcome the alteration of anger but mainly to not feel what she is feeling.

There are many other consequences of the education of emotion regarding the nature of emotion such as to ignore the diversity of emotional processes within the same emotion-word and how details make the same emotional process slightly different and hold a different meaning, or how it leads us to take valence simplistically, or how it makes us overlook the importance of meta-emotions. In addition, it also gives us a misguided conception of the connection of time with emotional processes in general.

In an article entitled "Emotion" Peter Goldie has made an effort to provide a summary of all the facts that a theory of emotion needs to accommodate in order to be acceptable. Among these fact we find the notion of diversity and that a theory of emotion must somehow make sense of the wide range of diverse dimensions of the diversity of emotions. The first of these dimensions is the diversity of duration. Goldie writes that "[i]n this dimension, emotions range from short termed emotional reactions such as surprise at the sudden and unexpected noise behind you, to longer-term emotions such as enduring love of your children or parents. The

former is best understood as an emotional episode or experience, and the later as an emotional disposition, but we typically use the term 'emotion' for both episode and disposition" (Goldie 2007, 928).

The connection of time and emotion matters not just for a better understanding of how temporal dimensions are present in emotion but also for a deeper understanding of what emotion really is. In a article about the connection between time and agency, Flaherty states that time was a central issue for the intellectual agenda of George Herbert Mead because he understood that the way time is conceived determines the way in which human reality is understood. Accordingly, Flaherty writes, Mead begins his influential course Social Psychology with a critique of John B Watson behaviorism, among other things, because the behavioristic temporality (that a stimulus (i.e. cause) precedes a response (i.e. effect)) meant that the past determined the future, and "Mead recognized that this temporal framework excludes what he viewed as essential facets of human nature: choice, novelty, emergence, and improvisation" (Flaherty 2003, 17). Wierzbicka also testifies this in her book *Emotions across Languages and Cultures* (1999) when she points out that the idea that feelings have a very brief duration is the reflection of a strong behavioristic bias in the semantics of the English word emotion (Wierzbicka 1999, 20). Thus, accordingly, we find Ekman writing that his proposal is

"that emotions are typically a matter of seconds, not minutes or hours, is supported by some preliminary evidence. Examining spontaneous emotional events suggests a short time span. When subjects have reported experiencing an emotion for 15 or 20 minutes, and I have access to a videotaped record of their preceding behavior, I found that they showed multiple expression of that emotion. My interpretation of such incidents is that people summate in their verbal report what was actually a series of repeated but discrete emotion episodes" (Ekman 1992, 186).

Similarly many others will reserve the label of an emotion for emotional experiences that last a matter of a few minutes and a few hours (Oatley & Jenkins 1996, 124-125). In an introductory book on emotions, Oatley and Jenkins say that emotion is, in general, taken to have a limited time and that those emotions that are reported by people are typically described as lasting between a few minutes and a few hours. And they show a spectrum of affective phenomena (Oatley and Jenkins 1996, 124) where emotions last from seconds to minutes, while moods last hours and emotional disorders and personality traits are alteration that last from days to months.

Interestingly in the same book Oatley and Jenkins claim that there also seems present in the literature the idea that emotions do not happen as a whole, nor do they all happen in the same chronological time. That is, emotions do not usually

happen all at once because they are usually caused, then they run through a process, and then have consequences. Thus, Oatley and Jenkins point out, there is a widely accepted proposal made by Frijda (1986) of an emotion as a set of stages (Oatley & Jenkins 1996, 98). Oatley and Jenkins never explain how it is possible to hold both accounts true and both sustain that emotions last only seconds and that they have stages, so perhaps they assume these stages last for half or less than half of a second. Regardless it raises many questions: how we can measure the duration of theses different stages? Do different emotions have different durations of the different stages? Can the same emotion have two or more different states and different durations of the different stages? Regardless, it becomes clear that emotion is a temporal occurrence and that it implies that there is feeling it and being altered by it and then returning to a point of absence of feeling. Even if we assume that one never returns to the initial equal state after the alteration suffered from feeling an emotion, it still is the case that an alteration had been experienced and that there is a process, which with time becomes something else. It is this sense of time that emotions demand that is missed from the practice of the education of an emotion. That is, time interferes with emotions not only in their duration but also in the way in which people get out of the emotion

And that even if we maintain that an emotion lasts a few seconds it is still the case that a longer period of time must be considered to fully understand the alteration it causes. Rolls, for example, writes that emotion is a helpful concept to understand the mechanisms of reward and punishing decoding for a number of reasons, in which the fifth is that "[e]motional states have the important property that they persist for times in the order of minutes or hours, thus, maintaining persistence of behaviour and consistency of action even after the emotion-provoking stimulus has disappeared" (Rolls 2005, 37) pointing out that the alteration promoted by an emotion can last longer than the few seconds one can identify to be in a certain emotional state.

In sum, the fact that our education of emotion assumes and accepts that emotions are episodes that last a few seconds ignores several aspects of emotional processes, namely that its effect can last more than the reaction to a specific stimulus.

Thus, the poor analysis of the connection of emotion with time is far more complex as it encompasses not only the duration of an emotion, but also its frequency, the sequence of events before and after and the timing in which it happens (Flaherty 2003, 21). Consequently, a proper analysis of anger must include not only the alteration the subject suffers when one feels angry but also the return to a subsequent mode of something like a non-alteration after the occurrence of anger.

That is, if we define anger as an emotion that relates to one's psychological interpretation of having been offended, wronged, or denied and a tendency to react through retaliation, it is necessary to describe what happens after its occurrence, how does that feeling change afterwards. Similarly to the way one would like to say that feeling fear ends when one subsequently feels safe after having acted on fear. If a man feels fear when a bear is attacking, to use James's famous example, then the same man will only feel safe and out of the claws of fear when is his no longer within the reach of the bear's attack and has changed to another situation which is clearly safe. That is, we should read Aristotle famous statement that "A man who is angry at the right things and with the right people and, further, as he ought, when he ought, and as long as he ought, is praised" (Nichomachean Ethics 1125b32-34), as indicating that this alteration of the self is desirable within the proper and adequate limits and that the return to a non-alteration state is part of the adequate feeling of anger.

More importantly, it might even suggest that if anger does not happen, it is a sign of a type of failure, an undesirable indifference that we should be attentive too. Underlying the previous statement is the assumption that adequate anger happens when there is a violation of deep ethical values.

Accordingly, in *Sound Sentiments*, Pugmire shows that our deep emotions are connected to our deep values writing that, "it does not seem that an emotion can be profound unless it is morally adequate, where its subject has a moral status. Its moral adequacy can turn on the quality of the feeling of it, which give it the specific character that may serve as an analogue to a value. Thus, for instance, sentimental emotion strikes us both as hollow and as contemptible" (Pugmire 2005, 63). Though we may question Pugmire conclusion that Depth of emotion is a sign of excellence of character (Pugmire 2005, 64), no doubt Pugmire seems to be on a right track that deep emotions reveal our deepest values, though given that "[b]y the same token, deep emotion is a reflection of the world" (Pugmire 2005, 64) it may be the case that those deep emotions reveal character is a much larger scope that only excellence in character and that it may be quite hard (perhaps even impossible) to identify ourselves with a value we have been taught not to hold. But if we take the lesson to heart our emotional processes should not take into account only its causes, effects, alterations, but also what they reveal about what we hold dear and crucial for our lives. Consequently, *when we look at anger we should be aware that we can only learn how to avoid becoming angry, how to cease being angry and how to deal with other person's anger by considering seriously the hidden values it displays*. That it is, an episode of anger is a way into a search bellow the surface and an opportunity to investigate values that are somehow structural and verify in what ways one connected to crucial aspects of our life (e.g. people, ideas, objects) as well as how one is connected to

crucial events (e.g. childhood experiences, traumatic experiences), and how one is connected to difficult aspects of ourselves and the world to face (e.g. that one can feel envy, one can be mistaken, that the world is indifferent to our singular existence), and also how one is connected to emotional habits that sustain emotional life structure (e.g. righteous indignation promotes sense of self-worth necessary for daily function; being sad, miserable and pessimistic makes the events that go well feel thrilling), and what our character reveals.

Yet maintain that situations and finding oneself in certain situations is clearly connected to our feelings for a person may not know she is courageous until she is presented with an opportunity to exhibit courage and therefore whether or not one continues to be angry over a period of time does depend on the circumstances in which one finds oneself (Gert 1998, 255). Similarly it is possible to conceive a life in which someone's values are never threatened nor offended and the insight of this absence may be due both to absence of certain circumstances or lack of insight into them and consequently, education of emotion should include education of what to do in certain circumstances. As Kristjansson puts forward in "Can We Teach Justified Anger?" (2005) when he writes that there are behavioral strategies that allow the regulation and modification of behavior.

In addition, education of emotion should provide tools of identification and interpretation of situations so as to allow both a better grasp of selection of situations and their modification. For example, we can interpret diversion of attention as a way to avoid feeling a certain emotion, like the way young people use music to forget their unpleasant emotions, or the way older people are more likely to tend to chores or to engage in spiritual activities for the same purpose (Ben-Ze'ev, 2000, 238; Gross, 1998, 284). Yet, we can also read it as a way to transform the situation at hand and a way to modify the situation by bootstrapping ourselves in which we force ourselves, or are forced by others (de Sousa, 1987, 11) to transform an emotion-situation into another emotions-situation such as we can read the inhibition of a behavior that accompanies such bootstrapping as a way to modify the meaning of the situation. Accordingly, the young person who listens to music may be described as someone who is avoiding feeling angry and the demands of life and can also be described as developing a strong connection to music. Therefore, though "conative strategies of emotion regulation are more profound, and at the same time more difficult to implement, than are behavioral strategies because they involve a modification of desires and attitudes. (Ben-Ze'ev, 2000, 234–235), they may better be understood with a situational approach to emotions (Mendonça 2012) for it may be more clear that that the way in which they require "a serious overhaul of our self-assessment and our whole evaluative structure: of what we take to be our deepest engagements" (Ben-Ze'ev, 2000, 234–235) is also determined by how

we can interpret and analyze emotional processes within situational wholes. For these interpretations or reinterpretations of situations are ways in which we may change our beliefs or their meaning. As Kristjansson points out some of these strategies can assume the form of cognitive manipulations where one seems to be deceiving oneself, for instance "by managing to convince oneself, without good reason, that someone who leveled a bad insult did not really intend to do so. Flight attendants apparently learn an array of such techniques to avert anger towards obnoxious passengers by construing them as victims of suffering and hence as deserving of compassion rather than anger" (Kristjansson 2005, 678)

The description above indicates that the education of emotion is not like schooling (only obligatory up to a certain level) and that it continues throughout life. Therefore, just like education of emotion is not just telling children to not feel what they are feeling but also telling them stories, the fact that they experience seeing other people live their emotions, their experiencing of art, and other ways certain emotions are fostered in which we are not even aware such as the way fostering the ability to raise questions cultivates hope (Mendonça 2009) in the same manner our continuing emotional education is maintained in the same fashion. Of course, the openness to its possibility is an important part of the educational potential. If someone thinks there is only learning to control emotions then this will be the kind of learning that is mostly cultivated and perhaps the one with the most visible success. Regardless, a lot of our educational practices about emotional processes ask for hiding, accommodating others, redoing our emotional responses and perpetuates the denial of expression of certain emotions, the denial existence of other emotions and it is perhaps within the research on meta-emotions that its effect can best be understood (Mendonça 2013). Yet, regardless of how one deals with one's outburst of feeling anger and how one manages to control them or redirect them they should also be sources of information about ourselves such that we can say that being angry works as a type of sonar device which senses and identifies the boundaries of my social world, the boundaries of the ability and power to change the social world, as well as the boundaries of my ability and power to accept (or not) events of various kinds, and the boundaries of my abilities, the boundaries of my relationships, and finally the contours and boundaries of the ways of mastering emotions. It is of course the case that anger (or any other emotion for that matter) does not work only as a sonar device. It is known that anger also allows us to set social limits and boundaries[4] as theories about the sites of anger emphasize the importance of the social threats to identity and fairness (Schieman 1999, 275) and its impact on economics testifies how it sets boundaries (Zizzo 2008) for "anger works as a commitment to reject low offers if these are made, and, since it is a credible commitment, potentially angry receivers end up better off even in terms of their material utility" (Zizzo 2008, 154).

However, its sonar ability and how it causes alteration of the self and how this alteration demands a recovery is not an easy matter. Therefore the suggestion is but a first step to further research for it indicates that perhaps we should be pointing out to people the "sonar capacity of anger" as a way to monitor how emotional processes alter self and thus reveal its boundaries instead of simply and only teaching others and ourselves to restrain and hide feelings of anger. At the beginning of the paper I pointed out how others have identified that link to subsequent action with anger does not always refer to the same type of action. In "Moderation or the Middle Way: Two Approaches to Anger" Peter J. Vernezze shows that our historical inheritance also testify this ambiguity of action in anger. He writes that besides the Aristotelian tradition, which admits anger within the right and adequate circumstances, lies the Stoic and the Buddhist traditions which accept the underlying possible disturbance cause by anger but totally reject acting on it (Vernezze 2008, 6). The last perspective indicates that anger provides useful information about ourselves and the world even if we choose not to act on it.

I want to conclude by pointing out that the fact that we can be angry with ourselves may perhaps provide an interesting terrain to explore the sonar ability of emotions further.

Endnotes

[1] This paper was made possible by a FCT Post-doc Fellowship SFRH/BPD/14175/2007 and was presented inside the Project "Cognitive Foundation of the Self". Thanks to the audience of the International Workshop „The Altered Self and Altered Self-Experience" for the lively discussion.

[2] Thanks to Myriam Kyselo for calling my attention to this question.

[3] Taken partially from the handout of Gregory Sadler on "The *Rhetoric* Analysis of the Nature and Causes of Anger"

[4] Thanks to Rob Clowes for calling my attention to this trait of anger.

References:

Aristotle (1941). Nichomachean Ethics. In: R. Mc Keon (ed) *Basic Works of Aristotle*. New York: Random House, 935-1125

Averill, J. (1982). *Anger and aggression: An essay on emotion*. New York: Springer

Ben-Ze'ev, A. (2000). *The subtlety of emotions*. Cambridge: The MIT Press

Brink, D. (2003). "Prudence and Authenticity: Interpersonal Conflicts of Value" *The Philosophical Review 112 (2)*, 215-245

Colombetti, G. (2005). "Appraising Valence" *Journal of Consciousness Studies* 12 (8-10), 103-126

Dent, N. (2000). "Anger is a Short Madness: dealing with Anger in Émile's Education" *Journal of Philosophy of Education 34 (2)*, 313-325

De Sousa, R. (1987). *The Rationality of Emotion*, Cambridge, Mass.: MIT Press

___(2007). "Truth, Authenticity, and Rationality" *Dialectica*, 323-345

Ekman, P. (1992). "An argument for basic emotions", *Cognition and Emotion 6 (3/4)* (Special issue on basic emotions), 169-200

Flaherty, Michael. (2003) "Time Work: Customizing Temporal Experience" in *Social Psychology Quarterly 66(1)*, 17-33

Gambetti, E., Giusberti, F. (2009) "Dispositional anger and risk decision-making" *Mind & Society 8*, 7–20

Gert, H. (1998). "Anger and Chess". P. French, H. Wettstein (eds). *Midwest Studies in Philosophy XXII Philosophy of the Emotions.* University of Notre Dame Press: Notre Dame: Indiana, 249-265

Johnstone, A. (2012). "The Deep Bodily Roots of Emotion" *Husserl Studies 28 (3)*, 179-200

Goldie, P. (2007). Emotion. *Philosophy Compass 2*, 928–938

Gross, J. (1998). "The Emerging Field of Emotion Regulation: An Integrtive Review" *Review of General Psychology 2 (3)*, 271-299

Kristjansson, K. (2005). "Can We Teach Justified Anger?" *Journal of Philosophy of Education* 39 (4), 671-689

Lerner, J. and Dacher, K. (2001). "Fear, Anger, and Risk". *Journal of Personality and Social Psychology 81 (1)*, 146–159

Mendonça, D. (2009). "Let's Talk about Emotions" *Thinking: The Journal of Philosophy for Children 19, (2 & 3)*, 57-63

___(2012) Pattern of Sentiment-following a Deweyan suggestion" *Transactions of Charles Peirce Society* 48(2), 209-227

___(2013) "Emotions about Emotions" *Emotion Review 5(4)*, 1-7

Oatley, K. Jenkins, J. (1996). *Understanding Emotions*, Malden, USA/ Oxford, UK/ Victoria, Australia: Blackwell Publishing

Potegal, M. and Raymond Novaco (2010). "A Brief History of Anger" In: M. Potegal et al (eds). *International Handbook of Anger. Constituent and Concomitant Biological, Psychological, and Social Processes*. New York: Springer, 9-25

Pugmire, D. (1994). "Real Emotion" *Philosophy and Phenomenological Research, LIV (1)*, 105-122

___(2005). *Sound Sentiments. Integrity in the emotions.* Oxford: Oxford University Press

Rolls, E. (2005). *Emotion Explained.* University Press: Oxford

Schieman, S. (1999). "Age and Anger" *Journal of Health and Social Behavior 40 (3)*, 273-289

Vernezze, P. (2008). "Moderation or the Middle Way: Two Approaches to Anger" *Philosophy East and West 58(1)*, 2-16

Vogt, K. (2006). "Anger, Present Injustice, and Future Revenge in Seneca's De Ira". In: K. Vogt, W. Gareth (eds). *New Developments in Seneca Studies*. Leiden, 57-74

Wierzbicka, A. (1999). *Emotions Across Languages and Cultures: Dioversity and Universals*. Cambridge: Cambridge University Press

Zizzo, D. (2008). "Anger and economic rationality", *Journal of Economic Methodology 15(2)*, 147-1

2

Regaining the 'Lost Self': A Philosophical Analysis of Survivor's Guilt

∎

Amber L. Griffioen

University of Konstanz, Germany

∎

Abstract: *Although there has been much discussion regarding shame and guilt, not enough has been said about the complexities of the relationship between the two. In this paper, I examine one way in which I take shame and guilt to interact – namely in cases of so-called "survivor's guilt" among victims of trauma. More specifically, I argue that survivor's guilt may represent a kind of response to feelings of shame – one which is centrally tied to the central philosophical notions of autonomy and integrity.*

Keywords: guilt, shame, selfhood, agency, autonomy, integrity, trauma

It is so unjust, so unreasonable, that of all people the survivor should have to struggle, all by himself, with some of the greatest psychological difficulties imaginable […] – why should he in addition be obliged to feel a special responsibility, be persecuted by guilt, tortured by obviously unanswerable questions?

—Bruno Bettelheim[1]

[M]an has a right to be considered guilty and to be punished. To explain his guilt away by looking at him as the victim of circumstances also means taking away his human dignity. I would say that it is a prerogative of man to become guilty. To be sure, it also is his responsibility to overcome guilt.

—Viktor Frankl[2]

Shame is inventive.

—Friedrich Nietzsche[3]

1. Introduction

There has been much philosophical ink spilled on the subjects of shame and guilt. However, although much has been said about both how these emotions resemble one another and how they differ, not enough has been said about the complexities of the relationship between the two. Yet we may be able to come to a better understanding of these emotions if we look more closely at the ways in which they actually interact with one another. In what follows, I will examine one extreme type of case in which shame and guilt may be said to interact – namely in cases of so-called "survivor's guilt" among trauma victims.

Although no longer considered an official diagnostic criterion, the phenomenon of survivor's guilt played a very prominent role in the post-WWII development and discussion of the concept of Post-Traumatic Stress Disorder (PTSD).[4] Indeed, researchers have found that "trauma survivors often experience guilt that relates to the trauma in some way" (Kubany et al, 1996, 428).[5] Ruth Leys (2007) proposes that the deprioritization of the role of guilt in diagnosing PTSD is, in part, indicative of a "shift in attitude" regarding the survivor experience from a predominantly guilt-centric, "mimetic" approach, on which the survivor is said to unconsciously identify with the perpetrator of the trauma (8), to a more shame-centric approach that focuses on the "antimimetic consciousness of being seen" (10). Characterized in this way, the debate regarding the nature of survivor's guilt reflects an unfortunate theoretical polarization that views survivor's guilt solely in terms of either guilt or shame and thus fails to make room for the interaction of shame and guilt in the survivor experience.

In what follows, I hope to show that both shame *and* guilt play an important role in the etiology and manifestation of survivor's guilt. In so doing, I will resist the idea that shame is a mere reaction to a subject's "consciousness of being seen" and that survivor's guilt, likewise, is not best cashed out in terms of straightforward identification (unconscious or otherwise) with the actual perpetrator(s) of the trauma. Nevertheless, I hope to show that survivors' feelings of shame and guilt are importantly related. More specifically, I will argue that survivor's guilt, properly understood, may represent a kind of *response* to feelings of shame – one which is centrally tied to the philosophical notions of autonomy and integrity – and that it may serve the psychological function of helping the trauma victim grapple with the difficult process of re-integrating a shattered self into and within a world that no longer makes sense to her, albeit at a high cognitive price.

2. Understanding Shame and Guilt

To understand the phenomena that we are dealing with here, it will be helpful to look at the concepts of shame and guilt in a little more detail. Although they often

go hand in hand (and are thus sometimes conflated in the literature), shame and guilt are clearly distinct emotions. They are aimed at different kinds of objects and have distinct phenomenologies. Whereas the intentional object of guilt is something the agent does (or fails to do), the object of shame, properly speaking, is something about the agent herself. In other words, guilt focuses on an *action*, shame on the *self*. Of course, we regularly talk about feeling ashamed of things we have done or failed to do, but we really only experience shame in connection with these actions insofar as we take them to reveal something negative about *who we are* (e.g., in light of who we would like to be). Likewise, although concepts like "original sin" or "white guilt" appear to have persons as their object, the "guilt" in both of these cases refers back to the actions of some party or another, even if not those of the agent herself.

This distinction between shame and guilt can help us see how these two emotions may sometimes come apart. For example, I may feel ashamed, though not guilty, for having missed a free throw in an important basketball game. I experience shame in this scenario because I take the action to reflect something *about me*, about who I am (in this case, a basketball "failure"), in contrast with who I would like to be (e.g., a team hero). Likewise, I may feel guilty for having failed to fulfill a commitment to my significant other in a particular scenario (perhaps a work emergency makes me late for our anniversary dinner) without thereby feeling ashamed for having so failed. That is, I may take myself to have done something wrong (here, having failed my significant other) without believing that this particular case of wrongdoing reveals something about who I am as a person (e.g., that I am an unreliable partner).

A second important distinction to make between shame and guilt is in the way they present themselves to the agent. Although the felt character of both shame and guilt is unpleasant, Deonna et al (2011) note that in shame one often experiences a sense of insignificance and lack of control (5), whereas in guilt this sense of feeling "small and helpless" may be lacking, though the persistent "gnawing" of guilt at the agent may be accompanied by a desire to make up for the wrongdoing (13). Indeed, shame usually involves a desire to hide or disappear, whereas guilt (except insofar as it is accompanied by shame) rarely does. This desire to flee that accompanies shame is often characterized as a response to a feeling of being *exposed* under the disapproving gaze of a felt other (whether that other be physically present or not), and indeed many cases of shame do appear to display this voyeuristic aspect. Guilt, on the other hand, is less associated with the metaphor of sight than with that of sound – that is, of "hearing" the "voice" of judgment. The guilty person feels *judged*, either by herself (as with the "voice of conscience") or by another.[6]

However, we must be careful how far we extend these metaphors. As Deonna & Teroni (2008) note, to interpret shame as a wholly "social" emotion (that is, as *always* involving the perceived disapproval of a real or imagined audience) might be to

exclude certain cases of shame, in which no real or imagined audience appears to be present (69) – e.g., cases in which I feel ashamed, despite knowing that no one else would disapprove if they noticed.[7] Similarly, it seems clear that guilt – although involving a definite evaluative component – need not be triggered by an explicit recognition (the "hearing") of a negative moral judgment. Indeed, in some cases, the very fact that I feel guilty may be the first explicit indicator I have that I view an action of mine as morally problematic.

Given these distinctions, in what follows I will understand *shame* as an agent's affective apprehension of herself as failing to exemplify a certain trait that she values in a situation in which exemplifying that trait is relevant to maintaining her desired self-image.[8] *Guilt*, on the other hand, I will characterize as an agent's affective apprehension of an action or omission of hers as violating a (usually moral) normative requirement to which she adheres.[9] Thus, as Deonna et al. (2011) point out, shame appears to be a primarily *value*-driven notion, whereas guilt is *norm*-driven (82). Understood in this way, it should come as no surprise that in many cases, guilt and shame go hand-in-hand, given the common overlap between adhering to a norm and valuing the corresponding character trait that following this norm would exemplify.

Moreover, understanding shame and guilt along these lines allows us to make one more crucial distinction between these two emotions, namely that the latter always appears to imply an attribution of *responsibility* to the agent, whereas the former need (and, indeed, often does) not. While there is much debate in the literature as to whether or not guilt is a purely moral emotion,[10] there seems little doubt that it is *primarily* moral. In general, we do not feel guilty for violating what we take to be non-moral norms, nor for unknowingly or unwillingly violating those we take to be moral. One does not generally feel guilty about using the dessert fork instead of the salad fork when dining at a fancy restaurant (though one may feel ashamed), and neither do we think one ought feel guilty for unwittingly giving someone wrong directions or for missing a friend's wedding because one's flight was cancelled. If one feels guilty in these scenarios, we are likely to characterize that feeling as irrational, inappropriate, or at the very least unnecessary. Shame, on the other hand, is often experienced regarding features for which the agent is not directly responsible, as when one finds out one has unknowingly committed a social *faux pas* or when one is ashamed of certain of one's physical features, like nose size or skin color. Indeed, as we have said, the phenomenology of shame often includes a feeling of *lack* of control, as opposed to a sense of responsibility. (I return to this point below.)

Shame, then, may be said to be inappropriate[11] when either a) the values at stake are themselves silly or misguided (e.g., being ashamed for having thumbs or for

being female) or b) the agent perceives herself as failing to exemplify an appropriate value relevant to her self-image, when in fact she does not (e.g., feeling ashamed of one's physique in front of a person who finds one's body sexy). Guilt, likewise, may be inappropriate in two different ways: a) when the adherence to the norm in question is inappropriate (e.g., feeling guilty for failing to prioritize one's fantasy football team over one's spouse), or b) when one attributes responsibility to oneself for an action for which one was not, in fact, responsible. Yet the latter is precisely what appears to be the case in instances of survivor's guilt, in which the survivor of trauma feels guilty about something he had little to no control over. In fact, in many cases, survivor's guilt – given its sometimes stubborn persistence over time – appears straightforwardly *irrational*, insofar as the agent herself judges that she is not responsible for the traumatic event(s) in question. So how are we to understand this phenomenon?

The mimetic theories mentioned in Section (1) above would maintain that the agent unconsciously identifies with the actual perpetrator of the trauma, yet this seems to leave out cases in which there is no identifiable perpetrator to be imitated, as when one feels guilty for having outlived one's own child. Likewise, it fails to fully account for the deep sense of shame typically exhibited by trauma victims. Antimimetic theories, on the other hand, do emphasize the survivor's shame at being exposed in a certain way, but this both fails to account for cases in which there does not appear to be any real or imagined exposure (as with the case above) and does not fully appreciate the extent to which survivors appear to take on actual responsibility for some action or actions relevant to the trauma in question. Thus, a further account is needed to make sense of the complex relationship between shame and guilt in survivors of trauma. And I intend to provide just such an account. However, it may be helpful to first briefly classify a few of the various manifestations of survivor's guilt, in order to get a tighter grasp on the cluster of phenomena we are dealing with here.

3. Varieties of Survivor's Guilt

A prevalent manifestation of guilt feelings emerges among trauma survivors in the form of what we might call *substitution guilt*, namely the feeling that someone else has paid the price for or taken the place of the survivor. Agamben (1999) quotes author and activist Elie Wiesel, himself a concentration camp survivor, as saying: "I live, therefore I am guilty. I am here because a friend, an acquaintance, an unknown person *died in my place*." Another well-known Holocaust survivor, Primo Levi (1989), writes, "It is no more than a supposition, indeed the shadow of a suspicion: […] that each one of us […] has usurped his neighbor's place and lived in his stead. It is a supposition, but it gnaws at us; it has nestled deeply like a woodworm; although unseen from the outside, it gnaws and rasps" (81-2). This type of guilt –

namely the feeling that one has done something to "usurp" the place of another, or that someone who died is more "deserving" of life than oneself – is often found among survivors of genocide or war, but it is also observable in other instances, e.g., when children experience the death of a sibling or parents outlive their children. At the heart of this kind of guilt appears to be the perception of an unfair inequality or a gross injustice.[12] The agent assumes that she must have done something *wrong* to be on the receiving end of this unmerited benefit, even if she cannot exactly pinpoint what the relevant action was.[13]

In another form of survivor guilt, however, an agent may know precisely what she feels guilty for – namely for an act or omission of hers that under normal circumstances might be considered morally suspect or even impermissible but which either is excusable under the circumstances or bears no moral relationship to the traumatic event itself. Call this manifestation of survivor's guilt *transgression guilt*. Thus, a concentration camp survivor may feel guilty for having hoarded or stolen food, or for having failed to keep the Sabbath, or for having (forcibly) worked for the enemy. A soldier may feel guilty for having killed enemy combatants. A victim of date rape may feel guilty for having gotten drunk. But the object of the survivor's guilt need not be so specific. Sometimes the act or omission in question is much broader. Levi (1989) writes that many Holocaust survivors feel guilty, not for some specific action but more generally for "having failed in terms of human solidarity" (78). In several places, Levi expresses his guilt over having failed to help, to resist, to revolt, both during the rise of fascism and national socialism in Europe and later in the Lager itself, despite admitting that "on a rational plane there should not have been much to be ashamed of" (77).[14] In all of these cases – whether the action be of short or long duration, of narrow or broad scope – the agent has the feeling that she has committed a morally unjustifiable offense, despite her acknowledgement that the circumstances in fact permit, exculpate, or even morally justify such action.[15] Here, we might find a survivor's feelings of guilt somewhat easier to understand, since she feels guilty about an action or omission in which – under normal circumstances – guilt might be appropriate. Yet insofar as the agent *sincerely* admits that she does not deserve blame and yet still continues to feel guilty, and these cases are no less irrational or inappropriate than the above cases of substitution guilt, although they may occur at a different psychological level.

A related but distinct form of survivor guilt – what we might call *causal role guilt* – is found among many survivors of rape and other violent trauma, in which the agent does not necessarily feel guilty for having survived where others did not, nor for having committed an otherwise morally questionable action, but rather simply for having in some way or another causally contributed to the traumatic event in question. Thus, a rape victim may feel guilty for, e.g., having attended the party at

which she met her rapist, or for having worn supposedly "provocative" clothing, or for having failed to fight back harder against her attacker. In other words, she may feel that she somehow "invited" the rape or at least failed to prevent it from happening.[16] Here, an agent may not view her actions as morally reprehensible, but she may nevertheless experience feelings of guilt insofar as she attributes to herself a key causal role in the bringing about of the traumatic event that she did not actually play or that is irrelevant to the moral status of the event in question. This type of guilt is even more puzzling than transgression guilt, insofar as the agent blames herself for having merely been a part of the causal chain that led to her trauma. Yet playing a causal role is not tantamount to having transgressed a moral norm, so the guilt here appears especially displaced.

These three species of survivor's guilt are not meant to provide an exhaustive list of the various forms that survivor's guilt may take. Moreover, in some cases agents likely experience a kind of "mixed guilt" containing elements of each of these variations. Nevertheless, I think that this list suffices to get at one possible way that survivor's guilt may be plausibly understood, namely as a response to a certain kind of *shame* that often follows the experience of trauma – one that is closely tied to the agent-centered notions of autonomy and integrity. It is to a discussion of this relationship between shame and guilt in the manifestation of survivor's guilt that I now turn.

4. From Shame to Guilt: Trauma, Autonomy, and Integrity

As mentioned above, survivors of trauma – especially of violent trauma – often experience a deep sense of shame, both during the traumatic incident itself and as a result of it. But in what sense does the trauma survivor experience an "affective apprehension of herself as failing to exemplify a certain trait that she values in a situation in which exemplifying that trait is relevant to maintaining her desired self-image"?

It is my contention that for many trauma survivors the "self-relevant value" in question is not just any value. Rather, it is *the fundamental self-relevant value* – the value that grounds or otherwise frames all other values that we have – namely, the value we place on *agency*. To be a rational agent, one has to be able to set ends for oneself and to govern oneself in light of those ends – that is, one must be able to exercise one's *autonomy*. Further, rational agents are not isolated agents; they are diachronic agents within an intersubjective context that, in part, establishes the framework within which they can meaningfully and willfully undertake their intentional projects. Thus, being an agent means also that one must possess a minimal sense of one's self as doing things that make sense within one's immediate environment and have a set of commitments that is at least minimally coherent over time – that is, one must possess a sense of *integrity*. When either of these basic

conditions is threatened to a severe degree, this is likely to threaten one's overall sense of agency.

It is not difficult to see in what sense traumatic events threaten one's sense of autonomy. Indeed, it is likely that a perceived lack of autonomy is often (at least in part) what makes an event *traumatic* for a subject in the first place.[17] In most cases of violent trauma, the victims are treated by the perpetrator(s) as though they were not autonomous persons. The physical and emotional violation of a rapist toward his victim, the inhumane treatment of concentration camp prisoners and prisoners of war, the attitude toward slaves as "subhuman" beings or toward women as "weaker specimens" of the human race, the dehumanization of torture victims by their torturers – these are all ways in which perpetrators of trauma may be said to restrict (sometimes even to the point of eliminating) the autonomy of their victims (at least temporarily). But the examples need not be so extreme. A father may feel helpless to save his terminally ill child. A soldier may feel unable to defy orders, even when this results in the death of an innocent.[18] Or a child may not be able to stop his father from abusing his mother. And that same mother may herself so greatly fear making her way on her own that she feels unable to leave her husband. All of these cases might represent situations that severely threaten one's perception of oneself as an full-blown agent and contribute to a very basic experience of shame, given one's perceived inability to instantiate one's autonomy to the degree required for one to have a sense of agency (at least regarding the event(s) in question).

Of course, not every experience of lack of control is perceived as traumatic, and neither do we feel ashamed every time we are unable to implement our autonomy in situations that we care about. However, the shame experienced in trauma often goes deeper than a mere feeling of lack of control, though the agent's perception of her own powerlessness still plays a crucial role. Traumatic events are those that do not fit neatly into an individual's implicit or explicit worldview.[19] Not only does the trauma survivor experience a lack of control or a threat to her autonomy – the traumatic event is, by its very nature, one that disrupts the "flow" or "consistency" of her self-narrative. The traumatized subject can no longer make sense of herself and the world in which she lives in the way she did prior to the traumatic event. In a very important sense, she has lost the *trust* in the world necessary to orient herself to and within that world. Gry Ardal Printzlau points to this feature of trauma in her poignant discussion of the effects of torture on the victim:

"What survivors of extreme violence try to suggest when they talk about the loss of trust seems to be this profound loss of orientation, meaning, and connection with a shared social world. The loss of trust is a symptom of the shattered embeddedness of the self, and perhaps inextricable from it. Without being connected to a shared world we cannot give coherency and order to our experiences, and disorientation

results. Trust is not only an instrument for manoeuvrings in a complex world or a weak kind of belief which one can hold when proper proof is lacking, but appears to be fundamental to how we perceive our world as meaningful and real." (Printzlau 2011, 62)

When one's fundamental trust is shattered in this way, one may lose the ability to see oneself as a full-blown, unified self living in a comprehensible world in which one's actions and ends still make sense. In other words, one may lose one's fundamental sense of *integrity*. Thus, the trauma survivor finds herself unable to relate adequately to herself and to others in the way necessary for viewing herself as an autonomous, integrated self – as an *agent-in-the-world*.

We are now in a better position to tell a plausible story about the origins and persistence of survivor's guilt in victims of psychological trauma. As a matter of empirical fact, in everyday situations the subjective negative experience of lack of control often gives rise to superstitions and superstitious behavior, in which engaging in a kind of "magical" causal thinking serves the psychological function of allowing an agent to maintain or regain a sense of control over that which she perceives to be uncontrollable in situations where she cares about the result.[20] I want to suggest that something similar occurs in survivor's guilt, though on a much more profound level (given the severity of the perceived lack of agency) – namely, that the survivor of trauma attributes a kind of "magical" causality to herself (here, *agent* causality) as a response to the shame that arises from the perceived lack of autonomy she experiences regarding the traumatic event and the lack of integrity she experiences in its aftermath.

Assuming a subject cares about being an agent, she thus (at least implicitly) takes autonomy and integrity as constitutive self-relevant values. An event may present itself as traumatic, then, at least in part because it is perceived as severely restricting her ability to exemplify these basic values in the sense fundamentally relevant to her self-image (or to her having any self-image whatsoever).[21] Her apprehension of this inability, then, may manifest itself in a deep and pervasive shame regarding the traumatic event – and perhaps about herself in general (if the perceived violation of autonomy cuts deep enough). Thus, shame may arise in the form of a negative emotion directed at an agent's perceived lack of autonomy and integrity – and it is this negative emotion, I argue, that may motivate the attribution of guilt to oneself. Survivor's guilt, then, serves the psychological function of allowing a traumatized agent to retrospectively maintain a sense of agency in the perceived absence of autonomy and integrity, where this guilt assists the agent in coping with her life post-trauma. It is a way of reinstituting a sense of meaning for the agent that may allow her to re-orient herself in the world. As Susan Brison (2002), a philosopher and rape survivor, writes in her account of the aftermath of her rape, "it can be less

painful to believe that you did something blameworthy than it is to think that you live in a world where you can be attacked at any time, in any place, simply because you are a woman" (13). That is, it is better to perceive oneself as guilty than to try to eke out an existence in a fragmented world that no longer makes sense.

However, one might wonder: If guilt serves this positive, self-(re)building psychological function, wouldn't a more positively-oriented emotion be better equipped to do the job? Guilt, like shame, is a negatively-valenced emotion, and even if it does serve as a coping mechanism, it comes with its own burdensome psychological baggage. Yet as opposed to denial or repression, in which a traumatic event is no longer recognized (or recognized as such) – and in contrast to positive thinking, which downplays the "badness" of the trauma itself – survivor's guilt appears to be a somewhat more *honest* way of grappling with the negative nature of trauma. It does not reject the evil or badness of the traumatic event in question but rather attributes at least partial responsibility for that event or its consequences to the survivor herself. This allows for the recognition of the event as something objectively *bad*, but makes room for the subject to retain a sense of agency in light of her trauma. That is, although survivor's guilt may transform the subject's recollection of the relevant event(s) into something she can understand and control, the nature of the trauma remains itself negative and something that cannot merely be "done away with" via some sort of emotional transformation.

Thus, from the point of view of the subject, in the face of the evil nature of the event(s) in question, *it is better to be a guilty agent than no agent at all*. Re-establishing this sense of agency may represent a step toward the subject's being able to affirm her own humanity and to re-integrate herself and her experience into a meaningful whole. The narrative of autonomy established via survivor's guilt can help her recover a sense of meaning and thus of integrity, even if it results in a radically different worldview than the one implicitly held pre-trauma. As Bruno Bettelheim writes:

"What more impressive demonstration could there be that only the ability to feel guilty makes us human, particularly if, objectively seen, one is not guilty? […] [Our experience] taught us that there is meaning to life, […] a much deeper meaning than we had thought possible before we became survivors. And our feeling of guilt for having been so lucky as to survive the hell of the concentration camp is a most significant part of this meaning—testimony to a humanity that not even the abomination of the concentration camp can destroy." (Bettelheim 1976, 52)

In this way, survivor's guilt may assist a trauma survivor in recovering an at least minimal sense of autonomy and integrity – one that may be psychologically

necessary for her to overcome the pervasive shame of a subject whose very selfhood has been severely compromised.

5. Dealing with Survivor's Guilt

For the reasons given above, survivor's guilt may be said to (at least initially) perform a therapeutic function for the trauma survivor, and as such should not be immediately underplayed or dismissed as wholly irrational by family and friends, and especially not by medical or psychological professionals involved in the treatment of the survivor of trauma. Indeed, there is some evidence that, from a therapeutic standpoint, subjects who experience a sense of action-centered guilt cope better with their trauma than those who remain mired in person-centered shame. Ronnie Janoff-Bulman (1979, 1985) has shown that what she calls "behavioral self-blame" (blaming oneself for a particular action or actions, or a specific piece of behavior) may, indeed, be therapeutically adaptive, whereas "characterological self-blame" (blame aimed at an enduring character trait or tendency) is primarily maladaptive and associated with depression. I submit that this is because in the former case the subject affirms her agency by positively attributing autonomy to herself regarding the action in question, whereas in the latter case the subject negatively evaluates her own person and character – a phenomenon that looks much more like shame as we have characterized it above.

Still, even as a coping mechanism, survivor's guilt can be extremely psychologically taxing. It is, as we have said above, in some sense emotionally "inappropriate," insofar as the subject's guilt is centered on something for which she is not responsible (or at least not culpable). Thus, even if it helps the traumatized subject to establish a retrospective sense of control and thereby to make sense of the ultimately senseless evil she experienced, there is still a sense in which she might be at odds with herself, especially insofar as she in some sense sincerely and honestly assesses herself as innocent – as a *victim*. Whereas the presence of guilt feelings and the resulting belief that one is guilty[22] may start out as the result of an automatic defense mechanism or even as a kind of "motivated mistake," a subject may come to psychologically rely on and identify with her guilt in unhealthy ways. Indeed, subjects who obstinately hold onto their guilt over an extended period of time in the face of strong, recognized evidence to the contrary may slip into a kind of willful self-deception, which often comes at an even higher cognitive price and which can ultimately be autonomy-*undermining*. Not only does self-deception threaten one's general truth-orientedness and thus one's ability to act in a reality-responsive manner within one's environment, it aims at undermining the very standards that establish one's epistemic (and perhaps thereby also one's moral) agency. Furthermore, it is not clear that the consequences of self-deception are always (or even ever) better for one's overall happiness and flourishing than more

self-honest approaches. For this reason, although survivor's guilt itself may not always be maladaptive, professionals treating survivors of trauma would do well to look for alternative therapeutic strategies aimed at fostering a sense of autonomy and self-integration that can take the place of survivor's guilt in the subject's response to and evaluation of her trauma.[23]

6. Conclusion

In this paper, I have presented an account of survivor's guilt that demonstrates one very interesting way in which shame and guilt may interact in the etiology of and response to trauma. My account neither assesses the phenomenon as a straightforward unconscious identification with the perpetrator of the trauma nor reduces it to an instance of reproachful shame under the gaze of another. Instead, I have argued, survivor's guilt represents a response to a very basic sense of shame arising from a disrupted sense of agency grounded in a perceived lack of autonomy and integrity. The adoption of a feeling of guilt, then, allows the agent to maintain or regain a sense of responsibility and therefore of autonomy, one which can assist in the re-integration of the self into a world that has become otherwise incomprehensible to her. There is, of course, certainly more work to be done on this and related issues. How is survivor's guilt related to paradigm cases of epistemic irrationality, such as wishful thinking or self-deception? How can we evaluate mixed cases in which subjects are simultaneously both the perpetrators and victims of trauma (e.g., perpetrators who are traumatized by their role in inflicting trauma on others, or victims who as a result of their traumatic experiences become perpetrators)? Similarly, in contrast to the innocent subject who experiences guilt in order to escape the shame of heteronomy, what about the responsible agent who denies autonomy, in order to escape guilt? These are all fascinating avenues for future research, and it is my hope that this paper will serve as a springboard for the further discussion of these topics.[24]

Endnotes

1 Bettelheim (1986), 42.

2 Frankl (1978), 51.

3 Nietzsche (1886/1966), §40.

4 For a more detailed discussion of the role survivor's guilt has played in post-WWII discussions of PTSD, cf. Ley (2007).

5 Nevertheless, although guilt is "associated" with PTSD, it is neither necessary nor sufficient for a diagnosis. However, for this paper's purposes, it is sufficient to note that feelings of guilt and the survival of trauma often go hand in hand.

6 I have borrowed these metaphors from Williams (1993), 89-91.

7 For a more detailed discussion of this point, cf. also Deonna & Teroni (2011), 146-52.

8 Here, I depart from Deonna et al.'s (2011) characterization of shame as occurring "when we apprehend a trait or an action of ours that we take to exemplify the polar opposite of a self-relevant value as indicating our incapacity to exemplify this self-relevant value even to a minimal degree" (102). I am not sure that the agent need apprehend herself as exemplifying the relevant value's *polar opposite* to count as experiencing shame, nor that she need apprehend her failure as indicating an incapacity to exemplify the valued trait *even to a minimal degree*.

  I also wish to note here that I do not intend my characterizations of either shame or guilt to imply a strong cognitivist theory of the emotions. I would like to leave room for the possibility that the apprehension in question may be non-propositional.

9 Here I remain somewhat closer to Deonna et al.'s (2011) definition (cf. 114). I have inserted 'affective' into my characterizations of both 'shame' and 'guilt', since it seems that in either case, one could apprehend the relevant facts in question and yet fail to feel the appropriate emotion.

10 On this point, see for example Morris (1987); Deigh (1999); Murphy (1999); and Teroni & Bruun (2011).

11 The use of the term 'inappropriate' here is meant to be non-pejorative, since the inappropriateness of the emotion may rest on a mere cognitive or factual error, on the one hand, or a psychological compulsion, on the other. In such cases, although the emotion may rightly be said to be inappropriate in the sense of 'misdirected', it is non-culpably so.

12 I am indebted to Steven Arkonovich for pointing me to this feature of substitutionary guilt. Blacher (2000), too, notes that survivor's guilt often involves the feeling of "getting more than one's share" (46).

13 In some cases of substitutionary guilt, the object of the agent's guilt may simply be "some action or another," where the object of the agent's guilt remains vague and undetermined. In other cases, the agent may focus on an arbitrary incident prior to the traumatic event(s) in question for which she was responsible and postulate a false causal connection between that incident and her unfairly profiting.

14 Levi often conflates shame and guilt in his writing, but in this passage he is talking almost exclusively of guilt. Of course, that he and his fellow survivors are likewise *ashamed* of the being the *kind* of people who "had not done anything, or not enough, against the system into which we had been absorbed" (76) should also come as no surprise.

[15] Whether the guilt feeling is "inappropriate" here will lie in the details of the action in question. In some cases, we might think agents *ought* to feel guilty – as when a soldier brutally beats a prisoner or shoots an innocent civilian. In such cases, the agent may rightly feel responsible for a piece of wrongdoing, and it may be this very recognition that (in part) leads to his traumatization by the incident. Such "mixed" instances, in which an agent is traumatized by his role in perpetrating trauma, are worth exploring in more detail. However, my main concern in this paper is with instances in which the agent cognitively evaluates herself as having technically done nothing wrong and yet continues to experience pervasive guilt for the action(s) in question.

[16] Of course, in many cultures, the blame *is* in part often either explicitly or implicitly placed on the victim herself, and this may also directly contribute to the survivor's feelings of guilt.

[17] In their chapter, 'The Defining Moment of Psychological Trauma: What Makes a Traumatic Event Traumatic?', Everly & Lating (2003) suggest that one factor that makes a particular event traumatic for an agent is a feeling of helplessness, of lack of self-efficacy or control over a situation (45-7).

[18] Alternatively, a soldier may admit that he willingly followed orders (that is, that he exercised his autonomy in this respect), but he may still feel a lack of control regarding the situation in general – or, say, regarding his ability to both follow orders and save the innocent person. Indeed, one of the reasons moral dilemmas – e.g., in which fulfilling one's duty to X rules out fulfilling one's duty to Y and vice versa – are so psychologically conflicting might hinge on the fact that one perceives oneself as unable to fulfill the prima facie demands of morality, namely to do both X and Y.

[19] The idea of a disrupted explanatory worldview is another central feature of Everly & Lating's (2003) discussion of what makes an event traumatic for a person (cf. 36-8).

[20] Cf., e.g., Malinowksi (1954) and Thompson et al. (1998). Cf. also Griffioen (forthcoming, 2013) for a closer discussion of the connection between lack of control and superstitious belief, especially in sports contexts.

[21] This may also explain why some experiences are perceived of as traumatic by some subjects but not by others. Certain subjects may not apprehend the event(s) in question as restricting their agency in ways as devastating as others do, or they may have other methods of coping with the perceived failure of autonomy or integrity.

[22] Whether or not feeling guilty is always accompanied by the belief that one is guilty is an interesting question for further examination. But the fact that subjects who experience survivor's guilt also tend to experience a significant amount of cognitive dissonance indicates that there is likely some perceived tension between feeling guilty and the assessment that one has done nothing wrong – a tension that might be best explained by recourse to conflicting beliefs.

[23] Brison (2002) suggests, for example, that rape survivors take self-defense courses or engage in political activism as means "to break out of the double bind of self-blame or helplessness" (76).

References

Agamben, G. (1999). *Remnants of Auschwitz: The witness and the archive*. D. Heller-Roazen (Trans.). New York: Zone Books. Holocaust. Flamingo/Fontana Paperbacks

Bettelheim, B. (1976). "Reflections: Surviving." *The New Yorker, 2 August 1976*, 31-52

___ (1986). Surviving the Surviving. *The New Yorker*, 39-52

Blacher, R. (2000). "It isn't fair": Postoperative depression and other manifestations of survivor guilt." *General Hospital Psychiatry 22*, 43-8

Brison, S. (2002). *Aftermath: Violence and the remaking of a self*. Princeton, NJ: Princeton University Press

Deigh, J. (1999). "All Kinds of Guilt." *Law and Philosophy 18(4)*, 313-25. doi:10.1023/A:1006380226393

Deonna, J., et al (2011). *In defense of shame: The faces of an emotion*. Oxford: Oxford University Press.

Deonna, J. & Teroni, F. (2008). "Shame's guilt disproved." *Critical Quarterly 50(4)*, 65-72; doi:10.1111/j.1467-8705.2008.00850.x

Everly, G. & Lating, J. (2003). *Personality-Guided Therapy for Posttraumatic Stress Disorder*. The American Psychological Association

Frankl, V. (1978). *The unheard cry for meaning: Psychotherapy and humanism.* New York: Simon and Schuster

Griffioen, A. (forthcoming, 2013). "'Don't step on the foul line': Baseball and the (ir-) rationality of superstition in sports"

Janoff-Bulman, R. (1985). "The aftermath of victimization: Rebuilding shattered assumptions." In: C. Figley (Ed.), *Trauma and Its Wake*. Bristol, PA: Brunner/Mazel, 15-35

Janoff-Bulman, R. (1979). "Characterological versus behavioral self-blame: Inquiries into depression and rape." *Journal of Personality and Social Psychology 37(10)*, 1798-1809

Kubany, E. et al (1996). "Development and validation of the trauma-related guilt inventory (TRGI)." *Psychological Assessment, 8(4)*, 428-44. doi:10.1037/1040-3590.8.4.428

Levi, P. (1989). *The drowned and the saved*. New York: Vintage Books

Leys, R. (2007). *From guilt to shame: Auschwitz and after*. Princeton, NJ: Princeton University Press

Malinowski, B. (1954). *Magic, science and religion: and other essays.* Garden City, NY:

Morris, H. (1987). "Non-moral guilt." In: F. Schoeman (Ed.), Responsibility, character, and the emotions: New essays in moral psychology. Cambridge: Cambridge University Press, 220-41

Murphy, J. (1999). Shame creeps through guilt and feels like retribution. Law and Philosophy, 18(4), 327-44. doi:10.1023/A:1006332331845

Nietzsche, F. (1966). Beyond good and evil. (W. Kaufmann, Trans.). Cambridge: Cambridge University Press. (Original work published 1886)

Printzlau, G. (2011). Trust and ethical selfhood. (Unpublished Ph.D. thesis). Faculty of Theology: University of Copenhagen

Teroni, F. and Bruun, O. (2011). "Shame, guilt, and morality." Journal of Moral Philosophy, 8, 221-43. doi:10.1163/174552411X563574

Thompson, S., Armstrong, W., & Thomas, C. "Illusions of control, underestimations, and accuracy: A control heuristic explanation." Psychological Bulletin, 123(2), 143-61. doi:10.1037/0033-2909.123.2.143

Williams, B. (1993). Shame and necessity. Berkeley: University of California Press

...

3

Acting and the Self

·

Sara F. Bizarro

Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

·

Abstract: *In this paper, Douglas Hofstadter's view of the self as a "strange loop" is used in order to understand how several acting techniques work. As examples of acting techniques I will use the work of Lee Strasberg, Constantin Stanislavski, Stella Adler and Sanford Meisner. I will argue that Douglas Hofstadter's view of the self as a strange loop allows us to understand how acting works. I will furthermore argue that because Douglas Hofstadter's view is successful in explaining how different acting techniques work, that is a good indication of its adequacy as a theory of the self.*

Keywords: Self, Acting, Acting Techniques, Douglas Hofstadter, Strange Loop.

"A strange loop is a paradoxical level-crossing feedback loop."

Douglas Hofstadter, *I am a Strange Loop,* p. 102

Douglas Hofstadter has created an interesting and original theory of the Self, which he defines as being a "strange loop", a circular multi-leveled looping structure in the brain that would explain our psychological experience of having a self. Hofstadter's Self is created not only with an internal loop, but also with the external interaction of the Self and the world, what happens to the individual becomes formative in the Self's strange loop. Furthermore, one of Hofstadter's most original points is saying that we create these strange loops not only to represent ourselves, but also to represent others, and that there is no essential difference between these two types of loops. Hofstadter's wonderfully complex ideas about the Self loops and loops for Others, can help us understand how acting techniques work and how actors are able to build many characters and personify these characters in their work. Hofstadter's ideas can also explain how *different* acting techniques, or different approaches to acting, work.

This paper will be divided in three parts, in the first part I will introduce Hofstadter's views about the Self as a strange loop, including their original background in Hofstadter's earlier work, *Gödel, Escher and Bach* (1979). In the second part of the paper I will suggest a quick summary of different acting strategies, going from an emotion based acting (Strasberg, 1987), mixed strategies involving imagination (Stanislavski, 1936; Adler, 1988, 2000), and ending with action based acting (Meisner, 1987). In the third part I will use Hofstadter's views of the Self to explain how these different acting techniques can all function in creating convincing Self loops that flesh out the characters built by the actors.

1. The Self as a strange loop

In his book, *I am a Strange Loop*, Hofstadter goes through great lengths to explain what he means by a strange loop. A loop is a closed circuit where there seems to be an ever increasing in level, but nevertheless we end up where we started. One of Hofstadter's favorite examples is what we calls an enchanting experience he had as a kid when closing a cardboard box and folding the flaps in a cyclical order: "It always gave me a *frisson* of delight (and even today it still does a little bit) to perform that final *verboten* fold and thus feel I was flirting dangerously with paradoxicality." (Hofstadter, 2007, p. 101).

There are many other examples of different types of loops, some are circular, some tend towards infinity. Among the circular loops, Hofstadter includes de "lap loop", a game where each individual is sitting on another individual's lap until a complete circle is achieved and it is hard to say who is holding the group, but the lap loop does not collapse. Another simple illustration of a circular lap can easily be exemplified when we arrange several pens in the following way:
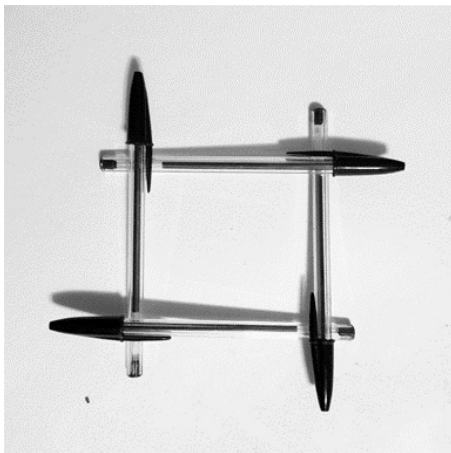
Image 1:
Pen loop: circular hierarchical system,

Photo S.F.Bizzaro

This is another version of the original example presented by Hofstadter when we close a cardboard box with overlapping flaps. The infinity generating loops are the ones that happen when we place two facing mirrors or when we have a camera that is plugged into a certain monitor that is showing what it is recording and we turn the camera towards the said monitor, in that case we get a video loop. Here is an example of the mirror loop:

Image 2:
Mirror loop: infinite regress;

Photo: Sara F. Bizzaro

But what does Hofstadter mean when he says the Self is a strange loop? Does he mean that there is an actual physical loopy structure in the brain that supports what we call the self? Yes and no. Hofstadter does think that there must be *some* structure in the brain that supports the self, but the exact nature of that structure is still unknown to us. What he means for the time being is that we can use this idea of a strange loop and create an abstract theory of the Self:

 "What I mean by "strange loop" is - here goes a first stab, anyway - not a physical circuit but an abstract loop in which, in the series of stages that constitute the cycling-around, there is a shift from one level of abstraction (or structure) to another, which feels like an upwards movement in a hierarchy, and yet somehow the successive "upward" shifts turn out to give rise to a closed cycle. That is, despite one's sense of departing ever further from one's origin, one winds up, to one's shock, exactly where one had started out." (Hofstadter, 2007, p. 101-102)

1.1. Gödel, Russell, Strange Loops, Self-reference and Paradox.

Hofstadter started thinking about strange loops in his earlier book, *Gödel, Escher and Bach* (1979). In this book Hofstadter talks about the first Gödel incompleteness theorem. Gödel incompleteness theorem was developed after Russell and Whitehead's *Principia of Mathematica*. In the *Principia of Mathematica* Russell and Whitehead were trying to reduce mathematics to logics, but the first Gödel incompleteness theorem states that there is a truth that cannot be proved in the system presented in Russell and Whitehead. Gödel was able to construct a self-referential sentence in PM by mapping the symbols used by Russell and Whitehead to Gödel numbers, which allowed him to construct what is called the Gödel sentence. The Gödel sentence is a statement that is true, but not provable in PM. Gödel statement is:

"This statement
cannot be proved"

Russell and Whitehead where well aware of the problems and paradoxes arrived at when constructing Self-referential sentences. In fact, their own system was developed after Russell encountered Self-referential issues in the previous work of Frege (1884) who had tried to reduce arithmetic to set theory. Russell had developed what is known as the Russell Paradox. Russell constructed a set that was defined as the set of all sets that do not belong to themselves, if this set belongs to itself, then it does not belong to itself and if it does not belong to itself, then it belongs to itself. To avoid these types of self-referential paradoxes, Russell had introduced what he called the Theory of Types that was build in order to prevent the construction of paradoxical self-reference logical sentences by keeping hierarchical levels separate, however, Gödel constructed a different interpretation of Russell's symbols broke the hierarchy blocking strategy that Russell had devised and the paradox reappeared.

As a young man, Hofstadter was fascinated by these are other examples of self-reference. However, what touches him the most is not that self-reference in certain systems can create paradoxical contractions, instead he is delighted by the idea that there is no way to block self-reference if the system presented is sufficiently complex, that is the uptake of Gödel's first incompleteness theorem for Hofstadter. However, the "self loop" that Hofstadter wants to introduce, is not necessarily of the self-destructive kind: "…negation plays no essential role in strange loopiness. It's just that the strangeness becomes more pungent or humorous if the loop enjoys a self-undermining quality. Recall Escher's Drawing Hands. There is no negation in it - both hands are drawing. Imagine if one was erasing the other!" (Hofstadter, 2007, p.159)

Hofstadter's is mainly interested in the idea that sufficiently complex systems cannot help themselves and become self-referential. This is the main vision that arose from Gödel's first incompleteness theorem: "What makes a strange loop appear in a brain (...) is an ability - the ability to think - which is in effect a one syllable word standing for the possession of a sufficiently large repertoire of trigger-able symbols. Just as the richness of whole numbers gave PM the power to represent phenomena of unlimited complexity and thus to twist back and engulf itself via Godel's construction, so our extensible repertoires of symbols give our brains the power to represent phenomena of unlimited complexity and thus to twist back and engulf themselves via a strange loop." (Hofstadter, 2007, p. 203).

1.2. Hofstadter's Self

The original intuition for Hofstadter is then that sufficient complex systems will inevitably be self-referential. But what is this and how does a Self appear and develop? What is a Self? For Hofstadter, the "I" represents a rather large and dominant structure in the brain that is not essentially different from other structures in the brain, but its presence has a larger impact because of its size and complexity. The structure that comprises the "I" includes endless episodic, emotional and introspective memories associated with it. The "I" is built as these memories are constructed and as the person interacts with the world the vision of the "I" changes accordingly:

"The vast amounts of stuff that we call "I" collectively give rise, at some particular moment, to some external action, much as a stone tossed into a pond gives rise to expanding rings of ripples. Soon, our action's myriad consequences start bouncing back at us, like the first ripples returning after bouncing of the pond's banks. What we receive back affords us the chance to perceive what our gradually metamorphosing "I" has wrought." (Hofstadter, 2007, p. 203)

For Hofstadter the Self is therefore a large complex looping structure that is not essentially different from other structures in the brain. The Self loop has a larger impact on our experience and behavior mostly because of its size. The Self loop consists of endless episodic, emotional and introspective memories. The Self also is affected and formed as the person interacts with the world, the Self loop is not completely internal, it reaches out to the world and changes as its experiences develop. One of the most fundamental ideas in Hofstadter is that the Self is not an essentially different from other brain structures and in fact the way we represent others is very similar to the way we represent ourselves.

1.3. Others are Strange Loops

According to Hofstadter strange loops do not only build Selves but they also represent Others. Besides being able to create our ever changing "I" with our

interaction with the world, we are also able to create similar structures to represent others, "you", "he", "she", "Robert", "Sandy" and so on. When we interact with other people, we assume that they perform the same creation of self that we do and we create a smaller loop to represent what we assume to be their own self-creation. However, since we have less information about others than about ourselves these loops, according to Hofstadter, are less elaborate, but not radically different from ours.

The Other for Hofstadter is a smaller structure, not essentially different from ours, only smaller in size, the fact that it has limited information makes it a less complex system and therefore it might be less rich and complete. One interesting result of Hofstadter's view of the Other as a strange loop is that one has to admit to some introspective access of the other. Our introspective access to others loops is limited by the amount of complexity of other loops, but in fact, if we know someone very well, we can introspect from the inside of our representation of the other. Of course this probably will not correspond with the experience the other has of his own Self loop, but there is no reason in principle that will necessarily stop the possibility of the introspection of other's Self loops. In fact, the more we have experiences with Others, the more we can tune our representations of them in our Other loops and the closer we can be having a similar introspective experience as they do. This ability to represent others as we do ourselves opens to way to understanding how acting works and how an actor can be successful in creating a character.

In the next section of this paper I will present several acting techniques in order to understand the tools actors use to create loops that represent the characters they want to embody. There are several types of acting techniques and I will try to give an overview of them in order to have a general perspective of how they work. Finally in the last section I will try to see how the acting techniques can be understood using Hofstadter's theories of the Self.

2. Acting

In a documentary for the PBS TV Series *American Masters* about Sanford Meisner, Sydney Pollack while speaking about Constantin Stanislavski, defined acting in the following way: "Acting has to do reproducing honest and truthful human behavior in imaginary circumstances." Even though this is a very good definition of acting, I would tweak it slightly in this way: "Acting is reproducing human behavior in imaginary circumstances in an honest and truthful way." I think this is probably closer to what Sydney Pollack meant, since not all characters played by actors are honest and truthful, an actor can also represent dishonest and misleading human behavior in imaginary circumstances in an honest and truthful way. After this small tweak I like this definition because it touches on three elements that pertain to all acting: behavior, imagination and truthfulness, which has space for the adequate

emotional content of the behavior being performed. Tentatively I will put forward the three essential elements of acting: emotion, imagination and action and I will try to show how different acting techniques use these three elements, although each kind of technique can stress one element as more important than the others.

In the next sections I will skim through the teachings of the four main acting teachers in theater history. I will organize them not chronologically but by how they stand in the line from emotion to action based acting with imagination being somewhere in between emotion and action. I will therefore start with Lee Strasberg (1987) because he has the most emotion-based approach also know as The Method. Then I will look at the teachings of Costantin Stanislavski (1936) who started as an emotion based acting teacher and later in his career is said to have stressed the power of imagination. After that I will look at the work of Stella Adler (1988) who stresses the importance of imagination in an actor's work. Finally, I will look at the technique proposed by Sanford Meisner (1987) that stresses the importance of action while teaching his acting technique. I will argue that even though each teacher stresses a different aspect of acting, they all agree with the triad of acting as involving emotion, imagination and action.

2.1. Strasberg

Lee Strasberg is an American actor and acting teacher who was the director of the Actor's Studio and who is considered the father of Method Acting in the United States, a method followed by many famous film actors: James Dean, Marilyn Monroe, Paul Newman, Robert De Niro, Al Pacino and many others are all said to be followers of Method Acting. Method Acting, as proposed by Strasberg, is a group of techniques and exercises that help an actor connect to a character by using their own personal emotions and memories. Method actors use what they affective, emotional and sensory memory, when they are building a character. A method actor will start by building a complete profile of the character he wants to portrait or embody, he will consider his history, relationships, experiences and so on, however, to be able to really embody the character, Method Acting encourages the actor to look into their personal emotional experiences in order to "flesh out" the emotional experiences of the character. The aim of the actor is to move the audience and the idea is that he can only move the audience if he is able to move himself. The basic idea is that a performance can only be experienced by the audience as moving, if it is experienced as moving by the actor and brining up deeply personal emotional memories is supposed to help in the genuine experience of the performance.

In his book, *A Dream of Passion* (1987), Lee Strasberg talks about the main dilemma of the actor, what is sometimes called "inspiration". How can an actor be moved day in and day out in a theater performance? Does an actor need to be moved in order to move the audience? This topic was analysed by Diderot in his book *The*

Paradox of Acting (1883), basically the problem for Diderot is that emotional performers cannot be expected to deliver a consistent performance, we can expect no "unity" from players who "play from the heart": "Their playing is alternately strong and feeble, fiery and cold, dull and sublime." (Diderot, 1883, p.8). However, Lee Strasberg did not conclude with Diderot that we must strip emotion from the performances, instead he looked for ways to preserve emotion in performances in a consistent way. He knew that this would be possible after seeing the Moscow Art Theater performances of Chekov plays in New York city guided by Stanislavski. Here is how Lee Strasberg describes these productions: "In the productions we saw in America, the plays were alive with vivid, intense, colorful experiences, while each moment was filled with marvelous creations of the experiences of the characters. There was never anything maudlin or pathetic or sentimental." It was after seeing the Moscow Art Theater performances that Lee Strasberg decided to become an actor and being that he knew Stanislavski had developed a consistent method for achieving performances of the stature that Lee Strasberg had witnessed, he later looked to develop such a method himself as director of the Actor's Studio.

After seeing the Moscow Art Theater, Lee Strasberg was convinced that "acting in the process of living on the stage" (Strasberg, 1987, p.63). There are two main parts of the stage experience that have to be dealt with by the actor, one is adequate preparation and the other is adequate expression. When setting up a character the actor in Stanislavski's method uses what Strasberg calls the "creative if", where he imagines himself in a certain set of circumstances and tries to see how he would react to such circumstances. Strasberg found that this only worked if the setting of the play was close enough to the actor's own experience, in classical plays were the experience is very far from the actor's lived experience, this strategy does not work. Strasberg suggested that the actor can imagine a different set of circumstances that he is familiar with and react with those in mind even if they have nothing to do with the play. By using this strategy the actor's performance will come across as credible because his emotions are sincere. Sincerity of emotion and experience on stage is the key to a good performance according to Lee Strasberg and the way to achieve this is by using one's existing previous emotional experiences.

There is a lot to say about each acting technique, but for the purposes of this paper I will just try to concentrate on the essential aspects of each technique and hopefully in the end we will be able to have a general view of acting technique in general and we will be able to try to understand the techniques in view of Hofstadter's theory of the Self. In the next chapter I will present a summary of the approach developed by the father of all the approaches described in this paper, Stanislavski.

2.2. Stanislavski

Constantin Stanislavski was the master of social realism in theater in Russia and inspired all psychological realism in theater in the United States. I am describing is work after Less Strasberg because even though he stresses the emotional content of a performance, he has a more balanced view on how to learn the art of acting. Stanislavski defines acting first and foremost as *action with a purpose*. His main instructional book, *An Actor Prepares*, is a fictional work written from the point of view of one of his students. In this book, the student describes going to the Moscow Art Theater and doing all the exercises proposed by the teacher. In many examples and situations the teacher shows the students that one can only perform an action convincingly if the action to be performed has a defined purpose, one cannot simply perform an action and arouse emotion, one needs a complete set up of the scene's background and character motivation in order to have the emotional content of the scene appear in a natural and honest way:

"Fix this for all time in your memories: *On the stage there cannot be, under any circumstances, action which is directed immediately at the arousing of a feeling for its own sake.* To ignore this rule results only in the most disgusting artificiality. *When you are choosing some bit of action leave feeling and spiritual content alone.* Never seek to be jealous, or to make love, or to suffer, for its own sake. *All such feelings are the result of something that has gone before. Of the thing that goes before you should think as hard as you can. As for the result, it will produce itself.*" (Stanislavski, p. 38)

So even though emotion is important for a performance to be moving, it cannot be the starting point for a performance. Stanislavski says that in order to perform a convincing action one must develop the character's situation, motivation, background and so on. In order to do this the actor will use the big "what if" guideline. However, since many of the plays will present the actor with situations that are far beyond what he has experienced in life and because the actor has to perform in a non-realist environment, the stage, one of the actor's main challenges is to develop his *imagination*. Stanislavski proposes several exercises regarding the development of an actor's imagination. In one example he asks the student to imagine that he is a tree, then asks the student to develop this in more detail: the student imagines he is a giant oak tree in the mountain standing alone in a large field, he hears the rustle of the leaves and he sees the birds on his branches but he can't imagine much else. The teacher then suggests that the student should explain why he is alone in a field and the student places the tree in a feudal era and says the feudal lord had all the other trees cut down for fear of the enemy and left that one tree because it created a shadow on a spring of water and in order to climb on it and look out for the enemy. At this point the student has a much more fleshed out idea of the tree. Now, since a tree does not move, it is a particularly difficult challenge

for an actor to create much more with this subject, but at this point the teacher asks the student to look into his personal emotional history and bring up something that stirs him in his real life. The student says he is particularly moved by fights and so the teacher says he should imagine the enemy's army is approaching, he is being attached with arrows and gun shots and so on. The student is able to bring up several strong emotions but since the tree cannot move, his physical dramatic expression is limited. Stanlislavski uses this example in order to illustrate how much background and preparation is before any action at all is performed. (Stanislavski, p. 61-66)

Imagination is important in order to create the background for action, but when action is being performed there are other techniques that are important to make the action look natural and real. Strategies like concentration of attention in certain objects and exercises in muscle relaxation are also a part of the actor's preparation. Finally, because an actor must be ready to repeat an emotional and truthful experience on stage day after day, Stanislavski stresses the importance of what he calls sensory and emotional memory. Sensory memory is memory of sensory inputs like sound, tactile memory, smells and so on. If an actor trains his sensory memory he will be able to bring it about on stage and he will be able to flesh out a scene with real sensory experience. Emotional memory is like regular action memory but its focus is on past emotions. If, for instance, one had a close friend who died, it is relatively easy to recall the emotion felt when one heard about their death. The actor must practice the recollection of all kinds of emotions. Then, besides practicing the recollection of emotion, the actor will associate certain actions to a certain emotion in order to create a conditioned reflex where a certain action brings forth a certain emotion. This requires a lot of training, but it is a way to achieve emotionally filled performances on stage that can be consistently repeated.

Summing up Stanislavsky's method: the aim of a theatrical performance is to achieve a truthful embodiment of a play on stage. All action on stage must be with a purpose and in order to establish that purpose actors need to exercise their imagination while fleshing out the characters background. After developing the character's setting the action on stage should bring forth the adequate emotional content for each scene. The actor then needs to develop his emotional memory so that he can repeat the emotional content of the play in a consistent way and in order to develop his emotional memory the actor associates certain actions with certain emotions. So, in Stanislavsky's method he uses the three aspects of acting: action, imagination and emotion, in a balanced way.

2.3. Adler
Stella Adler also witnessed Stanislavski's Moscow Art Theater performances in New York City, she was also part of the Group Theater with Lee Strasberg and

actually travelled to Paris to study with Stanislavski himself. Generally, Adler follows the Stanislvski's system, but her interpretation is different from Lee Strasberg in that Strasberg underlined the importance of personal emotions in the performance, while Adler underlined the importance of imagination. Adler had distain for the idea that actors have to look onto their personal emotional past in order to flesh out the characters. In most plays, the characters embodied by actors where so far from what the general middle class American experience had to offer, that it seemed to her absurd that the actor should look for the emotional content of his performances in such an experience:

"Nowadays a lot of what passes for acting is nothing more than finding yourself in some character. That doesn't interest me. Of course you have to bring your own experience to bear on the characters you play, but you have to realize right from the outset that Hamlet was not "a guy like you"." (Adler, 2000, p. 19)

Adler stressed that the actor should spend large amounts of time reading and becoming more and more knowledgeable about different cultures and different attitudes. The actor, according to her, should be constantly enriching himself and by doing so he will furnish more and more the most important quality an actor needs in order to succeed, his imagination:

"Ninety-nine percent of what you see and use on the stage comes from imagination. Onstage you will never have your own name and personality or be in your own house. Every person you talk to will have been written imaginatively by the playwright. Every circumstance you find yourself in will be an imaginary one. And so, every word, every action, must originate in the actor's imagination." (Adler, 1988, p. 17).

Adler also admits that emotion is very important in order to construct theatrical truth or a sense of being genuine on stage, however she claims that emotion and action in theater must have their root in imagination and therefore most of her exercises are constructed in order to develop this faculty in the actor.

2.4. Meisner
Last, but not lease, Meisner is another follower of Stanislavski's system, however, instead of stressing emotion or imagination, he stresses action. Meisner thought that the most important thing in order to achieve truthfulness on stage was to be completely present in the moment while doing the action. Meisner's motto was:

"The foundation of acting is the reality of doing" (Meisner p. 16)

In order to achieve this presence in the moment Meisner's students start by developing the ability to improvise first, the ability to live in the moment of the scene, to react to the other actors in the moment of the scene. Meisner's most famous

techniques are the repetition exercises where the actor's repeat something to each other until something impels them to move forward, for instance, one actor will say: "Your smart" and the other will answer: "I'm smart", then the first actor will repeat: "Your smart" and the second: "I am smart." The inflection in both actors will change as they repeat this exercise and certain emotions will arise, they might get irritated, bored and so on. The most important thing for Meisner is moment to moment spontaneity. When learning the lines, a Meisner actor will say them with no intonation, the intonation will only occur in the moment and can be radically different from performance to performance. This does not produce consistent performances, but guarantees honesty and truthfulness in each performance. Emotion and imagination are also present in Meisner's method, but the over stressing of each of these made for actor who had a very deep presence but acted with themselves and not with others, not in the moment of the scene.

There is much more that can be said about each approach to acting that I sketched here. However, for the sake of this paper I only wanted to try to show that these approaches all use the same basic elements in constructing a character: emotion, imagination and action. The starting point and the focus is in either one of them in each case, Strasberg started from emotion, Adler from imagination and Meisner from action, while the original teacher being this realist method, Stanislavski, balanced all three. Here is a graphic illustration of this idea:

Emotion	Imagination	Action

→

Lee Strasberg / Constantin Stanislavski / Stella Adler / Sanford Meisner
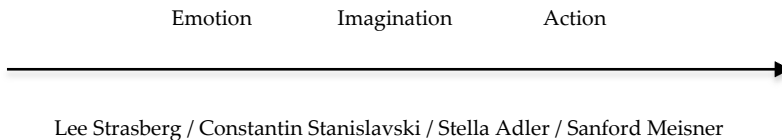
In the next and last section I will try to understand why this triad works in building a character by using Hofstadter's view of the Self. If the Self is somehow a unique system in the brain that belong only to the person in who it exists, it cannot be easily understood how actors can build convincing and non-artificial characters. If the Self is unique and exclusive of the bearer, then acting is doomed to fail and the aim of truth on stage that the realist Methods try to achieve is a red herring. However, if the Self, like Hofstadter claims, is not fundamentally different from other structures in the brain that we use to represent others, then all the techniques employed by these Methods make sense and we can understand how an actor builds a character.

3. The Actor's "Self"

How do Hofstadter's views allow us to understand what happens in the types of acting techniques presented in the previous section? Hofstadter's view of the Self

can help us understand how acting techniques work in several interesting ways: the idea that Self and Other are not substantially different in the brain allows us to understand how it is possible for an actor to create and embody a convincing character; self-referentiality can only be achieved in sufficiently complex systems, this is why the actor needs to create a sufficiently complex body of experiences and background for the character in order to incorporate it; when we attribute selves to others we do not need complete access to their intimate experiences, so we are also willing to attribute selves to the actor's characters when we witness them, however, there is a minimum of complexity in the actor's character self experience, if that is not present we experience the performance as fake or unconvincing; acting techniques also help us understand how selves are constituted by inner and outer relationships, actions and reactions, in the case of real life we have the cyclic triad: circumstances, action, emotion. In acting the circumstances are replaced by imagination, but the cyclic experience is the same; finally Hofstadter's views also allow us to understand why the techniques work no matter what aspect of the experience is underlined as long as all aspects are taken into consideration.

3.1. Self and Other

One of Hofstadter's main points about the Self is that the loop that constitutes the Self in the brain is not substantially different from the loops that represents Others. According to Hofstadter, the only difference between our Self loops and those that represent Others is that our loop is much larger and much more detailed than those that represent others. If this is so, then it is easy to understand that actors can *indeed* create and assume a character in their acting roles. Furthermore, Hofstadter's view explains why the actor has to go through such lengths in order to create that character, why he has to include endless circumstances, emotions and actions before the character can come to life. The complexity of the loop an actor creates to represent the character he wants to embody is somewhere in between the loops we create for ourselves and the loops we create for others.

Another of Hofstadter's main insights is that, following Gödel, any sufficiently complex system can become Self-referential, therefore if the actor works long enough and creates a complex enough character, self-referentiality will appear. In other words, an actor does not have to worry about "creating" a character out of the blue, what he needs to do is work on several aspects of a character, circumstances, emotions, actions, experiences, and the self-referential loop will necessarily appear once sufficient complexity has been achieved. In this way, what seems like an incredibly difficult achievement, creating a living character within ourselves, is actually a pretty innocuous result of the acting practices. Acting techniques help in that they speed up the process by picking out on the most essential elements and activities that bring about character creation and therefore make it easier and faster

for an actor to find the character in a theatrical setting. The character, therefore, is not created by the actor as a sculptor might create a statue, it appears from the work of the actor because the complexity of the psychological experience and content will necessary bring forth self-referentiality. When this happens, it is easy for a performance to look convincing and moving since in fact the character they are witnessing is a sufficiently embodied Self and can move others just as regular Other selves do.

3.2. Self and the World

Yet another aspect of Hofstadter's views about the Self is that the external world, including other selves, also has a constant constitutive role in our ever-changing loops. The loop is not only internal, but also external. The Self constituting loops can be illustrated as something of this sort:

Real World

Circumstances
(world, others)
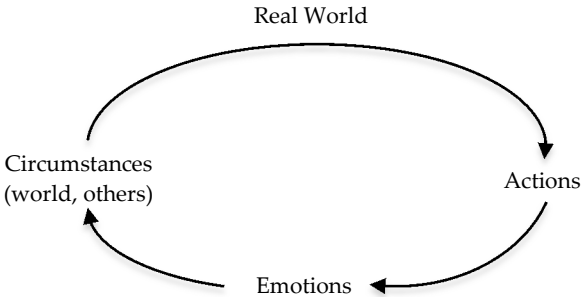
Actions

Emotions

In the real world, our Selves are constituted by elements of the ever-changing circumstances we find ourselves in, our interactions with others, our emotions and our own actions. The Self is a constant construction from these elements and an ever changing loop that becomes more and more complex with every step. In acting what happens is very similar, except the circumstances, the real world, are replaced by imagination:

Acting

Imagination
(imaginary circunstances, others

Actions

Emotions
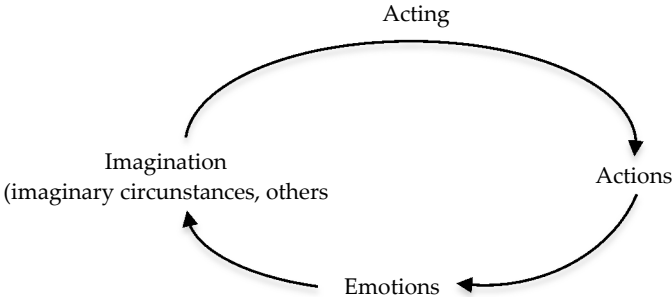
This understanding of acting also explains how the different strategies presented by Strasberg, Adler and Meisner work. It seems that one can start building character from either of the elements in the above graph, as long as all elements are present, each element is as good a starting point as any other. According to this interpretation of how acting techniques work, it seems that the most important aspect in acting is the extensive work an actor can do in all these elements for only with a sufficiently complex array of them can the character's Self naturally appear. Finally, the fact that Hofstadter's views about the Self allow us to create an understanding of how acting techniques function, can be a good argument for the accuracy of Hofstadter's views on the Self.

References:

Adler, S. (1988). *The Technique of Acting*. New York: Bantam Books

Adler, S. (2000). *The Art of Acting*. Montclair: Applause Books

Diderot, D. (1883). *The Paradox of Acting, Translated with Annotations from Diderot's "Paradoxe Sur Le Comedien"*. by Walter Herries Pollock with a preface by Henry Irvin. London: Chatto & Windus

Frege, G. (1884). *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Translated as The Foundations of Arithmetic: A logico-mathematical enquiry into the concept of number, by J.L. Austin, Oxford: Blackwell, second revised edition, 1974. Breslau: W. Koebner.

Gödel, K. (1962). *On Formally Undecidable Propositions of Principia Mathematica and Related Systems*. translated by Meltzer and Braithwaite. New York: Dover Publications

Hofstadter, D. (1979). *Godel, Escher, Bach: an Eternal Golden Braid*. New York, Basic Book

___ (2007). *I am a Strange Loop*. New York: Basic Books

Meisner, S. & Longwell, D. (1987). *Sanford Meisner on Acting*. New York: Vintage Books

Pollack, S. (Producer), & Doob, N. (Director). (1990). *Sanford Meisner: The American Theatre's Best Kept Secret*. [American Master's Documentary]. United States: PBS

Stanislavski, C. (1983[1936]). *An Actor Prepares*. New York: Theater Arts Books

Strasberg, L. (1987). *A Dream of Passion, The Development of the Method*. Little: Brown and Company

Whitehead, A. and Russell, B. (1910, 1912, 1913). *Principia Mathematica, 3 vols*, Cambridge: Cambridge University Press. Second edition, 1925 (Vol. 1), 1927 (Vols 2, 3). Abridged as Principia Mathematica to *56, Cambridge: Cambridge University Press, 1962

…

4

Cutting the Cords - Rethinking Judith Butler's Position on Transgender with René Girard's Theory of Mimesis

▪

Iwona Janicka
University of Cambridge

▪

Abstract: *One way of self-alteration is the alteration of the subjective experience of gender identity. In this case the alteration is defined in relation to an alleged social normality. Psychoanalysis tries to think this alteration but does so in an extremely conservative and normalizing way. These are limitations that prevent Judith Butler, who works with some of Sigmund Freud's key texts, to have a more open approach to thinking about the alterations of gender undergone by the self. In order to overcome these limitations and think about gender beyond sexual orientation I propose combining Judith Butler's insights with René Girard's work. In this paper I confront Judith Butler's ideas on gender and sexuality with René Girard's concept of mimesis. Through this theoretical encounter I wish to show how certain paths of interrogation are foreclosed for Butler because psychoanalysis does not offer conceptual tools to think about them. We will be able to observe how Girard's model accounts for constellations that Butler is interested in, such as the transgender, but which she does not and cannot fully explore using psychoanalytical framework. Girard's model is able to account for different modalities of the 'self': for being a man, a woman or a third entity; and for different modalities of desire: desiring homosexually, heterosexually, and polysexually. In mimesis the spectrum of gender behaviour is as broad as the scope of the model's desires and behaviours. This is where the disjunction between sexuality and gender is both possible and necessary. The idea of alteration based on mimesis is based on bodily repetition of models that are available to the subject in her socio-historical context.*

Keywords: Judith Butler, transgender, René Girard, mimesis, triangular desire

Judith Butler has been deeply preoccupied throughout her work with the question of transgender. Her theoretical thinking about gender and sexual identity is mainly anchored in psychoanalysis where she criticizes, reformulates and reuses psychoanalytical framework. From the transgender perspective, Butler has been

harshly criticized for not accounting for transgender identity as an identity operating in its own terms.[1] Namaste (1996) goes as far as to say: "Clearly, as scholars and as activists, we need to challenge Butler's negation of transgender identity" (Mamaste 1996, 188). In this paper I would like to confront Judith Butler's ideas on gender and sexuality with René Girard's concept of mimesis. Through this theoretical encounter I wish to show how certain paths of interrogation are foreclosed for Butler because psychoanalysis does not offer conceptual tools to think about them. We will be able to observe how Girard's model accounts for constellations that Butler is interested in but which she does not and cannot fully explore using psychoanalytical framework. She deals with the prohibition on incest and homosexuality but with psychoanalysis she lacks the tools to account for transgender position. René Girard will help us to propose an alternative way to conceptualize the relationship between sexuality and gender and so distance Butler from psychoanalysis.

Butler's account of gender identity in *Gender Trouble* draws on Freudian paradigm of ego formation through melancholia in an Oedipal constellation. She brings to our attention the fact that Freud in "The Ego and the Super-Ego" does not only describe "character formation" but also acquisition of gender identity. Freud claims that identification happens as a response to loss of a loved object. In order to preserve it we install the lost object in the ego. Freud claims that in the process of ego-formation a child desires one of its parents but the taboo against incest means that the desire has to be given up. Like the melancholic, who takes the loss into her and thereby preserves it, a child preserves the desired and lost parent through identification. The ego in Freud is then a repository of all desires that had to be given up. As Freud (1923) puts it: "the character of the ego is a precipitate of abandoned object-cathexes and […] it contains the history of those object-choices" (368 quoted in Salih 2002, 53). If the primary object of desire is the mother then the identification will be with her and if it is the father then with him. Freud does not, however, determine why a child desires one parent rather than another but he attributes it to a child's "disposition": either "feminine" or "masculine". The disposition of the child is its innate desire for the member of the opposite or the same sex. However, Freud himself is hesitant about dispositions and gives an example of a "little girl" who identifies with her father that brings "her masculinity into prominence". It would seem that the primary object choice is the result of primary disposition, feminine or masculine to start with, and Butler rejects such postulations (cf. Salih 2002, 52-56; Freud 1923, 372). "Feminine" would mean identifying with the mother and desiring the father and "masculine" - identifying with the father and desiring the mother. Butler (1990, 82) asks: "to what extent do we read the desire for the father as evidence of a feminine disposition"? Even though Freud suggests bisexuality as "originally present in the children", it seems to Butler that bisexuality in Freud "*is*

the coincidence of two heterosexual desires within a single psyche" (Butler 1990, 82). The masculine disposition is never directed towards the father as an object of love and the feminine disposition is never directed at the mother.[2] Therefore, there must be according to Butler another prohibition at work that directs this gender identification and is the source of problems that Freud encounters in accounting for identification.[3] Butler is interested in precisely those dispositions and how this primary object choice works.[4]

She proposes in this context the concept of primary prohibition on homosexuality where heterosexuality and gender identification are established not only through implementing the incest prohibition but, prior to that, through the prohibition on homosexuality (cf. Butler 1997, 135). As we know, for Freud identification with the same sex other is a psychic form of preserving the lost object. Due to cultural taboos against incest and homosexuality our passionate attachment to the same sex person (mother or father) is rejected. A girl becomes a woman by losing her mother as a primary object of love. This giving up an object of love is only possible through a melancholic "bringing inside" of her mother through identification with her and so becoming feminine (Butler 1997, 134). In case of a heterosexual union with the parent the object of desire is denied but not the modality of desire (heterosexual desire). In consequence, desire is deflected to other objects of the opposite sex constituting the normal process of mourning. In case of homosexual union with the parent both the object and the modality of desire require renunciation and so become subject to internalization through melancholia. In this way gender is acquired as internalization of prohibition on homosexuality. This identity is constructed and maintained by the consistent application of the taboo against homosexuality (cf. Butler 1990, 86). Thus, "homosexuality is *not* abolished but preserved, though preserved precisely in the prohibition on homosexuality" (cf. Butler 1997, 135-142) and gender identification is based on this set of disavowed attachments. Femininity is then formed through the refusal to grieve the feminine (the mother) as a possibility of love - "an exclusion never grieved, but 'preserved' through heightened […] identification" (Butler 1997, 146). A woman becomes a heterosexual melancholic where she refuses to acknowledge the attachment to the same sex and so a strictly straight woman is the truest lesbian melancholic (cf. Butler 1997, 147).[5] Butler finds it interesting because this account seems to follow from Freud's own claims (cf. Butler 2004, 199-200).[6]

To this Freudian interpretation, Butler importantly adds the surface of the body as the site where identification takes place (cf. Salih 2002, 56; Butler 1990, 86). Identification happens through bodily acts of "incorporation of the Other by mimetic practice" (Butler 1990a, 132-133): "Identifications are always made in response to loss of some kind, and […] they involve a certain *mimetic practice* that

seeks to incorporate the lost love within the very 'identity' of the one who remains" (Butler 1990a, 132). As we said above, the choice of the model is directed by the taboo against homosexuality and gender identification, therefore, follows from the "acting out" this prohibition. This constitutes precisely the difference between melancholia caused by separation, death or breaking of an emotional tie and melancholia in the Oedipal situation. As she puts it in *Gender Trouble*: "In the Oedipal situation [...] the loss is dictated by a *prohibition* attended by a set of punishments" (Butler 1990, 87). A child that enters the Oedipal drama with incestuous desires has already been subjected to prohibition of homosexuality. "Hence, dispositions that Freud assumes to be primary or constitutive facts of sexual life are effects of a law which, internalized, produces and regulates discrete gender identity and heterosexuality" (Butler 1990, 87). For Butler, therefore, gender identity is produced by incorporation of the cultural prohibition on homosexuality.[7] Our bodies enact this incorporated prohibition as femininity or masculinity constituting a residue of this primary, most important loss. The question that one needs to pose to Butler, however, is what about other prohibitions apart from incest and homosexuality that we incorporate in our bodies and which contribute to our gender identity? What about all that the melancholic phrase "I have lost nothing" incorporates? Does gender has to be always already sexual? What about the transgender position where sexuality and gender identity are disjoined?[8]

In order to see how Girard tackles the problems ingrained in psychoanalytical matrix let us first make explicit the undercurrent assumptions that Butler works with. First, for Freud and Butler desire is always already there. That one desires is a precondition for the functioning of their systems. Second, Butler assumes that taboo against homosexuality permeates our culture and affects the structure of desire and identity irrespectively of our immediate collectivity. The Foucaultian prohibition produces the primary desire as homosexual. Third, she takes up the psychoanalytical assumption that the disa**v**owed, like the Lacanian *Real*, structures our psyche. And fourth, that sexual desire is central to the formation of gender identity.

With Girard we can explain identification without recourse to primary prohibition or initial desire or, to put it in more psychoanalytical terms, without turning to castration or lack. Desire for Girard is produced in result of the mechanism of mimesis – desire is never autonomous but originates in the desire of the other that we see and unwittingly imitate. Therefore, desire always operates in triangular constellations. Girard accounts for the Freudian Oedipus complex with the *double bind* of mimesis. In his view, there is no way for a child to distinguish between the behaviour that is "good" to imitate (i.e. *non-acquisitive mimesis*, imitating a model) and behaviour that is not good to imitate (*acquisitive mimesis* that leads to mimetic

rivalry and thus violence because it wishes to appropriate the object of desire). In the case of a disciple-master relationship, which Girard gives as an example of a double bind, the disciple imitates the master. This imitation is clearly expected in the educational process. However, once the disciple surpasses the master, the master will become hostile towards this disciple due to emerging rivalry. The double bind takes place when a subject is incapable of correctly interpreting the double imperative that comes from the other person: as a model, imitate me; as a rival, do not imitate me (Girard 1987, 291). In Oedipus complex, Freud assumes that a child has the intrinsic desire for the mother and inherent narcissism. Girard refuses Freudian essences and claims that the farther is a model for apprenticeship and this inevitably involves also being a model in terms of sexual desire. The child, in Girard's view, does not know what it does when it imitates the father's desire for the mother. It does not distinguish between imitation and rivalry. It is the adult's interpretation, in this case Freud's, that accuses the child of sexual desire for the mother, whereas the child, according to Girard, only imitates unwittingly a model, i.e. the father (cf. Girard 1987, 352).

In Girard's system, therefore, gender identity would come from non-acquisitive imitation of a model or models: mother or father, mother or mother, father or father. This would mean imitating a whole range of conducts, norms embodied and repeated in their behaviours, and also, but not necessarily, the modality of their desires (be it homosexual or heterosexual). A disjunction of gender identity and sexual interest would be possible here. Let us, however, first think through the options that Girard's system proposes for the conflation of desire with identity which psychoanalysis makes. Let us also analyse the modalities of desire that are involved in triangulation. It has to be noted that these are purely structural considerations.

The triangular mimetic structure involves both heterosexual and homosexual desire at the same time. Depending on the constellation: heterosexual configuration implies homosexual desire or homosexual configuration implies heterosexual desire. Let us think it through step by step and first consider a triangular mimetic structure involving two men, as models and rivals, and a woman as the object of desire. It has to be pointed out that we are thinking here strictly in terms of triangles of desire that are characteristic for Girardian thinking. The constellation of departure here involves a heterosexual desire: both men desire a woman. One man admires the other and imitates his desire for the woman and so becomes a rival. Both imitate each other's desires for the woman and intensify the involvement with each other. The fierce rivalry leads to the obsession with the rival rather than the object of desire and leads to eroticizing of the rival. This is how heterosexual constellation involves inevitably homosexual structure.

However, if the constellation of departure is homosexual then there are two options with different consequences. Let us start with the simpler one. If all three tips of the triangle are of one gender, all parties involved are either only women or only men, then the structure would inevitably be of exclusively homosexual nature. This poses no further problems. If, however, the constellation of departure is homosexual yet mixed: a man and a woman are rivals over a woman as an object of desire then the rivalry is heterosexual in structure. Therefore, Girard claims that "there is no structural difference between the type of homosexuality and the type of heterosexuality that we are discussing at this point" (Girard 1987, 347). Both desires are involved here. The interesting part, however, in this constellation is the question of identification that results from such a structure and the question of the nature of desire. As Girard claims, obsessive rivalry can lead to becoming the double of the model-rival. If we conflate for a moment sexual desire with gender identity, as Butler does, then in the proposed "mixed homosexual structure of departure" this would mean that a man by imitating a desire of a woman for another woman would become a male lesbian and a woman by imitating the desire of a man for the woman would become a female straight man. This would then posit a structure of desire of a transgender person: a lesbian trapped in a male body, a straight man trapped in a female body. The interesting part in such a triangular structure is that Girard's structure of mimetic desire involves both desires operating at the same time if the parties are of different gender. And that it is able to account for desire in a transgender person whereas psychoanalysis is not. It could also account for a transgender desire of wanting to be a different gender irrespectively of the modality of desire (homosexual or heterosexual) or irrespectively of any sexual desire at all if we do not conflate sexual desire with gender identity.

We can observe here how Girard's model accounts for constellations that Butler is interested in – the transgender – but which she does not and cannot fully explore using psychoanalytical framework. She deals with the prohibition on incest and homosexuality but with psychoanalysis she lacks the tools to account for transgender position. The question of identification in transgender does not make itself available for conceptualization with Freud and we wonder if it is a necessary spectre that has to haunt Butler's system. Transgender desire is unintelligible from within psychoanalysis because there are only two options possible in that framework: homosexuality and heterosexuality, the man and the woman. However problematic Butler has tried to make those psychoanalytical presuppositions, her subversive reading could only lead her reformulation to occur within the bounds of psychoanalytical framework: questioning heterosexual matrix and binary logic of male and female rather than actively exploring other modalities. This is perhaps why Butler does not account for transgender as much as she could if she employed a different framework, despite the fact that transgender and violence towards

transgender people is such an important preoccupation to her thinking. Transgender serves as the abject that panics heteronormativity. Yet, the mechanisms of transgender identity are not explored as much as mechanisms of homosexuality and heterosexuality, masculinity and femininity.

There is a profound conservative normativity in psychoanalysis that is at variance with Butler's theoretical interests and Girard offers a possibility to overcome it. Girard's model is able to account for different modalities of being and desire: for being a man, a woman or a third entity and desiring homosexually, heterosexually, and polysexually. Butler makes the connection desire-loss and prohibition-identity through Freud and Foucault to show how unacknowledged loss structures who we are. For Girard, the pathways of desire and identity do not have to necessarily overlap. Desire is constructed through mimesis and identity is constructed through mimesis but there is a whole spectrum of gender identity that is not constructed as a result of sexual desire, prohibited or not. This is where the disjunction between sexuality and gender would be possible. Hence the link desire-identity is much weaker in Girardian theory than it is in Butler and forecloses the possibility of a formative primary experience. The question that one is compelled to ask Butler is if this "ungrievable loss", if this melancholia which constitutes our gender identity, has to be necessarily related to sexuality. And more importantly if an asexual position is possible both in our culture and in Butler?[9]

Endnotes

1 On transgender critique of Butler see Prosser 1998 and 2006; Namaste 2000 and 1996, 183-190; Bettcher 2009. For Butler on transgender see many references scattered throughout Butler 1993; Butler 2004; Butler 1990; in particular: Butler 1993, 128-137; Butler, 2004, p. 76-101 and Butler 1990, 130-150; also Butler 1993a; Butler 2006; Butler 1998. Unfortunately, we lack space to comment in detail on all the vehement disagreements between transgender studies and Butler, as one of the founding figures of queer studies.

2 We can, however, ask provokingly, this time perhaps more of Butler than of Freud, why this masculinity of the little girl could not possibly be read as a case of foreclosed transsexuality rather than foreclosed homosexuality? "Sigmund Freud never discussed transsexualism *per se*" (Rosario 1996, 38), yet Butler has and this foreclosure in Butler could be considered to be problematic from transgender perspective. Butler pays little attention to the case of the masculine "little girl" in Freud's text and so, in a way, erases the figure of the transgenderist here.

3 For the problems Freud has with positing the source of identification see (Freud 1923, 371-372; Butler 1990, 78-89).

4 Butler often uses examples of female Oedipus complex structure because, from the perspective of her reading, the male Oedipus complex has a more complicated double displacement (cf. Butler 1997, 127-138). From Girardian perspective, however, the difference between female and male Oedipus complex is irrelevant and that is why we do not follow the psychoanalytically based distinction between male and female Oedipus.

5 Butler acknowledges that "a homosexual for whom heterosexual desire is unthinkable" will maintain his or her heterosexual desire through melancholic incorporation. However, due to different cultural sanctions towards homosexuality and heterosexuality the two types of melancholia are not equivalent (Butler 1990, 95; cf. Salih 2002, 57).

6 A similar structure of melancholia is posited in *Antigone's Claim* where Antigone becomes, in a way, her brother because she lost him. However, the question is then why she chooses particularly her brother Polyneices as a figure of identification if she had also lost other people including her other brother Eteocles, her sister (whom she repudiates), her mother or her father-brother Oedipus, who are all left ungrieved (cf. Butler 2000, 67; 79-81). There is an implication of incest in the relation with Polyneices. Does the object have to be sexual?

7 Although in 1990 she explicitly concedes a doubt: "whether loss or mimetism is primary [is] perhaps an undecidable problem" (Butler 1990a, 133), it seems that a couple of years later, in 1997 in *The Psychic Life of Power*, she decides on loss.

8 From transgender position there is no necessary relationship between gender identity and sexuality: "Trans people generally undergo gender reassignment without changing their sexual orientation – transsexual people are just as likely to be straight, gay, lesbian or bisexual in their new gender role as any people are. […] There is no relationship to sexuality" (Whittle, 2000, 19).

9 Butler problematizes the separation of gender from sexuality I n her "Against Proper Objects" (1994) and argues that "an analysis of sexual relations apart from an analysis of gender relations is [not] possible" (9). Although we agree that analyses of gender and sexuality should not be strictly separated, the point of contention is that sexual orientation (irrespective of whether hetero- or homosexuality) should not be posited in a causal relation

to gender identity. This is, however, what Butler implicitly does in her interpretation of heterosexual melancholia. Although she underlines the importance of categories of "class" and "race" as equally important as gender to feminist analyses and criticizes the binary frame, in which the gender-sexuality discussion entraps feminism, it is not race or class that are reinscribed into her re-reading of Freud but foreclosed homosexuality (see Butler 1994 and Prosser, 2006, 277-280 for a transgender critique of Butler's ideas in "Against Proper Objects").

References

Bettcher, T. (n.d.). *Feminist Perspectives on Trans Issues*. Retrieved 18th of April 2013, from http://plato.stanford.edu/archives/win2009/entries/feminism-trans/

Butler, J. (1990). *Gender trouble: feminism and the subversion of identity.* New York: Routledge

Butler, J. (1990a). "Imitation and gender insubordination." In: J. Butler & S. Salih (eds) (2004). *The Judith Butler reader*. Malden MA & Oxford: Blackwell Publishing, 119-137

Butler, J. (1993). *Bodies that matter: on the discursive limits of 'sex'*. New York & London: Routledge

Butler, J. (1993a). "Critically queer." In: E. Striff (Ed.) (2003), *Performance studies*. New York, NY: Palgrave Macmillan, 152-165

Butler, J. (1994). "Against proper objects." *Differences. A Journal of Feminist Cultural Studies*, 6(2+3),1-26

Butler, J. (1997). *The psychic life of power: theories in subjection*. Stanford, Calif.: Stanford University Press

Butler, J. (1998). "Afterword for Transgender in Latin America: Persons, Practices, Meanings." Special issue of *Sexualities*, *I*(3)

Butler, J. (2000). *Antigone's claim: kinship between life & death*. New York: Columbia University Press

Butler, J. (2004). *Undoing gender*. New York & London: Routledge

Butler, J. (2006). "Doing justice to someone: sex reassignment and allegories of transsexuality." In: S. Stryker & S. Whittle (Eds.). *The Transgender studies reader*. New York & London: Routledge, 183-193

Freud, S. et al [1923](1991). *On metapsychology: the theory of psychoanalysis : 'Beyond the pleasure principle,' 'The ego and the id' and other works* (Vol. 11). Harmondsworth: Penguin

Girard, R. (1987). *Things hidden since the foundation of the world*. Stanford: Stanford University Press

Namaste, K. (1996). "'Tragic misreadings': queer theory's erasure of transgender." In: B. Breemyn & M. Eliason (Eds.). *Queer Studies. A Lesbian, Gay, Bisexual, and Transgender Anthology*. New York & London: New York University Press, 183-203

Namaste, V. (2000). *Invisible lives: the erasure of transsexual and transgendered people*. Chicago, IL & London: University of Chicago Press

Prosser, J. (1998). *Second skins: the body narratives of transsexuality*. New York: Columbia University Press

Prosser, J. (2006). "Judith Butler: queer feminism, transgender, and the transubstantiation of sex." In: Stryker, S., & Whittle, S. (Eds.), *The transgender studies reader*. London: Routledge, 257-280

Rosario II, V. (1996). "Trans(Homo) Sexuality? Double Inversion, Psychiatric Confusion, and Hetero-Hegemony." In: B. Breemyn & M. Eliason (Eds.). *Queer Studies. A Lesbian, Gay, Bisexual, and Transgender Anthology*. New York & London: New York University Press, 35-51

Salih, S. (2002). *Judith Butler*. London: Routledge

Whittle, S. (2000). *The transgender debate: the crisis surrounding gender identity*. Reading: South Stree

...

5

On the Cinematic Self. Cinematic experience as "Out-of-Body" experience?

.

Alexander Gerner[1]
CFCUL, Universidade de Lisboa, Portugal
Michele Guerra
University of Parma, Italy

.

We are our bodies—but in that very basic notion one also discovers that our bodies have an amazing plasticity and polymorphism that is often brought out precisely in our relations with technologies. We are bodies in technologies.
—Don Ihde, *Bodies in Technologies*

Keywords: cinematic experience, immersiveness, cinematic self, perspective- taking, OBE

Introduction

What we generally suppose in this chapter is that film experience is, from its very beginning, thought of as an experience rooted in our brain-body system, aiming to *externalize* our cognitive processes, feelings, emotions, and motor behavior (Münsterberg 1916; Freeburg 1918) and that certain modalities of perspective taking, switching perspectives, dynamic emotional empathetic behaviour with other bodies are part of a general cinematic experience that might be also present especially in technically induced or manipulated Out-of-Body Experiences (OBE´s).

Therefore we propose an explorative journey in the territory of the cinematic (bodily) self in relation to (1) Out-of-body experiences and (2) different modes of "(inter-) embodied" cinematic experience in relation to perspective taking and immersion. This will be persued in a theoretical "bifocal vision" of *plastic and polymorphic bodies and selves in technologies*.

Therefore we explore the notion of self in OBE experiences in relation to the natural medium of the human body that sometime has been iconically related to the

cinematic apparatus. We give a short overview on a debate on film rooted in the idea that this relatively new form of art is to some extent rooted in a "neurological conception of modernity" (Singer 1995) and that its appeal is basically related to what Georg Simmel would call an intensification of the nervous stimulation resulting from the swift and uninterrupted change of outer and inner stimuli. During the first decades of XX century, several physicians start studying movies because they guess that something physical is happening to a viewer *whose mental and bodily faculties are altered cinematographically*. Such a position is basically shared by the tradition of French filmology – gathered around the "Revue International de Filmologie" –, by people like Erich Feldmann (1953), who quite clearly talked of a *bilocated* mind (cf: Furlanetto et al 2013) during film-watching, describing film experience like formed by two ellipses, one related to the real world and the other to the fiction with a small area in between intersection enacted by the viewer still on his seat in the darkness. Henri Wallon (1953) then linked this discussion to viewers' motor behavior and mirror mechanisms.

 Finally we will make a first explorative journey into the concept of "*Cinematic experience as a temporally limited immersive self-loss in the other*" or inside –the-other-(body) experience a) suspended in its status nascendi b) a flight interrupted when the lights are switched on, or c) a morphing that regresses when we "drop out of the game".

1. Cinematic experience
According to a consolidate tradition within film studies, film experience challenges our spatio-temporal cognition and implies an alteration of viewers' self and body and thus their embodied self, by using film style and editing to trigger in some cases something very similar to Out-of-Body-Experiences (OBEs), switching the spectator's viewpoint and her emotional and empathetic identification or better: his or her immersivenes into the *film's body*.

With such a term – which has been used both in studies within film phenomenology and in those within cinematic subjectivity (Sobchack 1992, 2004; Barker 2009; Chateau 2011) – we refer to the layers of resonance represented by the bodies depicted on the screen, but also to the gestuality simulated by the complex experience with the movie itself, that is to the ability of *film techniques of conveying a peculiar form of subjectivity*, including the immersiveness of the self of the spectator in the cinematic experience. Let´s recall editing style that conveys such a form of cinematic (inter-) subjectivity as proposed by Walter Murch (2001):
According to the American film-editor's *Rules of Six* an "ideal" editing style conveys six criteria at once and in a certain hierarchy of importance: 1) emotional entanglement with the emotion of the cinematic moment ("51%") 2) Advancement

of the story ("22%") 3) Rhythm ("10%") occuring at the rhythmically right moment 4) Attentional Eye-Trace ("7%"), acknowledging the audience's focus of interest at each moment 5) Planarity ("6%") of the screen and 6) the three dimensional continuity of actual space ("4%") where people are in the room and in relation to one another.

Interestingly, one could add the *suspension of feeling one´s somatic body* as one of the self-evident rules of cinematic experience together with the *getting emotionally entangled within* not only the *character of the movie* but *with the whole "body of the film"*: cinematic experience becomes an outcome of a specific embodied technique (Ihde 2002,2010).

▪

If we vary our perspective, we can ask: Could we consider OBE as a *cinematic experience without a screen* in which the proper body image would be projected outside?

Why does perspective matter (Petkova et al 2011)? How do alteration of perspective introduce change in the 1st Person Perspective (1PP) realized by a) alienation (OBE) or b) appropriation (avatar identification) (see: Ganesh et al 2011) as well as by the switch in between 2nd, 1st and 3rd PP in order to better understand (altered) cinematic self experience: The egocentric reference frame, our orthodox 1PP might be not as clearly as it seems our principal perspective we can assume. Beccio et al (2011) calls the First person perspective "egocentric perspective" while *imagining* another opposite perspective of our own would be for her a "disembodied perspective taking", while a second person actually sitting in front of someone would be an "embodied perspective taking". For her, perspective taking needs the presence of another person to function plainly. However, we suppose that by our image-consiousness (Husserl 2006) we are enabled to take the embodied perspective of another embodied person also in his artificial presence on a movie screen (in all its degrees of embodiment). If we talk here of perspective taking we have to clearify that we can distinguish at least the following basic forms:

1*) Visuospatial perspective taking 2) affective perspective taking 3) kinaesthetic perspective taking 4) motivational or volitional perspective taking*.

All four should seen as joint/ coordinated and sometimes segregated as in altered self -experience.

Clearly the doubeling of the somatic /virtual self in autoscopic experiences, in which the virtual body or body image is doubled and the attentional self-location between the constitutional virtual body image (see: Ihde 2002) and the somatic body schema

may switch as in OBE, should be considered different from the orthodox film experience, where a moviegoer sits still in a dark movie theater.

And nonetheless, the viewer's empathetic relations and social perspective taking with the other on the screen, that is from a fixed body position to a virtually mobile one, could lead to something similar to the feeling of displacement of one's own body (as in OBE), or transformation of one's own body (OBT; see Gardner 2013).

Our proposal is connected to what Don Ihde (2002, 2010) designates as "embodied technics" of cinematic experience: our embodied and mediated experience with and through contemporary technologies, in our case mainly cinema (a 2D moving picture in a dark room experience in which our body is quietly sitting on a chair) and derived new forms of cinematic experience and immersiveness (from frontal, stereo-sound to sensouround-sound in *Apocalypse Now* to several screens to being immersed in a 3D atmosphere).

If we start from the position that our self in its dynamic constitution is actually mediated by and through our body and the technologies we experience our bodies though, we can´t favor a position of media-technologies, imaging, digital-compuational or virtual reality and film being responsible for just "disembodying ourselves". Neither could we hold a position that our bodies might be reducible to machinic bodies in which we can utopically "upload" our minds, but with John Ihde we can realistically argue that the somatic human experience of embodied technics is actually embodied or re-embodied in new interactive ways and thus our self experience is transformed through new somatic externalization and reinternalization: *we become a cinematic self in its technically instilled and mediated loops of intersbjective and empathetic embodiments- disembodiments and reembodiments.*

.

What if, the lights never turn back on again in the embodied cinema, or if we turn them on and we were in another place or in another body? Thereby different *degrees of self-loss* and its different modes of the suspension/alteration of the self in filmic (2D, 3D) or virtual reality can be thought of. These ideas are not so far from contemporary research by Visch, Tan, and Molenaar (2010) on film immersion (2D, 3D, VR), and from Don Ihde's proposals on embodied techniques (2010), where he describes the position of the viewer respect to movies and videogames according to three different typologies: embodied, disembodied, avatar. These proposed triads of cinematic experience help us in understanding the different possible modes of the self in altered self-experience and we might include OBEs as a forth mode of the body-image in the sense of a doubled self.

2. The of Out of Body (OBE) Experience and the general mechanisms of perspective and perspective-taking

2.1 OBE´s
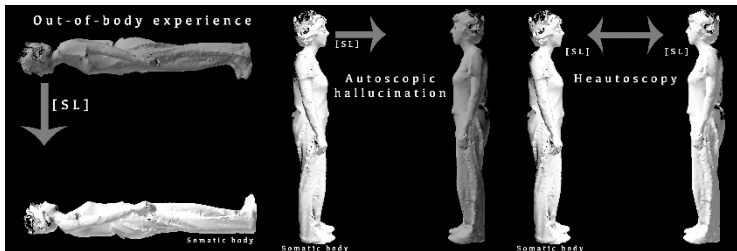
Diagram©2013 Alexander Gerner, In: Gerner (fc 2014)

> This diagram (adopted from: Blanke & Metzinger 2008, 10) shows the dynamics of the attentional self-location [SL] or point of view in autoscopic experience: Three cases of direction of the attentional point of view in *natural cinematic experience without a screen*, either from the hallucinated body towards the somatic body in Out-of-body experience (virtual observer perspective) or from the somatic body towards the hallucinated virtual body (somatic 1P perspective) in Autoscopic hallucination or both ways in Heautoscopy (switching between somatic 1P perspective and virtual observer perspective). The somatic and the virtual body in these three cases of observer perspectives always face each other on the contrary to a "felt presence" from behind (existential feeling perspective of autoscopy), another form of autoscopic experience

OBE is often described as a breakdown of several necessary aspects of bodily self-consciousness (see: Brugger et al. 1997; Blanke et al. 2002; Aspell et al 2012, Gerner fc 2014). Thus Out-of-body experiences challenge and alter our everyday experience of the spatial unity of self and body and the identity of self and body, whereby the body is given as the most complex multisensory "object" in the world (Aspell et al 2012).

Out-of-body experiences are conventionally analyzed as "autoscopic phenomena" that- as Bolognini et al (2012) put it- "refer to complex experiences involving the illusory reduplication of one's own body". According to Mohr&Blanke (2005) autoscopic phenomena (AP) are rare, illusory visual experiences during which the subject e.g. has the impression of seeing a second own body in extrapersonal space. AP - in their view- consist in "out-of-body experience, autoscopic hallucination, and heautoscopy".

"The main forms of doubles are the visual own-body reduplications: autoscopic hallucination (AH), heautoscopy (HAS), and out-of-body experience (OBE) as well

as the rarer forms including polyopic heautoscopy and inner heautoscopy. These are referred to here as visual doubles. Other own body reduplications include feeling of a presence (sensorimotor doubles), hearing of a presence (auditory doubles), and negative heautoscopy (negative doubles)." (Blanke[2] et al 2008, 451)

For Brugger et al (2006) Heautoscopy is the encounter with one's double (the reduplication of a single body and self and thus a breakdown of integrative processes that let me identify with my body or my self), in the sense of a multimodal illusory reduplication of one's own body and self. The phenomenon of *polyopic heautoscopy* (a multiplication of body and self) according to Brugger et al (2006) "points to the multiple mappings of the body, whose disintegration may give rise to the illusory experience of multiple selves."

Autoscopic phenomena deal with viewpoint changes, illusionary self-identification, altered or abnormal self-location(s) and changes in the first- person perspective (cf. Blanke 2012). They can be ideopathic, self induced or induced by non-invasive technological aid (Blanke & Metzinger 2009) using for instance video (Lenggenhager et al 2007), virtual reality (Ehrsson 2007) or robotic devices (Ionta 2011), inducing changes in the self-location, self-identification or first person perspective in healthy subjects; Moreover, recent research has not only described phenomenologically these strange doubeling, mirrowing or shadowing phenomena of a "disrupted" self (Mishara 2010) but has shown as well that invasive manipulation of the brain can even induce a "illusory shadow person" (Arzy et al 2006) by artificial brain stimulation.

Phenomenologically OBEs can be characterized by three elements:
(1) The impression that the self is localized outside one's body. This can mean a feeling of disembodiment or the impression of a virtual body phenomon- e.g. a doubling phenomenon of the body or an extrasomatic attentional self-location. A person experiencing an OBE in this sense would be absorbed by the experience of an (supplementary) allocentric self, besides the somatic self;
(2) The experience of seeing the world from an extracorporeal and elevated or even lowered first-person perspective, and thus changing and doubling the proper point of view: what we would generally describe as a dynamics of points of view and perspectives and their perspective switches. This dynamics in autoscopic and heteroscopic perspective-taking is of course not at all exclusive to OBEs by asking: "Could there be spatial situations in which people spontaneously adopt another's perspective rather than their own, even when not communicating to other person?" (Tversky and Hard 2009)
(3) Experience in OBEs are mostly accompanied with the non-unitary notion of the self as *doubeling or splitting of one's own body image*: the impression of seeing one's own body from alternated allocentered perspectives in relation to the somatic body.

OBEs in healthy persons are often related to sleep-paralysis, REM sleep (see Zippel 2014, this volume), lucid dreaming, trance and traumatic experience. Thus OBEs challenge and alter our "everyday" or expected experience of the spatial and temporal unity of self and body and the "identity" of global self and body in its multimodal constitution between senses, attentional self-location, action and imagination.

Autoscopic experiences in our view question:
1) the coherence of the body as one and only and exclusively mine: the "naturally" assumed somatic self-"identification" (i.e. the degree(!) to which humans identify with their own bodies: "What I experience as my body" (Blanke and Metzinger 2009)"
2) the stability and "permanence" (Merleau Ponty) of my self-experience (visual or proprioceptive feeling or vestibular) in relation to my body and to self-location (i.e., the volume in space: "Where I experience to be")
3) the relation between my experiential point of view of my self and the point of view of the somatic body: first-person perspective (i.e. , the directedness of conscious experience: "From which vantage point I experience the world")

Autoscopic experiences can help elucidate body ownership[3]. According to Tsakiris (2011) body ownership "gives somatosensory signals a special phenomenal quality, and it is fundamental to self-consciousness: the relation between my body and "me" differs from both the **relation between my body and other people's bodies** and the **relation between external objects and me**" (Tsakiri 2011, 181).

We argue that on the one hand we can partly describe autoscopic experiences such as OBE or heautoscopy as cinematic experiences of a *body without a screen*. On the other hand we can observe that the viewer's film experience in a movie theater resembles *an extended sense of "out-of-body" experience* that diverges in the sense of being more of an extended "inside-of-the-other-(body)-experience" What differences do we feel when the subjective perspective puts us in the situation of emphatizing with the murderer, someone of another sex etc.? Could we actuallz put this switch of perspective to test with neuroscientific methods?

What we want to draw attention on by describing film experience as an out-of-body experience is something detectable as well in an experiment by Slater et al (2009), which shows strong evidence for the *plasticity of the body image*. Hereby male participants in a virtual reality situation even perceived the avatar of a young girl as their own body (Slater et al 2009). This brings us up to the point that these drifts of perceptions, imaginations and affects of the perspective spatially and psychologically are important to be studied by sitations of cinematic experience including 2D cinema, 3D virtual reality and avatar studies that will help us to

understand the drifting attribution of a certain body image to me or to another person or character depending on the perspective we are taking on or shifting away from. In the analysis of film experience including virtual reality and avatar studies, nonetheless we have to stay conscious about the fact that such experiences imply completely different positions of the viewer and this is crucial as we talk of dynamics of perspective taking, besides attentional self-locations (SL). The very goal of our proposal is to problematize some relevant aspects of our behavior during film watching, when we are challenged to move ourselves to a virtual environment populated by virtual agents with whom we can interact through the peculiar behavior of the camera.

Why OBEs could be interesting to be treated as cinematic experience in itself and why OBEs can be called "cinematic experience without a screen"? We need to ask further: why we not only have a) a sense of self-agency, the "prereflective experience that I am the one who is causing or generating a movement or action or thought" (Gallagher 2012) b) a sense of "self-ownership" the pre-reflexive experience that I am the one who experiences, but also the c) Perspective switches, the possibility to consciously or mostly prereflectivly switch perspectives in relation to the body-location and different extension levels of embodiment and its relation to others: my body, the other body on the screen and the film as body, the general body of the cinematic experience (that can have interbodily components) and then come back to our core body self after the screen is dark again and the lights are switched on.

How can we temporarily identify and lose ourselves/ our feeling of embodiment in an absorbing or immersive experience and then come back to a self-localisation of our own core body? These questions bring us to the topic of perspective taking.

2.2 Perspective –Taking
For Thomas Fuchs (2012, 2013) every encounter is based on the capacity to switch between your own embodied perspective and the perspective of the other and at the same time to distinguish both perspectives that is to assert yourself in front of the other. Hereby Fuchs (2013) quotes an interesting point of Blankenburg that we will take up here: That is, one has to able to integrate the egocentric and the allocentric perspective without loosing one´s own bodily center *permanently*.

"Or as Blankenburg 1965 says this to the point: Every taking over of perspective implies already a potential self-loss that however is suspended in its *status nascendi*." (Fuchs 2012b).

According to Fuchs schizophrenia is best analysed as the alienation of its own body or as a "disembodiment" (Stanghellini 2004, Fox 2005, Fuchs and Schlimme 2009).

This refers to the concepts embodied subjectivity (Embodiment), as currently used in the cognitive sciences (Varela et al. 1991, Gallagher 2005, Thompson 2007, Fuchs 2012c). Disturbances of embodiment may be classified according to Fuchs & Schlimme in two fundamental categories:"(1) as primarily affecting the subject body or prereflective embodied sense of self; such is the case, for example, in schizophrenia or depression, or (2) as being more related to the bodyimage or explicit body awareness. These include, for example, body dysmorphic disorder, hypochondriasis, somatoform disorders or eating disorders such as anorexia nervosa" Fuchs & Schlimme 2009, 571, and we could add the second type is important in cinematic experience.

Schizophrenia thus includes, according to Fuchs & Schlimme the weakening of the basic sense of self. This means a disruption of implicit bodily functioning and a disconnection from the intercorporality with others: "As a result of this disembodiment, the prereflective, practical immersion of the self in the world is lost" (Fuchs & Schlimme 2009). We could call this the *natural media immersion of the bodily self in the world* in difference to artificial technologically induced immersion as by cinematic experience or virtual reality environments. For Fuchs there is a foundational role of second person interactions for the development of social perspectives (Fuchs 2012). He argues that embodied second person interactions are not only an enabling, but also the constitutive condition for the development of an explicit first and third person perspective. This elevates the possibility of OBE´s and different kinds of perspectives and perspective taking to fundamental importance not only in social cognition but as well in the proper idea of a cinematic self and its technologically mediated existence, one of it foundational part is the switch of perspective.

Perspective taking is a developed "natural" technology of a lived human body. Perspective taking can mean the embedded ability to follow the eye gaze of the other and get empathically entangled and experience the other´s complex perspective (visual, empathic-affective, motoric etc.) and in a metaperspective describe the presentation of a scene, object, event or atmosphere from different situated vantage points in the world. This ability of perspective and orientation isdeeplz related to the possibility of switch of perspective and involving the attention to another self or ourself from another point of view, but does persepctive always imply the feeling of being grounded in a somatic body?

A) Perspective taking involves the perspective *from x* such as a situated subjective either extended spatio- (somato) corporal self location or a virtual or imagined selflocation as in autoscopic experiences and B) the perspective/ angle *towards y*, for instance the objectifying one's own view of the object, and anticipating that moving to another situated vantage point. C) These changes of vantage points can

result in specific changes in presentation of an event, object, scene or atmosphere, such as a feeling towards someone, one´s own body-image, or self-concept. D) The differentiation and self/other coordination of viewpoints is an important feature of multiple perspective-takings as well as E) the constancy/stability/permanancy of perspectives towards a scene/atmosphere/object event or world during taking on a perspective and accordingly the point of view of the self. We can as well distinguish perspective-taking in the following modalities:

a) Spontaneous, involuntary or effortless Perspective-Taking (without volition or intention of taking on the perspective of oneself, the you or the other), and that is what Fuchs (2012b) calls *implicit* 1PP, 2PP or 3PP and we can also call *transparent* perspectives for the one taking on the corresponding perspective. For example spontaneously feeling oneself in someone else´s shoes, but also effordless attentional switches between the 2nd, 1st and 3rd PP.

b) Non-spontaneous, voluntary, effordfull, self-conscious or explicit (Fuchs 2012b) perspective taking (1PP, 2PP and 3PP) goes far beyond the feeling of empathy; it involves for example active effordful figuring out what others feel, perceive and think. The effordful acquisition of a perspective – as in an actors work on a role) is based on on many of the brains executive functions. It may require inhibitory control over our thought and feelings to consider the perspectives of others, and thus in a metacognitive reflection to consider the possibility of someone else besides our own thinking, cognitive flexibility to be able to see and interpret a situation in different ways can be seen.

OBE and other forms of autoscopic experiences are first of all unorthodox forms of spectatorship and they entail different forms of immersion which have to be researched on and considered within a dialectic between alienation (of self from own body, loss of somatic self) and appropriation (of avatar body among other possibilities), according – for instance – to Ganesh and colleague's works on the human brain and the virtual (2011).

Hereby we can refer to cinematic experiences as immersive stretegies in modulating and enhancing the possibility of taking on involuntarily and later also reflectively the perspective of a certain "fictional" point of view of the other.

2.3 OBE, the self and cinematic experience

So far we have been wondering whether it would be possible to compare OBEs to a sort of cinematic experience without a screen and to detect something similar to OBEs in traditional cinematic experience, by mainly focusing on the dialectic between different forms of perspective taking. As we have demonstrated so far, OBEs could be thought of as a weird form of spectatorship, which implies a

dissociation of the viewer from her own doubled (visible and somatic) body and the observation of it from an imaginary location in the extracorporeal space. Alienation and self-loss are constitutive elements of an experience that we would describe as an illusory, phantasmagoric and fictious experience. The body is bilocated and we see our own body from another dimension and position, and although this position is an illusory one, we experience our illusory second body as our own body and our real body as an image of it. This is the reason why OBE represents a very particular type of immersion of the self: the OBEer is immersed in an imaginary space, perceived within an illusory sensory-motor perspective, responding to very particular stimuli, which holds as well for cinematic experience in general.

If a non-OBEer should try to imagine an OBE, he or she would inevitably end up to refer to his or her experiences as spectator, he or she would imagine of occupying an impossible position, usually above the real body, and they should imagine an absurd point-of-view moving freely inside their room, and perhaps flying out of the window, and finally coming back to rejoin the real body. In short, they would imagine a cinematic experience: the camera is able to place us in unfamiliar positions, it can provide us an absurd point-of-view, it can wander across our room, and – as every moviegoer knows – it can make us fly with the means of the film's body. Moreover, film editing is able to regulate our attention, to elicit our emotion, to link very different places and environments in perfect continuity and transparency, and then – at the end of the movie – to allow us to re-enter our somatic body, we had forgotten about.

At the very beginning of cinema, when film theory was more a physiological matter than a cultural one, we find several writings in which film experience is described as an alteration of the human self, or as a loss of self-location and self-identification. We could recall here many writings from the first years of film history, to emphasize the sensorial novelty displayed by the movies and to demonstrate how impressive the new experience was from an affective point of view. Let us quote just two exemplar passages written in 1896 and 1919 by Maksim Gorki and Urban Gad respectively. Gorki offers a disturbing description of film experience:

"Cette vie grise et silencieuse finit par vous troubler et vous oppresser, vous avez l'impression qu'elle contient comme un avertissement, dont la signification vous échappe, mais qui est lugubre, et étreint votre cœur d'angoisse. Vous oubliez peu à peu où vous êtes, d'étranges images surgissent dans votre tête, votre conscience semble s'obscurcir, se perturber…" (Gorki 1896).

Gad's description of film experience is quite surprising and provides elements that make the comparison with the OBE's spectatorship even easier: "Les hommes dans leur grande masse naïve doivent se retrouver dans le film comme dans un miroir –

un miroir, il est vrai, suspendu en hauteur et qui contraint à lever les yeux." (Gad 1919).

The alteration of the viewer's self seems to be implied in these first descriptions of cinematic experience. On the one hand, Gorki talks of the distress of such an oppressive experience, emphasizing the alteration and perturbation of the viewer's consciousness, while on the other Gad describes the screen as a mirror placed above the viewer and capable to double her position and to put herself in a totally new dimension.

At the beginning of 20th century it is a common and widely shared idea that film experience should have been conceived as an altered state of consciousness, something between daydreams and a mysterious form of hypnosis. There is a very telling short story, published in 1907 by the popular Italian writer Edmondo De Amicis, through which we can perfectly grasp such a feeling about cinema. It is the story of a middle class Italian man, alone in his house since his wife and daughters went to theater. He sits on an armchair and thinks of his life. Gradually he starts having a weird sensation, like one who leaves his own body and floats through the room, and then along hills, mountains and valleys. Suddenly he thinks of a newsagent and he sees him quite clearly, and what is more interesting and scaring he feels the newsagent as a secondary self, having the impression that the newsagent's face has overlapped his own. If we should find a term to describe, nowadays, the experience narrated by De Amicis, maybe we would choose OBE. Nonetheless, in 1907, De Amicis chose another term: cinema. The title of the story is "Cinematografo cerebrale" ("cerebral cinema"), in which film experience and the brain are connected to emphasize the alteration of human cognition at the movies.

Although someone might sell off these judgments as naives and too strongly connected to a not yet well developed idea of film experience, it is a matter of fact that the relationship between the movie and the human mind is at the heart of what is considered to be one of the most insightful book of early film theory: Hugo Münsterberg's 1916 *The Photoplay: A Psychological Study* (2002). This distinguished Harvard psychologist – who moved from Germany to Massachussets at the end of XIX century on the invitation of William James – noticed that cinema could not be understood without referring to the effect it has on our brain-body system. Cinema, in other words, externalizes our affective-cognitive processes, feelings, emotions, and even motor behavior, by means of stylistic techniques and innovative narrative solutions: a flashback would be a kind of externalization/representation of memory, while a close-up would be the same for attention. Münsterberg's book, as we can read it nowadays, would represent the most clear reflection of a common and shared feelings about modernity as a form of intensification of the nervous stimulation, resulting from the swift and uninterrupted change of outer and inner

stimuli, to borrow Georg Simmel's description of the new metropolis at the very beginning of XX century. According to scholars like Ben Singer (1995), or more recently Christof Türcke (2002), film culture would be part of a new form of modernism based on the hyperstimulus and on a new form of *affective interaction*, and it would precisely grounded on a "neurological conception of modernity", as Singer calls it, *including the distraction from one´s own somatic body*. In other words, there is a new form of self-technique capable of shaping up our imaginary by affecting our mind and body and by challenging the viewer's spatio-temporal cognition in a totally new and impressive manner. Recalling the idea by Daney about the dialectic between two spaces and two kinds of vision – basically the same form of dialectic suggested by Feldmann and by two film phenomenologists like Sobchack and Voss –, we could observe how it is the film style which pursues and at the same time regulate this spatial negotiation, and how our multilayered and multimodal film cognition depends on the success of such a negotiation.

To better understand such a story, we should get back to the long neglected season of film-physiology, that is a period – from the beginning of 20th century to the 1920s – during which many physician start working on film in order to evaluate the impact of the new medium on the human brain (Guerra 2013) and cognitive make-up such as the "attentional self" (Gerner upcoming). Both in Europe and in US we have important studies on this way, like those of the French physician Edouard Toulouse, who, for instance, was convinced that the impression of reality largely depended on the viewer's motor simulation of the events depicted on the screen – some passages in Toulouse's works seem to anticipate the research on embodied simulation promoted after the discovery of mirror neurons (Toulouse 2010; Gallese and Guerra 2012) and its empathy in relation to a film character or a virtual object (Fuchs 2014).

Among physicans, we could mention the case of the Italian neuropsychiatrist Giuseppe D'Abundo, who wrote a paper in 1911 entitled "Sopra alcuni particolari effetti delle projezioni cinematografiche sui nevrotici" (Concerning the effect of film viewing on neurotic individuals). His idea was that a movie can determine states of psychic instability in patients like neurotics, hysterics, or paranoids. According to him, the responsible of this state is not the film plot, but "the rapid and vibratory movement of the cinematic action" (D'Abundo 1911: 434), which is able to transport the viewer in another dimension, giving him the impression to be at the same time here and there. He concludes saying that film projections should be considered dangerous for many categories of subjects.

To us it is important to rethink such an experimental background, since we know how crucial it has been for film theories like those proposed by Sergej Eisenstein (we know today how close to Aleksandr Lurija he was) in the 1920s and 1930s –

think of his idea of film editing as a form of "ek-stasis" – or by Walter Benjamin in the 1930s. An affective and sensory approach to film experience was also implied in Antonin Artaud's few writings on film, where he said that cinema "acts directly on the grey matter of the brain" (Artaud 1972, 166). It is not by accident if, in the 1950s, the new school of French filmology will restart from here, trying to shape up a field of research on film where psychology and anthropology would converge.

In 1956, the German philosopher Erich Feldmann wrote a brilliant article on the "Revue Internationale de Filmologie", in which he claimed that film experience basically depends on the viewer's ability to move from a real environment – that she occupies in the dark movie theater – to an imaginary dimension – that provided by the world depicted on the screen –, feeling herself localized outside her body. Feldmann stresses that such an experience, during film watching, is elicited without what he calls "modifications psycho-physiologiques". These are his words: "Le film réclame du spectateur ce qui semble à première vue impossible : se transporter, sans l'aide d'excitants, de stupéfiants, ni de modifications psycho-physiologiques engendrées par la seule projection lumineuse, dans une situation irréelle, tout en demeurant dans la situation réelle de la salle un être éveillé qui croit à la réalité du film qui l'absorbe." (Feldmann 1956, 84). After few lines, he adds that "l'individu voit surgir dans le cadre de sa vie une combinaison de conditions qui modifient son attitude habituelle et qui demandent une accomodation." Feldmann seems to suggest that this "accomodation" would need a kind of ability that the viewer should have in order to enjoy the movie. His assumption implies not only a change of perspective and a doubling of the viewer's presence-to-a-world, but also an alteration of the viewer's consciousness, as if she would need a cinematic consciousness provided by film techniques to enter the fictitious world of film. Without entering a debate on what we mean when we talk of cinematic consciousness (see Morin 2005, who was strongly influenced and inspired by French filmologists, Pepperell and Punt 2006, McGinn 2005), we could borrow Serge Daney's theory (1993) of film viewing and describe it as a form of alternation between a "vision bloquée" (meaning body centered, situated), and a "vision liberée" (meaning disembodied, experienced in an extrapersonal space).

We could as well recall some of the theories according to which cinema implies an alteration of the self based on a doubled spectator capable of living simultaneously in two different environments- inside a world which is unreal or all too hyperreal. Such positions are still widely shared if we think of how contemporary film phenomenology basically oscillates between the idea that the reciprocity between the viewer and the screen would originate a strange subject to be denominated "cinesthetic subject" (Sobchack 1992), and the idea that the viewer's body, resonating with the events on screen, would loan a three-dimensional body to the

screen, making the viewer nothing less than a "surrogate body" for the screen (Voss 2011), an idea amplified by the independence view of a doubled self, or artificial extention of "secondary persons"(Bainbridge 2014).

Also, more recent theories of cinematic subjectivity (Chateau 2011) seem to wonder how and whether the movie can be endowed with subjectivity, at least a simulated ot enacted form of intersubjectivity assured by the degree of immersion and self-loss of the viewer into the movie and its "extended empathy"[4] (Fuchs 2014) and immersion towards the characters represented on the screen (without their real bodies being present).

▪

For instance seeing Pedro Costa´s film "Juventude em Marcha" for the first time in the movie theatre in Lisbon and noting a switch/twist or crack in the way of perceiving while empathetically taking on perspective and feeling inside the shoes of the main character *Ventura* getting disoriented in a fragil world of his cultural survival. The anti-hero Ventura that already lost the one he loved and is disoriented in the cleanness of social housing that have windows and doors that close but don´t let the common life or the community occur, shows what could be formulated with Peter Handke as the loss of the open image, or as we could call it the whitening out of the image as a bleeching out a stain and belonging to a life world: The loss of an image of the world ("Der Bildverlust" Handke 2002) is as well a *self loss of existential feeling of belonging* to a fundamental self-world-image at the time of its loss. In the movie *Colossal Youth* [original: Juventude em Marcha] (2006) of Pedro Costa something of a life (Ventura) and culture (capverdian) is milled. This empathetic world-image that is installed in the viewer is proposed to open up the viewer´s self in the sense of an "existential feeling of being" (Ratcliffe 2008). Ventura becomes you and me, we are affected by him and become part of the film´s body: we become the one that lost a friend a woman, a life a home a culture, or simply a human belonging to this world. The film of Pedro Costa describes a bleaching of the image that simultaneously opens and makes the viewer empathetically belong to film's body, a property of pain, a memory, a march against the loss of an image and a body that holds, resists, that is valid. The film starts with the emptying of the interieur of a house by defenestration, a concrete (and political) act of ejectment of an old furniture (an interior- or *a former self is defenestrated*, through a window frame that never had a window in the first place, a place without the protection against the cold, but as well a place of social possibility of permanent openings for instance if there is need to call someone for help from the street or for company and the one who is called answering right away. In the film Ventura is obliged to accept social housing, all white walls and double windows, which is shown as the opposite as to the inicial fenestration of old furniture. The loss of viscerally felt entanglement with

the living by institutionalized, whitened and cleaned atmospheres that are aseptic and disorienting living spaces in which the doors close on their own and the window are double and soundproof, that no life sign can be detected by a call of a friend from the street: what this movie shows is a double autoscopic self-closure, a losing myself immersively in the character Ventura and the body of *Collosal Youth*, and the expression of a cultural self-loss in the narrative of the movie: a double cinematic OBE experience.

Let us therefore explore the concept of "*Cinematic experience as a temporally limited immersive self-loss in the other*". What could be the reason why players prefer to see their own avatar character entirely from an overview perspective from above (OBE perspective) and not from within a direct and absorbing 1st PP? How can we relate this back to the cinematic experience as a proto-OBE? For this we have to have a look at immersive experience.

3. Immersive experience

It seems phenomenologically more correct to not simply equal OBEs with Avatar experience- as technologic self- extension of self concept in which a distance towards the technical virtual double can be reinstalled at any given moment. However, we have to ask the following question in relation to immersive experience: *Can we be distracted from our somatic body and immersed into a fictional or cinematic body- the body I feel and am affected by as long as the cinematic experience lasts?*

Let's start to tackle this question by looking back on a phenomenological account of the basic bodily self: The basic bodily self before having a referable 1st PP or any kind of self-knowledge or self-concept in relation to contents or objects is characterized by a (pre-predicative) bodily affective self feeling, of an immediate pre-reflective self-presence. The question is if this self-feeling is already attributed to my somatic body and if this is always a conscious, or in the sense of Thomas Fuchs, *explicit* 1PP?

The self in this first ipseity or immanent account of radical self-affection, can also be described as an existential feeling of being (Ratcliffe 2008) –a self, an worldly awareness of being that is not already an object, an emotion or a mere disposition of something or itself. *I feel through bodily appearing and by being affected, but this constant floating feeling of being makes me myself.* This does not – however- mean that "I" feel exceptionally from a pure 1st PP and always consciously a "what", that is "my body". This in turn means my body through which I feel and am by what I am somatically and physicologically affected with is already part of something else than me- an internal other- and thus is experienced in different ways, that is in *Altered Self-Experiencs*. Ratcliffe[5] defines existential feeling therefore as a *relational*

bodily awareness in which the body is a seen as a kind of a *natural medium* of this existential feeling "through which" something is experienced.

Therefore Irene Mittelberg (2013), when referring to bodily movements as gestures calls these already natural media of an "exbodied mind". How strong the bond towards this natural medium – the bodily self- is can be the degree of natural immersedness inside our bodies. This feeling of being affected for Fuchs is inherent in all conscious processes and thus is able to be differentiated into a a) *primarily bodily self* b) an *ecological self* and c) a *social self*. All of these are seen under the concept of *ipsiety* or the experience of self-affection. How can we see now the immersivenes inside our body being altered by cinematic experience and its different technically produced modes by film style and editing and the atmosphere of the cinema theatre?

For a short while we start by the feeling of absorption by the cinematic dark room, leaving the conscious nexus with our somatic body by being affected by entering this cinema world. In this darkness your senses go to the light, your body is fixed on a seat, your bodily self is still, stillness as if being in a temporary tank of sensory deprevation, in which your gaze is channeled by image sound and rhythm as film editing and film style: your gaze onto the screen, and reducing the complexity about the self awareness "about" your own body by the immersion with other bodies. From the start of cinematic experience people wanted to see bodies like ours moving in space, people, animals running leaving the factory, the first kiss in film -"that´s how people leave the factory, that's how people kiss", "that's how people walk", and that´s how people escape from the train coming towards us", the scientific instinct of seeing, and feeling and running with the others on the screen as a necessity to see ourselves (as explored in the thought provoking Elias Canetti's theatre play "The comedy of vanity" (Canetti 1981) in which the only reason for a revolution is the *sensory deprivation of seeing ourselves in other* people, thus the people loosing themselves by not seeing themselves in the others, in their *own double image of a self-other*; we still want to see ourselves autoscopically in arts in the movies in a TV showbut also to immerse our body within the other– as we have a technique of social perspective taking and joint intentionality (Tomasello 2014): or let´s better say: we constitutionally need to see ourselves in order for me to be able to *form, maintain or alter myself*. We could call this the natural media immersion of the bodily self in the other. This natural immersiveness of the body can be technologically enhanced and become an artificial technologically induced immersion by external media, both natural and artificial imersiveness are part of cinematic experiences that alter the experience of self.

For Fuchs (2012, 2012b) there is a foundational role of second person interactions for the development of these immersions in the other that he calls social

perspectives. He argues that embodied second person interactions are not only an enabling, but also the constitutive condition for the development of an explicit first and third person perspective. Maybe cinematic experience can give us back a basic –almost childlike- belief in the world as if for the first time looking at the other; cinematic experience then would be another form of reembodied self experience with the other that is me on the screen: that is what Deleuze talks about in his cinema books.

When we see a bad movie, we "drop out" of the state of immersion, or bounce back to our seat and feel the uncomfortable position our body is squeezesd in, we start moving on our seat, or we look at details that are unimportant – in bad theatre we start noticing the lighting equipement- we let our attention stray in the image instead of being immersed and transported, entertained or even transformed. Instead of loosing our self we loose track of the narrative the plot, the character the situation, the film´s body, and we start thinking about something else, ruminating. The cinematic guidance of a psychogenic flight stops or crashes we are on land of our own reality instead of being in the air and out and away with the film.

.

The cinematic phenomenon of immersion into a new cinematic body has similarity to the technically induced »body swap illusion« (Petkova/Ehrson 2008), in which manipulation of the visual perspective and correlated multisensory information (passive tactile information) from another person's body is sufficient to create the illusion of inhabiting a non-somatic body by means of a continous match between visual and somatosensory information about the state or location of the new humanoid body and the adoption of a 1PP moving the person´s perceived center of awareness from the somatic body to another artificial body.

We can describe immersion as a voluntarily induced bodily or embodied somatic self-loss experience in the other – on the contrary to involuntary or pathologic self absorption. This is exactly what aesthetic cinematic experience, in the large sense, seems about: "a new method to move a person's perceived centre of awareness from one body to another". We become part of the body of the movie, we are carried along by hands of narrative and the face of empathy and rhythm of editing style and all the characters affects, performatively presenting us with our actually unfolded complexity, in which by taking on the perspective of the other. We discover new knowledge about us and the world by the unfolding of the characters on the screen that become my other self, the story relates to my life, myself and other´s around me: in a cinematic experience the you – I-here relation: the somatic body, the body of the characters and the film´s body, as well as the narrative, mingle into a complex technically induced and experimentally felt cinematic self and with

the next illusionary continuity shot we don´t cut out parts of our body or our position in space or change our point of view etc., but we might drift or switch into the film, by shortening the loops between the filmic double of my body and the body that is the cinematic experience of the film. As Becchio et al put it: "perspective-taking entails an altercentric remapping of space, i.e. remapping of objects and locations coded with reference to the other person's body" (Becchio et al 2011).

But: *can we become a cinematic self who's self-location can be temporarily be shifted out of its body?* The notion of a cinematic -self and its possible mode OBE looking back on us, is still to be discovered: the screen becoming my virtual body, then looking back onto myself, might be described as a "proto" heautoscopy in cinematic experience. The magic latern lights up on the skin of my body- still a metaphoric way of speaking- but, maybe in the future, this might change and become reality.

Experiences as the bodyswapping art/gender project of "The Machine of Being another" hints into this direction by the direct swap of female/male bodies and the looking back onto my body from the viewpoint outside of myself. To avoid nausea the gender pairs are asked to choreograph their movements while looking onto themselves from another outside body of the other.

screenshots (p/w) from: http://www.themachinetobeanother.org/?p=1062

Through the use of Oculus Rift headsets, and first-person cameras, the Gender Swap experiment creates a visual-perceptive enhancement that partner A can see what is being recorded by the visor worn by partner B, and vice versa. A technically mediated cinematic Out-of-Body Experience created by Barcelona-based artists BeAnotherLab, in which both are asked to coordinate their body movements.

Getting lost, drawn into or involved in a plot, a narrative, a character a feature of a body or even a full body, our filmic-somatic loop becomes alive. *Does being dragged into a movie or being absorbed by a virtual cinematic experience mean that I loose myself?* The lights would never turn back on again exactly the same way as time passes in

the screening room, or in my visual field. What if this temporary self-loss experience in an extreme situation would never stop? Blankenburg describes the self-loss in the other in the case of schizophrenia, and that we are always in general already in a state of self-loss by *being with* the other- and we can add by being *in the movie with* each other- however: we can suspend this self-loss in the other by going out of the movie or taking off the VR googles, a fact that schizophrenia patients in their condition can't, they are stuck to their episodic hallucination for the time being. We usually can suspend this self-loss in its status nascendi and reassure that my somatic body is mine and your somatic body is yours. What if our imaginative flight does not land in the same place again? A question has to be answered: Is our bodily self the same/ identical after having had a cinematic experience?

Two other fundamental questions arise at the very beginning of this work in progress:

1) *How can we provide empirical data in order to test the plausibility of these proposals with respect to theories of the self and film phenomenology, and technically enhanced cinematic experience and virtual immersion?*

2) *Would this approach be the right one in order to disentangle the complex articulation of the embodied, disembodied, and re-embodied relation to cinema and, more generally, visual media in an immersed cinematic self?*

Endnotes

[1] Research of Alexander Gerner is supported by a FCT Post-Doc grant: SFRH/BPD/90360/2012

[2] According to Blanke et al (2008) we can distinguish several forms of autoscopic phenomena or illusiory doubles of a bodily self, visual, auditory the sensomotor. He also includes negative heautoscopy, the impossibility to see onself when looked up directly in the mirrow (see: Menninger-Lerchenthal, 1935) also called negative doubles, but we will not include these phenomena in this paper.

[3] Bolognini et al (2011) report the third long-lasting case of autoscopy in a patient with right occipital lesion in their study "Spatial perspective and Coordinate Systems in autoscopy. A Case Report of a "Fantome de Profil" in Occipital Brain Damage". Instead of the commonly reported frontal mirror view (fantôme spéculaire), the patient saw her head and upper trunk laterally in side view (fantôme de profil). The autoscopic image changes in relation to movement of the body. While the body is still just the perfile of the face and the upper trunck are visible, in arm movements also the arm gets visible and in full movement (walking) all correspective body parts get visible in the autoscopic image. This is important for the fact that autoscopy may come in degree and thus we should as well consider immersive degrees of embodiment in one's own or another body.

[4] Empathy towards the "virtual other" in Fuchs account is seen as captured "*notions of (1) phantomization as a media-based simulation of direct reality which undermines the as-if-consciousness, and (2) disembodied communication which shifts the modes of empathy towards the fictional pole at the risk of merely projecting one's own feelings onto the other.*"(Fuchs 2014)

[5] "I argue that most, if not all, bodily feelings are relational- they are seldom, if ever, directed exclusively at the body. Indeed, there are "bodily feelings" that do not involve the body as an object of experience at all. Instead the body manifests itself as that *through which* something else is experienced."(Ratcliffe 2012, 38)

References

Artaud, A. (1972). *Collected Works. Volume Three*. Paule Thévenin (ed.). London: Calder and Boyars

Barker, J. (2009). *The Tactile Eye.Touch and the Cineamtic Experience.* Berkeley and Los Angeles: University of California Press

Bainbridge, W. (2014). "Identity Expansion and Transcendence". In: F. Bataglia and A. Carnevale (Eds.). Reframing the Debate on Human Enhancement. *Humana.Mente 26*, 117-140

Becchio, C. et al. (2011). "In your place: neuropsychological evidence for altercentric remapping in embodied perspective-taking". *SCAN*, doi: 10.1093/scan/nsr083

Blanke, O., Metzinger, T. (2009). „Full-body illusions and minimal phenomenal selfhood." *Trends in Cognitive Science 13*, 7-13

Boellstorff, T. (2011). "Virtuality. Placing the Virtual Body: Avatar, Chora, Cypherg." In: F. Mescia-Lees. (Ed.). *A Companion to the Anthropology of the Body and Embodiment*. Hoboken NJ: Wiley-Blackwell, 504-520

Canetti, E. (1981). *The Comedy of Vanity*. PAJ playscripts: New York

Chateau, D. (ed.) (2011). *Subjectivity: Filmic Representation and the Spectator's Experience.* Amsterdam: Amsterdam University Press

D'Abundo, G. (1911), "Sopra alcuni particolari effetti delle projezioni cinematografiche nei nevrotici", *Rivista Italiana di Neuropatologia, Psichiatria ed Elettroterapia 10*

Daney, S. (1993). *L'exercice a été profitable, Monsieur*. Paris : POL

De Amicis, E. (1995). *Cinematografo cerebrale*. Roma: Salerno Editrice

Dor-Ziderman, Y., et al (2013) "Mindfulness-induced selflessness: a MEG neurophenomenological study." *Frontiers in Human Neuroscience 24*, doi: 10.3389/fnhum.2013.00582

Feldmann, E. (1956), "Considérations sur la situation du spectateur au cinéma", *Revue Internationale de Filmologie*, 26

Fuchs, T. (2014). "The Virtual Other. Empathy in the Age of Virtuality." *Journal of Consciousness Studies 21 (5-6)*, 152-173

___ (2012). "Selbst und Schizophrenie". *DZPhil 60 (6)*, 887-901

___ (2012b). "The phenomenology and development of social perspectives" *Phenom Cogn Sci* (online-first). DOI 10.1007/s11097-012-9267-x

___, Schlimme, J. (2009). „Embodiment and psychopathology: a phenomenological perspective". *Current Opinion in Psychiatry 22*, 570–575

Furlanetto, T. et al (2013). "The bilocated mind: new perspectives on self-localization and self-identification." *Frontiers of Human Neuroscience 7, 71*

Gad, U. (2008 [1919]). *Un miroir qui contraint à lever les yeux*. In: D. Banda & J. Moure (eds.), *Le cinéma naissance d'un art*. Paris: Flammarion, 416-18

Gallagher, S. (2012). Phenomenology. Hampshire: Pellgrave Macmillan

Gallese, V., Guerra, M. (2012). "Embodying Movies: Embodied Simulation and Film Studies", *Cinema: Journal of Philosophy and the Moving Image 3*

Ganesh, S., et al (2011). "How the Human Brain Goes Virtual: Distinct Cortical Regions of the Person-Processing Network Are Involved in Self-Identification with Virtual Agents", *Cerebral Cortex* Sept. 2011; doi: 10.1093/cercor/bhr227

Gardner, M. et al (2013). "Strategy modulates spatial perspective taking: evidence for dissociable embodied and disembodied routes." *Front. Hum. Neurosci.13* August 2013 ; doi: 10.3389/fnhum.2013.00457

Gerner, A (forthcoming, 2014). "Conceptual Personae of the "attentional self"". In: J. Fonseca, J. Gonçalves (eds.). "*Philosophical Perspectives on the Self*", Frankfurt, New York: Peter Lang, 198-233

Gorki, M. (2008{1896}), "Au royaume des ombres", in D. Banda et J. Moure (eds.), *Le cinéma naissance d'un art*. Paris : Flammarion, 48-52

Guerra, M. (2013), "Film-Physiology", *pH. 3*

Handke, P. (2002). *Der Bildverlust oder Durch die Sierra des Gredos*. Frankfurt: Suhrkamp

Husserl, E. (2006). *Phantasy, Image Consciousness and Memory (1898-1925)*. Trans. John B. Brough. (=Husserliana: Edmund Husserl - Collected Works; Book 11) Dordrecht: Springer

Ihde, D. (2002). *Bodies in Technology.* Mineapolis: University of Minnesota

___ (2010). Embodied Technics. Automatic Press / VIP

Maguire, E. et al (1998). "Knowing where and getting there. A Human Navigation Network." Report, Science 280, 921-924

McGinn, C. (2005). *The Power of Movies: How Screen and Mind Interact*. New York: Pantheon Books

Mittelberg, I. (2013). "The Exbodied Mind: Cognitive-Semiotic Principles as Motivating Forces in Gesture." In: C. Müller, et al (Eds.) (2013). *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction*. Handbooks of Linguistics and Communication Science (38.1). Berlin, New York: Mouton de Gruyter, 750-779

Menninger-Lerchenthal, E. (1946). *Der eigene Doppelgänger*. Bern

Morin, E. (2005), *The Cinema, or the Imaginary Man*, Minnesota University Press, Minneapolis

Münsterberg, H. (2002), *Hugo Münsterberg on Film: The Photoplay: a Psychological Study and Other Writings*. Allan Langdale (ed.). New York: Routledge

Murch, W. (2001). *In the Blinck of an Eye*. 2nd rev. edition. Los Angeles: Silman-James Press

Pepperell, R., Punt, M. (eds) (2006), *Screen Consciousness: Cinema, Mind and World*, Rodopi, Amsterdam-New York

Petkova, V., Ehrsson, H. (2008). "If I Were You: Perceptual Illusion of Body Swapping." *PLoS ONE 3(12)*, e3832, doi:10.1371/journal.pone.0003832

Ratcliffe, M. (2008). *Feelings of Being. Phenomenology, Psychiatry and the Sense of Reality.* Oxford: Oxford University Press

___ (2012). "The Phenomenology of Existential Feeling". In: J. Fingerhut, S. Marienberg (Eds.) *The Feeling of Being Alive*. Berlin: de Gruyter, 23-54

Singer, B. (1995). "Modernity, Hyperstimulus, and the Rise of Popular Sensationalism". In: *Cinema and the Invention of Modern Life*. L. Charney & V. Schwartz (eds.). Berkeley and Los Angeles: University of California Press

Sobchack, V. (1992), *The Address of the Eye: A Phenomenology of Film Experience*. Princeton: Princeton University Press

Slater, M. et al (2009). "Inducing Illusory Ownership of a Virtual Body." *Frontiers in Neuroscience 3*, 214–220

Tomasello, M. (2014). *A Natural History of Human Thinking*. Harvard: Harvard University Press

Türcke, C. (2002). *Erregte Gesellschaft. Philosophie der Sensation*. München: C.H.Beck

Tversky, B., Hard, B. (2009). "Embodied and Disembodied Cognition. Spatial Perspective Taking". *Cognition 110*, 124-199

Voss, C. (2011), "Film Experience and the Formation of Illusion: the Spectator as "Surrogate Body" for the Cinema", *Cinema Journal, 50 (4)*

The Altered Self in Dreaming State related to the Out-Of-Body Experiences. A Phenomenological Account interfacing with Neuroscience

.

Nicola Zippel

Department of Philosophy, Sapienza University of Rome, Rome, Italy

.

Abstract. *The dream state represents a very peculiar condition of bodily experience: while during Non-rapid eye movement (NREM) sleep one experiences a muscular-skeletal relaxation, in Rapid eye movement (REM) sleep one undergoes the paralysis of her own body, coupled with a vivid mental process triggered by the ponto-geniculo-occipital (PGO) waves. During this phase of the sleep, dreaming takes the shape of a bizarre narration that, despite its incongruity, uncertainty, and discontinuity, inserts itself in the network of neural connections already in place, and for this reason it is another way of performing personal emotions, thoughts, and associations. Precisely by the inability to move, the embodied self experiences such a rich mental performance as an alienating dimension of his/herself, since the body becomes an inactive though often invoked part of the dream (e.g. one tries to escape or to jump in a dream without succeeding in doing it). On the other hand, being the motionless counterpart of the frantic activity of the dreaming mind, the body, as it were, is outside the dream events.*

This complex neurophysiological situation is extremely interesting from a phenomenological perspective, for it offers the opportunity to consider the condition of the embodied subject as regards the unaware mental process of dreaming. As a general goal, the paper aims at showing the basic features of the relationship between the dreaming self and his/her own body in the light of the most significant empirical dream's research findings.

The analysis will be based on some common marks of dreaming state highlighted by both phenomenology and neuroscience such as engrossedness, entanglement, self-splitting, and transformed relation between the subject and his/her space/time. More in particular, the paper will deal with the very peculiar state of the out-of-body-experiences, which will be considered in conjunction with both sleep paralysis and vestibular-motor hallucinations.

Such an unusual condition of dreaming state, involving the experience of seeing his/her body from a location outside the physical body, entails an alteration of the selfhood with regard to his/her relationship with one's own bodily features and spatial position.

The paper will propose a phenomenological description of dreaming in general and of OBE in particular, aimed at working out a conceptual framework able to explain the essential structures of mental processes involved in the dream state. As a conclusion, I will stress the relevance of the notion of the self for the proposed approach and suggest a possible cooperation among phenomenology, neuroscience, and psychology in dream research.

Keywords – Dreaming Self, Out-Of-Body Experience, Sleep Paralysis, Bodily Selfhood, Phenomenology

"Either the well was very deep, or she fell very slowly,

for she had plenty of time as she went down to look about her and to wonder

what was going to happen next. […]

'Well!' thought Alice to herself, 'after such a fall as this,

I shall think nothing of tumbling down stairs!

How brave they'll all think me at home! Why, I wouldn't say anything about it,

even if I fell off the top of the house!' (Which was very likely true.)"

Lewis Carroll, *Alice's Adventures in Wonderland*

1. Neurophysiology of embodied dreaming

Thanks to the most relevant findings of the last 60 years regarding the dreaming experience, it has also been possible to shed light on the physical elements of the dream, about which Freud openly confessed his lack of knowledge[1]. I am especially referring to the discovery of REM (rapid eye movement) sleep in 1953 by Aserinsky and Kleitman and the number of important studies addressing it. Based on these studies, we can explain also the physiological or, more precisely, the *neuro*physiological processes that underpin the very particular state of the body in dreaming.

Before proceeding, we must clarify and better express the still unclear notion of "body in dreaming". As we saw with Freud, it is possible to talk about two kinds of bodies involved in dream: the dreamt body and the dreaming body. In the dream Freud speaks of, we deal with the dreamt body, but in the general discourse on motility presented by Freud, he is talking about the dreaming body. Nonetheless, I would suggest that this distinction has no reason for being. It is clear that we are taking into account two forms of bodies corresponding to the two levels of dreamt content and dreaming mind, wherein we find in the former the dreamt body and we identify the dreaming body as the one of the dreamer. Nevertheless, given the intrinsic interrelation between the two dimensions of what is dreamt and what is

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

dreaming, the motility of the dreamt body is not lacking in the relationship with the motility of his dreaming counterpart, i.e. the body of the dreamer. Expressed simply: the neurophysiological conditions of Freud's dreamt body, glued to the steps, are one and the same of the neurophysiological conditions of Freud's dreaming body, glued to the bed. In this sense, it is possible to talk of "embodied dreaming" as a homogeneous phenomenon.

The close connection between the two bodies, the dreamt one and the dreaming one, can be explained on the basis of the aforementioned research on REM sleep. In their famous article on this issue, Dement and Kleitman observe, that "(d)uring the actual experiments it was seen that rapid eye movement periods often started immediately after a series of body movements and that a body movement often coincided with their termination while the sleeper seemed relatively quiet during the eye movement period itself". The study of the EEG makes clear "the incidence of body movement in terms of the onset and termination of REM periods". The graph proposed by Dement and Kleitman shows "the tendency of body movement to increase up to the beginning of eye movements at which point it decreases sharply and rebounds at the end of the period of eye activity" (Dement and Kleitman 1957, 682).

In his seminal book on the dreaming activity of the brain, Allan Hobson devotes some enlightening pages to the issue of body motility. Referring to the studies by David Hubel and Edward Evarts, Hobson states that "(d)reams are characterized by a sense of continuous movement, and brain neurons concerned with movement fire intensely during REM sleep" (Hobson 1990, 171). This constant "sense" of moving, i.e. the pervasive presence of motility in our dreams, is so strictly linked to the eyes movement characterizing the REM sleep, that "the brain is both seeing and moving in REM sleep" (Hobson 1990, 171)[2]. The neurophysiologic source of this sense of moving can be identified in the ponto-geniculo-occipital (PGO) waves triggering the REM sleep, whereas the pons (P) instigates waves that go to the lateral geniculate body of the thalamus (G) and to the occipital lobes (O).

Hobson bases this description on his hypothesis of a physiological structure enabling the dream process that he worked out starting from the 1970s calling it both the "activation-synthesis hypothesis" (Hobson and McCarley 1971, Hobson 1990) and, enriched with new empirical findings, the "AIM model" (Hobson 2009). The basic assumption of such a hypothesis is the "neurally determined genesis" (Hobson and McCarley 1971, 1346) of dream state that entails the *activation* of the forebrain and, more specifically, of the pontine brain stem. This process, entirely endogenous, involves the discharge of reticular, vestibular, and oculomotor neurons that proceed and interact with a double and *simultaneous* effect of inhibition and excitation that, in turn, *synthetically* integrates diverse sensory, motor and

emotional elements. Precisely such a contradictory process has been examined in depth by Hobson in further studies leading to the renewed hypothesis of "AIM model", whereas the activation (A) is thought along with an "input-output gating" (I), i.e. the process that inhibits both external stimuli and internal motor activity, and "modulation" (M), i.e. the peculiar chemical balance of de-enhanced aminergic neurons and hyper-enhanced cholinergic neurons. This complex physiological course involving both visual and motor areas is the origin of the formation of images in dreams and of the sensation of moving[3].

The source of the sense of moving, as purely physiological, implies that we don't voluntarily move ourselves while dreaming; this means the dreamt body, for instance Freud's dreamt body jumping hastily over the stairway, isn't set in motion from any will of moving, but rather moves according to a neural unwitting activity[4]. Nonetheless, if one carefully looks at this situation, one has to recognize that there is *no body* that is moving. During the REM sleep the real body, i.e. the dreaming one, is totally in rest, absolutely paralyzed; what is moving is indeed the dreamt body, the one jumping over the stairs or, for instance, flying in the air (Schönhammer 2004). Nevertheless, being the dreamt body the product of neural activity, i.e. of an activity entirely occurring within the brain, there is no body, neither real nor unreal, that is moving. This condition is about only what Hobson calls a "curious state of the brain-mind", in which "motor systems act independent of input and output. Since output is blocked by inhibition, the operation of the central motor commands is free of feedback from the consequences of the movement" (Hobson 1990, 171). This means that it is about the movement, i.e. the activity, not of the body, but rather of the brain.

And yet, this movement and the related sense of movement concern the body as well, since the body is not only structurally involved in the process of dreaming, but is *determined* in its own state by the process of dreaming. Though it can be appear paradoxical, both the involvement and the determination by the movement of the brain take the shapes of the immobility of the body. Such a motionless condition of the body has to do with how the sleeping individual relates to the outer world. As Freud already observed, the detachment from the world and the suspension of any activities dealing with worldly commitments that characterize the dream state are linked to the bodily paralysis of the REM sleep. As Hobson explains: "Access to the internally activated brain-mind by input from the outside world has to be excluded in order for sleep – and the illusions of dreaming – to be maintained. This appears to be accomplished in at least two ways. First, active inhibition of the nerves denies stimulus signals of peripheral origin access to the central nervous system; […]. The second mechanism for excluding external sensory signals is competition, or *occlusion*, by which the higher levels of sensory and

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

associative circuits are kept so busy processing the internally generated messages that they ignore external signals" (Hobson 1990, 206. See also Hobson 2009). To be detached from the outer world doesn't mean to be internally switched off, as if the sleep were a death-like state, but, on the contrary, the inner activity of the sleeping – and dreaming – brain is so full that it can't need a relationship with the outer reality. The sleeping /dreaming brain creates its own living world. This peculiar condition of the brain-mind, that Hobson accurately denotes as "internally activated but sensorially disconnected" (Hobson 1990, 206), doesn't allow the body to establish a relation with the outside and so it remains in rest for all the duration of the REM sleep phase[5]. This coexistence of two contrasting processes is always due to a very particular neural state: on the one hand, at the upper levels of the motor system some reticular neurons send "go" commands usually generating the urgency to move; on the other hand, other reticular neurons are sending *at the same time* counter-commands of "no-going" and so it is creating a situation Hobson calls "a 'zero-sum' game", in which "internal motor commands are generated, but their external activation is effectively canceled by concomitant inhibitory signals" (Hobson 1990, 206). This complex process of inhibition, that occurs at the level of both the spinal cord and the brain stem, indicates that the *sensation* of moving doesn't involve the *performance* of moving[6]. If the brain is the locus of origin of such an entirely mental process, the body is the main subject of this intricate play, since it directly *feels* the contrasting game of different neurons, as it lies necessarily immobile though it refers to a fully active occurring. The contrast is vividly felt by the body in those dreams, in which we try to move, for instance to escape, without succeeding. At the level of brain-mind processes, the dreamt body aims at moving according to the "go" commands, but the dreaming body, the only one real, is not able to follow such a will because of the "no-go" commands occurring simultaneously. Each of us perfectly knows the result of such a situation that Hobson clearly describes basing on its neurophysiological conditions: "When the need or desire to escape becomes intense and one's voluntary effort increases, the muscular system is not likely to respond to that command. When dreaming, we experience this inhibition of motor output as heaviness in our legs" (Hobson 1990, 171)[7].

2. Non ordinary ways of living one's own body: Out-of-body experience and sleep paralysis

In this paragraph, I will deal with some very peculiar manners of experiencing one's own body that allow us to stress the involvement of the self in the dreaming process due to her relationship with the body. Moreover, the strangeness of these experiences will help us to face the many ways in which one relates not only to one's own body, but also to one's own self *in conjunction with* one's own body. Thus, it

will offer the opportunity to highlight the *unity* of self and body according to the idea of the embodied self, which I will discuss in the last paragraph while presenting my phenomenological theory.

The human body can undergo further even stranger mental situations, such as the so called Out-of-body experiences (OBEs) that represent a subject matter discussed in several recent neuroscientific studies. OBE can be described as "the experience in which a person seems to be awake and to see his body and the world from a location outside the physical body" (Blanke et al. 2004, 243). A strictly connected experience is autoscopy (AS), "which is characterized by the experience of seeing one's body in extrapersonal space" (Blanke et al. 2004, 243). The main difference between the two kinds of bodily experiences is that "during the OBE, the experiencer appears to 'see' himself and the world from a location other than his physical body (parasomatic, visuo-spatial perspective), whereas the experiencer during AS remains within the boundaries of his physical body (physical, visuo-spatial perspective)" (Blanke et al. 2004, 244). The study focuses its attention on the position of the real body *and* the imagined one, the visual-spatial perspectives, and the emotional reports associated with vestibular-motor sensations[8]. In their theoretical considerations, the authors come back to the issue of the disruption of self representational unity: "We speculate that, during AP, the integration of proprioceptive, tactile, and visual information of one's body has failed due to discrepant central representations by the different sensory systems. This might then lead to the experience of seeing one's body or body parts in a position that does not coincide with the felt position of one's body" (Blanke et al 2004, 255)[9].

Though their observations regard awake patients and don't deal with the dream state, in a passage of the paper Blanke and co-authors remark: "Interestingly, most techniques that are used to induce OBE voluntarily propose that the subjects use a *supine and relaxed position*" (Blanke et al. 2004, 254, emphasis added). I stressed the reference to the position assumed by the participants in Blanke's experiment, because it is about the *same* position assumed by the sleeper during the REM sleep, in particular during that phase of REM sleep that immediately precedes the waking and that can assume the shapes of sleep paralysis (SP). This similarity in the body position between who undergoes OBE's and who sleeps was grasped by Allan Cheyne and Todd Girard (2009, 203), who devote a very analytical paper to the interrelation of vestibular-motor hallucinations, out-of-body experiences, and sleep paralysis. Cheyne and Girard define SP as follows:

"Sleep paralysis (SP) refers to a brief period at sleep onset or offset during which a person is unable to move or speak while remaining awake and aware of the immediate environment. Consistent with the hypothesis that SP represents an intrusion of a rapid eye movement (REM) state into waking consciousness caused

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

by defective coordination of sleep-wake and REM-NonREM transitions, polysomnographic studies report mixed REM and waking electroencephalographic components during SP episodes. […] SP is frequently accompanied by diverse and often vivid hallucinations" (Cheyne and Girard 2009, 201)[10].

Such hallucinations take different forms of expressing that entail not only OBE, but also the phenomenon of felt presence (FP), strange bodily sensations, and illusory movement experiences. Referring to Ronald Melzack's notion of "neuromatrix" (Melzack 2001) Chayne and Girard state that

"OBEs may, in common with phantom-limb experiences, reflect a failure of integration or binding of tactile, proprioceptive, vestibular, motor, and visual experiences of the bodily-self, likely implicating relevant regions within parietal, temporal, and frontal cortices. […] In essence, we argue that OBEs represent a failure of the neurosignature, the pattern of activation of the neuromatrix that signals that the body is intact" (Cheyne and Girard 2009, 202).

 In their study, based on the interaction between SP and OBEs, the authors present various hypotheses to explain such a relationship, concluding that the diverse states implicated in it lead to a general feeling of disembodiment along with a parallel disruption of bodily-self processing. As to my interest in this research, the results show a strict connection between bodily experience and sense of self hence indicating that both the body and the self can be understood only in their reciprocal correlation. Such a correlation emerges in a very fascinating way in the dreaming state and in the peculiar bodily phenomena related to the dream state we are dealing with.

Further neurological implications of these findings are that "hypoactivation of cortical vestibular areas during SP may render these areas unable to cope with the conflicting information coming from different sensory sources as well as potentially increased activity of subcortical vestibular centers. In general, this evidence implicates parietal areas in the generation of vestibular dream imagery and, by extension, in vestibular-motor (V-M) hallucinations during SP" (Cheyne and Girard 2009, 209). In another contribution, Cheyne did already stress the link between SP and dreaming, with particular reference to the experience of hallucinations as a nightmare: "These hallucinations constitute a waking nightmare (w-nightmare) REM experience and are the original referents of the term 'nightmare'. W-nightmare hallucinations are described by a three-factor structure involving experiences consistent with 1) threatening intruders, 2) physical assaults, and 3) vestibular-motor (V-M) bodily sensations" (Cheyne 2003, 163).

The theoretical ground of the speculation by Cheyne and Girard is the already cited notion of "neuromatrix" by Melzack, who proposes the following description of

such a notion: "The anatomical substrate of the body-self, I propose, is a large, widespread network of neurons that consists of loops between thalamus and cortex as well as between the cortex and the limbic system. I have labeled the entire network […] as a neuromatrix" (Melzack 2001, 1379). The neuromatrix produces a cyclical processing and synthesis of nerve impulses called by Melzack "neurosignature". To these two features, Melzack adds two more, i.e. "the sentient neural hub, which converts (transduces) the flow of neurosignatures into the flow of awareness", and "the activation of an action neuromatrix to provide the *pattern* of movements to bring about the desired goal" (Melzack 2001, 1379). What is very interesting in Melzack's proposal is the philosophical assumption underlying it, according to which the sense of self is meant as the *result* of neurophysiological processes constituting the bodily experience *whereby* one feels herself *also* as a self. In a certain way, it represents an upsetting of the Cartesian perspective, as Melzack says: "The neuromatrix, distributed throughout many areas of the brain, comprises a widespread network of neurons that generates patterns, processes information that flows through it, and *ultimately* produces the pattern that is felt as a whole body possessing a sense of self" (Melzack 2001, 1380).

Following Melzack's schema, Cheyne and Girard observe that:

"disruption of the widely distributed neuromatrix can affect the processing of a variety of bodily senses involving vestibular, motor, and proprioceptive systems as discussed. This disruption may directly produce anomalous bodily experience as well as contribute to the breakdown of the neurosignature, particularly when central integration is compromised" (Cheyne and Girard 2009, 210).

The SP, interacting with these neuro-disruption, influences some sensory processes such as the "environmental change picked up by exteroceptors" (very attenuated during SP), and "the kinesthetic feedback from the body periphery" (unavailable during SP). At the same time, SP leaves intact the "enhancement of sensory areas either directly via corollary discharge/efference copy" or "indirectly through the construction of a forward model (emulator)" (Cheyne and Girard 2009, 210).

3. Phenomenological account of embodied dreaming

I will present an interpretation of the body in dreaming, both as a normal experience of a paralyzed body during REM sleep and a pathological experience of OBEs and FP during SP, grounded on the philosophical perspective of phenomenology.

According to phenomenology, any mental act is characterized as "intentional", which means the inborn trait of the act to be directed towards an object and in doing so to confer sense to it. The basic scheme of "intentionality" can be briefly explained as follows: to any act that "performs an intention", an object as "intentioned" corresponds. To further clarify by means of concrete examples: as to a perception,

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

one has "the perceiving" (i.e. the act that performs the perception) and "the perceived" (i.e. the object that is the target of perception); as to a memory, one has "the remembering" (i.e. the act that performs the memory) and "the remembered" (i.e. the object that is the target of memory). It is the same for any act of mental life, such as expectation, imagination, and dream. Within this general feature, consciousness' acts receive a further division between: 1) the "presenting" acts, i.e. the acts that present an object, the acts that confer sense to an object that is in turn present. The perception is the basic shape of "presentation" (*Gegenwärtigung*): I can perceive only something present. This means that the perceiving and the perceived are simultaneously. And 2) the "presentifying" acts, i.e. the acts that "make something present", the acts that make present something that is not currently present. Memory, expectation, and imagination are basic shapes of "presentification" (*Vergegenwärtigung*): I can remember, expect or imagine only something *not* present. This means that remembering, expecting, and imaging work on a double level of temporality: as mental acts they occur in the present, but their objects (the remembered, the expected, and the imagined ones) belong to the an unreal – non actual, i.e. non present - dimension (and so they are "presenti*fied*", "*made* present", from the Latin "facĕre").

I propose to define the dream as a presentification too[11]. In the case of the dream, such a relationship between the double level of consciousness is even more peculiar: the first level is the one of sleeping, in which there is no aware activity, while the second level, the one of presentification, is the layer of dreaming, in which there is no aware activity *either*. While dreaming, one makes present something (presentifies) without being conscious of it like when imagining or remembering.

Few, though relevant thinkers of the phenomenological tradition highlight some common traits of the dream state, such as: absorption, entanglement, ego-splitting, transformed relation between the subject and her space/time, the re-constitution of a world[12].

I would propose a phenomenological description on dreaming state based on the idea of a phenomenological *formalism* for describing the consciousness' mental acts, referring to the concept of *intentional mental process*. Starting from this phenomenological perspective, Eduard Marbach (1993, 2006, 2010) provides a conceptual framework aiming at describing the essential, fixed features of consciousness' experience. He calls such a framework "phenomenological notation", which is a formal description of mental acts inspired by Frege's concept of "Begriffsschrift" (Frege 1964) and consisting in a series of symbols and signs. As Marbach explains, "The notation should serve the purpose of expressing with precision *reflectively* elaborated forms of consciousness following methodological devices of Husserlian phenomenology" in order to "make concretely intelligible

that the very ways of intentional referring to something are simultaneously so many modes of inner consciousness" (Marbach 2010, 61-62). Basing on the rich notion of presentification, Marbach elucidates that it regards all acts which are "modifications of the basic form of perceptual consciousness" (Marbach 2010, 62) and so presentification stands for a very useful conceptual feature to understand the life of consciousness occurring in conjunction with the presentness of the ego while referring to something absent or unreal. Following Marbach's proposed phenomenological formalism, I suggest explaining the dreaming experience as a compound of consciousness' intentional acts[13].

The phenomenological description of dream experience based on the notation appears as follows:

$$i_____(PRE^2 \vdash [PER]\,) \vdash x$$
$$(\ \top\ PRE^1)\,s$$

This formula means: I (i), while based in potentially presenting (PRE^1) my surroundings (s), am actually presentifying (PRE^2) something that is dreamt (x), believed to be real, by means of presentifying (dreaming) a perceiving (PER) of x believed to actually occur.

The elements of this description are symbolized as follows: the triplets of upper case letters serve to denote mental activities; a pair of parentheses '(...)' together with 'x' indicates the intentional correlation between a mental act that actually occurs and its object; pairs of square bracket, '[…]' surrounding expressions for mental activities designate the occurrence of an intentional implication or modification of a mental activity contained within another activity; the sign of the shape ' $\vdash$ ' stands for the "belief-stroke" that, put in front of the symbol of intentional correlate, means that this is taken for something actually existing; the sign of the shape ' $\top$ ' stands for the "negation-stroke" that, put in front of triplet upper case letter, means the non actuality of mental act; the horizontal stroke '____' is called the "foundation-stroke" and signifies the presentational experience (the presentness) that always happens at the very same time of the presentifying performance as its basis either as perceived, conscious (in the case of a presentification as imagination) or unperceived, unconscious (in the case of a presentification as dream); the letter "i" is put at the beginning of the formula, in order to indicate the I-consciousness involved in the presentificational experience[14]. The only sign I added to Marbach's original work is the "negation-stroke" ' $\top$ ' that I draw from Frege's notation wherein it indicates that something doesn't take place (Frege 1964, 10). I introduce this sign in order to formally describe the peculiar condition of sleeping self who is presenting *potentially* (she can wake up at any time maintaining the perceptual

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

features though "sensorially disconnected" (see: Hobson 1990, 206). It is about "the having the world in the extreme manner of absorption" (Fink 1966, 64); hence, it is not about an actual presentation, but rather about a potential one.

Now, through the elements of the "notation", I would suggest a phenomenological description of dreaming as it occurs during OBE's in conjunction with SP (Fig. 1).

Keeping in mind the aforementioned description of SP, I propose:

$$i\underline{\qquad\qquad}(PER\text{-}PRE^2 \vdash [PER]) \vdash x$$
$$(\;\top\; PRE^1)\,s$$

This formula means: I (i), while based in potentially presenting (PRE^1) my surroundings (s), am actually perceiving/presentifying (PER-PRE^2) something (x), believed to be real (and actually moving, flying, floating), by means of presenting/presentifying (hallucinatory perceiving/dreaming) a perceiving ([PER]) of x believed to actually occur.

Let us consider the two formulas put together:

1) $$i\underline{\qquad\qquad}(PRE^2 \vdash [PER]) \vdash x$$
     $$(\;\top\; PRE^1)\,s$$

2) $$i\underline{\qquad\qquad}(PER\text{-}PRE^2 \vdash [PER]) \vdash x$$
     $$(\;\top\; PRE^1)\,s$$

When comparing both formulas, one notices that in the second one both the hallucinatory trait of the OBEs and the hybrid nature of SP as a mixed state of REM sleep and waking exclude a performance by dreaming only. For this reason, the second formula lacks an authentic presentificational act. Nonetheless, the common signs in both descriptions are the sign '$\vdash$', which indicates that the experienced 'x', either as content of dream or an illusory moving body, is believed to be actually existing, and the sign ' $\top$ ', which means that it is about sleeping state.

As to both formulas, it is worthy highlighting the coexistence of different mental acts, which characterizes the dream state as a presentificational process. While dreaming, the self and her body experience an overlapping of mental layers that should perform in diverse ambits of consciousness' life. In remembering and

imagining, the presentational (perceptual) level and the presentificational one, though overlapping, don't pretend to the same grade of validity nor confuse the respective ambits of validity. On the contrary, in dreaming the basic layer of presentation is unaware (sleeping), while the layer of presentification (dreaming) hosts in itself a series of presentation (i.e. the dreamt seeing, hearing, touching, and so on) that are believed as actually occurring and to which a series of dreamt bodily movement are associated. As a result, while sleeping one is basically presentifying through a sequence of acts that take the shapes of presentations. By means of the notions of presentation and presentification and of their overlapping interaction, phenomenology is able to offer a conceptual description of the complex state of dreaming meant not only as a neurophysiologic or psychological process, but also as an *intentional* one. The formal stage of description allows showing through commonly recognizable signs the basic features of this process: the formulas can be from time to time adjusted to the peculiar dream state to be described and, thanks to the recurrence of the same signs and symbols, they can help to conceptually read the specific dreaming process taken into account[15].

Finally, it is very relevant to stress the presence of an 'I' at the very beginning of the discussed formulas. The reference to the ego/the self is of basic importance for the phenomenological account, since it allows it to anchor the complex interplay of mental acts (and of the neurobiological processes underpinning them) to the central dimension of individuality[16].

Conclusion. The body and the self as a whole: Towards a meeting point between phenomenology and neuroscience

My proposal is based on the idea of an interdisciplinary approach to the issue of dreaming, and of consciousness' states in general, i.e. an approach in which there is no leading thought as *primus inter pares*, but rather each field of inquiry contributes to the whole explanation of the subject matter[17]. As to phenomenology, it can find in the neuroscientific studies not only a realm of empirical data lacking in a purely philosophical inquiry, but also a fascinating way of accounting for the consciousness' life through a different language than the philosophical one. In this case, the differences in speaking must be more a reciprocal enrichment than a source of misunderstanding. The latter can be avoided if one keeps in mind that both phenomenology and neuroscience deal with the same topic – what consciousness is like – adopting diverse manners of approaching it. Precisely in this regard, neuroscience can find in phenomenology a conceptual clarity able to make its assertions more solid and legitimate.

As our reflection on dreaming body has shown, this centrality of the self doesn't imply a potentially disembodied soul or a potentially "disensouled" body, but

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

rather it is about an embodied ego – a person -, i.e. an ego whose understanding is possible only by referring to the body/self *as a whole*[18].

I wish thus to offer the following reflections, which don't represent an arrival point, but rather a foundation for a new beginning in the process of inquiring the dream state *together*:

1) The interrelated features of phenomenological description of the existence, i.e. self and body, are represented in dreaming

2) Dreaming is a very fascinating and complex mental experience that can be explained without forgoing an idea of selfhood as either metaphysical or naturalistic entity, but rather as embodied selfhood

3) The phenomenological notation is a conceptual tool aiming to formally explain the essential structures of mental process of dreaming, and

4) By doing so, it becomes possible to recover and interpret the objective validity of neuroscientific results according to a perspective that aims at emphasizing the subjective layer of mental acts rooted in the bodily existence[19].

As a conclusion, I would like to briefly outline a hypothesis on how the phenomenological notation could contribute to further research on dreaming. If one thinks of an integrated study on the complex phenomena of OBE's and SP that would entail both neurophysiological analyses and psychological interpretations, such formalism would allow an interpretation of the data beyond just a medical reading or psychological explanation. Basing on the empirical findings obtained by neuroimaging techniques, the phenomenological notation would aim at describing on a conceptual level the nature of mental acts in any single dream related to OBE and SP, the interplay among such acts, the emotional involvement of the subject (due to the philosophical assumption by the notation of the constant reference to the "I"), and the complex relationship of the presentifying act of dreaming with the present state of sleeping.

In this way, very peculiar mental and bodily events like OBE and SP would receive a global description referred to their neurobiological, psychological, and conceptual elements concerning the dream state. The added theoretical reading would supply the medical and psychological analysis with the formal account of the composition of mental acts involved in dreaming.

Endnotes

1 Allan Hobson highlights Freud's early purpose to develop a scientific psychology in 1890s. See: Hobson 1990, 60-65

2 The stress put on the connection between the sense of moving and the eyes movement aims at emphasizing the neurophysiologic basis of both the visual dream and the dreamer's sensation of moving. It doesn't entail the adoption of the scanning hypothesis, which is currently under debate. See e.g. Arnulf 2001.

3 "[…] during dreaming, PGO signals are used in the construction of the hallucinoid visual imagery of dreams. The robust evidence for the generation of feedforward information, the excitability and functional significance of which varies systematically with brain state, is an important element of dream consciousness theory. In dreams, we see – in the absence of light – by activating the visual system of the brain from the motor side up" Hobson 2009, 809

4 In the present paper I will deal only with those dreams in which there is no active involvement by the dreamer for controlling somehow the dream course. Hence, I will exclude the complex and very fascinating case of the so-called lucid dreams (LaBerge 1985; LaBerge and Rheingold 1990)

5 I am referring to the "phasic" stage of REM sleep which is the specific trait of PGO activity, thus I am not taking into account the "tonic" stage of REM sleep that is characterized by an ocular quiescence

6 Commenting Edward Evarts' study (Evarts 1960) on the neurobiological basis of the similarity between dreaming and waking states, Hobson remarks: "This finding meant that, during dreaming, motor commands are actually being issued by the cortex but not being enacted, and thus could be related to our sense of constant movement during dreaming. Such 'fictive' dream movements have a nonfictive base!" Hobson 1990, 169.

7 Such a peculiar condition of high brain activity and bodily atonia represents a paradoxical though normal state of sleeping whose loss can determine a serious sleep disorder known as REM Sleep Behavior Disorder (RBD). RBD was carefully investigated by Schenck et al. (1986) and Schenck and Mahowald (2002, 2005). Revonsuo (2000) takes into account RBD and NREM-related sleep disorders in conjunction with his original theory of dream as threat simulation, observing that such disorders "show that threat simulation during sleep includes realistic and adequate motor activation in the brain in response to the perceived threats." (Revonsuo 2000, 890)

8 More to the point, paying attention to the vestibular sensations implies considering the basic role played by the vestibular cortex, whose core region "is situated at the temporo-parietal junction (TPJ) and/or the posterior insula", Blanke and co-authors stress also "the importance of the TPJ in normal and pathological visuo-spatial perspective taking […]". The basic implication of the vestibular processes makes clearer the involvement of the body – both as real and as unreal, illusory – in these special mental states and the related involvement of the sense of self: "the present data suggest that AP might also relate to a failure to integrate proprioceptive, tactile and visual body-related information (disintegration in multisensory personal space) in a coherent central representation of one's body (body schema)" (Blanke et al 2004, 253)

9 The disruption of the bodily self implies the co-disruption of the spatial self, which seems to be the condition to experience a parasomatic condition. See Blanke et al 2004, 255 and 256. The interactions between spatial and psychological dimensions in autoscopic phenomena and their relationship with the bodily self have been highlighted by Brugger 2007

10 It is about a hybrid state very similar to the one recognized as a lucid dreaming, in which the boundary between wakefulness and dreaming becomes uncertain (Voss et al 2009). Such

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

an uncertainty between both states gets lucid dreaming close to OBE's too. See: Metzinger 2009, 133 ff

[11] Edmund Husserl, the founder of phenomenological thought, doesn't consider dream a presentification, since he states that a presentification needs experiencing an aware contrast with perceptual reality. Such a contrast is experienced in the case of remembering or imagining, but it doesn't occur in the case of dreaming (Husserl 2005). In defining dream as a presentification, I follow Eugen Fink's (1966) remarks. In his considerations on imagination as a presentification, Fink, the last Husserl's co-workers in the 1930s, affirms that the *absorbed* status of the ego stands for the other *fundamental* character of subjectivity, beside the mental state of full awareness. As a consequence, the dreaming process is meant by Fink as a *more* absorbed (*versunkener*) mental state that assumes the shape of a presentifying act. See Fink 1966, 54-55

[12] Fink 1966, Sartre 1972, Conrad 1968

[13] Marbach applies such a phenomenological notation to the acts of remembering, imagining, picturing something, and their combinations, since, according to the Husserlian standpoint, he doesn't consider the dream as a presentification either (Marbach 1997, 607)

[14] As regards all the aforementioned signs and further ones I didn't deal with, see Marbach 1993, 25-27; Marbach 2010, 64-65

[15] As to the descriptive function of these signs, I refer to the clear words by Marbach himself: "The artificial signs of the phenomenological notation will enable us to separate the particular moments of an activity clearly from one another and simultaneously make their phenomenological unity, i.e. their intentional relations more readily visible than all too longwinded expressions from ordinary languages would permit. […] Once the notation will be more developed, improved, and hopefully standardized, the logic of the systems so discovered may even lead to new phenomenological discoveries" Marbach 1993, 23. See also: Frege 1964, 97

[16] As Marbach properly states, the "complexity of the components of a representing [presentifying] activity is simultaneously unified in an I-consciousness. The unification of temporally and spatially discontinuous moments takes place in an I-consciousness. […] I-consciousness in acts of mentally representing [presentifying] x can be said to occur as a subjective unifying reference of the plurality of the actual moments and of the non-actual ones that are intentionally implied in the unity of performing one and the same activity converging objectively on one and the same x" (Marbach 2010, 67-68. See also: Marbach 1993, 89-92)

[17] Valuable proposals to apply the phenomenological method to consciousness' study have been suggested by Amedeo Giorgi (2009; 2012) and Don Kuiken (Kuiken et al 1989) referring both to Husserl and Merleau-Ponty's inquiries

[18] See: Merleau-Ponty 1962, 408

[19] My attempt is not the first one to put phenomenology in conjunction with neuroscience. Already in the 1990s, Francisco Varela proposed the so-called "neurophenomenology", i.e. a phenomenology that practices its methodological insights interfacing with the scientific studies of consciousness. My proposal is somehow situated in the field of inquiry opened by Varela whose inspiring principles I completely agree with. I am especially referring to some Varela's statements (1996) such as the mutual constraints between phenomenology and neuroscience, the thought of the two ambits of research as different ways to consider the same world, the existence of shared topics like attention, time-consciousness, body image, among others. Nonetheless, I prefer to speak of my idea as a phenomenological formalism of dreaming instead of a neurophenomenology of dreaming, because of the relevance I confer to the formal description of the mental acts constituting the dream experience

References

Arnulf, I. (2011). "The 'scanning hypothesis' of rapid eye movements during REM sleep: a review of the evidence." *Archives Italiennes de Biologie 149*, 367-382

Aserinsky, E., & Kleitman, N. (1953). "Regularly occurring periods of eye mobility and concomitant phenomena during sleep." *Science 118*, 273-274

Blanke, O. et al (2004)."Out-of-body experience and autoscopy of neurological origin." *Brain 127*, 243-25

Brugger, P. (2002). "Reflective mirrors: Perspective-taking in autoscopic phenomena." *Cognitive Neuropsychiatry 7 (3)*, 179-194

Cheyne, A., Girard, T. (2009). "The body unbound: Vestibular-motor hallucinations and out-of-body experiences." *Cortex 45*, 201-215

Cheyne A. (2003). "Sleep paralysis and the structure of waking nightmare hallucinations." *Dreaming 13*, 163–179

Cramer Bornemann, M. et al (2006). "Parasomnias. Clinical Features and Forensic Implications." *Chest 130(2)*, 605-610

Conrad, T. (1968). *Zur Wesenslehre des psychischen Lebens und Erlebens*. The Hague: Nijhoff.

Dement, W., Kleitman, N. (1957). "Cyclic Variations in EEG during Sleep and their Relation to Eye Movements, Body Motility, and Dreaming." *Electr. Clin. Neurophy. 9*, 673-690

Evarts, E. (1960). "Effects of Sleep and Waking on Spontaneous and Evoked Discharge of Single Units in Visual Cortex." *Federation Proc. Supplement 4*, 828-837

ASASE
The Altered Self in Dreaming State related to the Out-Of-Body Experiences.
A Phenomenological Account interfacing with Neuroscience

Fink, E. (1966). *Vergegenwärtigung und Bild.* In: Idem, *Studien zur Phänomenologie 1930-1939*. The Hague: Nijhoff, 1-78

Frege, G. (1964). *Begriffsschrift und andere Aufsätze*. I. Angelilli. (ed). Darmstadt: Wissenschaftliche Buchgesellschaft

Giorgi, A. (2009). *The Descriptive Phenomenological Method in Psychology*. Pittsburgh: Duquesne University Press

___ (2012). "The Descriptive Phenomenological Psychological Method." *Journal of Phenomenological Psychology 43(2)*, 3-12

Hobson, J., McCarley, W. (1971). "The Brain as a Dream State Generator: An Activation-Synthesis Hypothesis of the Dream Process." *The American Journal of Psychiatry 134 (12)*, 1335-1348

Hobson, J. (1990). *The Dreaming Brain*. London: Penguin Books

___ (2009). REM sleep and dreaming: towards a theory of protoconsciousness. *Nature Reviews/Neuroscience* 10 (11), 803-13

Husserl, E. (2005). *Phantasy, Image Consciousness, and Memory (1898-1925)*. Trans. J.B. Brough. Dordrecht: Springer

Kuiken, D., et al (1989). "Numerically aided phenomenological method. A demonstration." *Journal of Mind and Behavior 10*, 373-392

LaBerge, S. (1985). *Lucid Dreaming. The Power of Being Awake and Aware in Your Dreams*. New York: Ballantine

___, Rheingold, H. (1990). *Exploring the World of Lucid Dreaming*. New York: Ballantine

Marbach, E. (1993). *Mental Representation and Consciousness. Towards a Phenomenological Theory of Representation and Reference*. Dordrecht: Kluwer Academic Publishers

Marbach, E. (1997). "Re-presentation." In: L. Embree et al. (eds.). *Encyclopedia of Phenomenology*. Dordrecht: Kluwer, 603-609

Marbach, E. (2006). "Wie sich Bewusstsein mit Hilfe der Husserlschen Phänomenologie in die (Neuro-) Wissenschaft einbeziehen lässt." In: D. Lohmar, D. Fonfara (eds.). *Interdisziplinäre Perspektiven der Phänomenologie*. Dordrecht: Kluwer Academic Publishers, 203-234

___ (2010). "Towards a Formalism for Expressing Structures of Consciousness." In: S. Gallagher and S. Schmicking (eds.). *Handbook of Phenomenology and Cognitive Science*. Dordrecht: Springer, 57-81

Melzack, R. (2001). "Pain and the Neuromatrix in the Brain." *Journal of Dental Education 65(12)*, 1378-1382

Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Trans. C. Smith. London: Routledge & Kegan Paul

Metzinger, T. (2009). *The Ego Tunnel: The Science of the Mind and the Myth of the Self*. New York: Basic Books

Revonsuo, A. (2000). "The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming." *Behavioral and Brain Science 23*, 793-1121

Sartre, J.-P. (1972). *The Psychology of Imagination*. Trans. M. Warnock. London: Methuen & Co. Ltd

Schönhammer, R. (2004). *Fliegen, Fallen, Flüchten. Psychologie intensiver Träume*. Tübingen: dgvt Verlag

Schenck C. et al (1986). "Chronic behavioral disorders of human REM sleep: a new category of parasomnia." *Sleep* 9, 293-308

Schenck, C., Mahowald M. (2002). "REM Sleep Behavior Disorder: Clinical, Developmental, and Neuroscience. Perspectives 16 Years After its Formal Identification in *Sleep*." *Sleep 25(2)*, 120-138

_____ (2005). Rapid Eye Movement and Non-REM Sleep Parasomnias. *Primary Psychiatry 12(8)*, 67-74

Varela, F. (1996). "Neurophenomenology. A Methodological Remedy for the Hard Problem." *Journal of Consciousness Studies 3(4)*, 330-349

Voss, U. et al (2009). "Lucid Dreaming: A State of Consciousness with Features of Both Waking and Non-Lucid Dreaming." *SLEEP 32 (9)*, 1191-1200

Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self: some Functional, Computational and Neurological Considerations

．

João Fonseca

Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

．

Abstract: *Full-Body illusions (FBI) and Out-of-Body Experiences (OBE) are two phenomenological experiences with increasing importance in current studies regarding the nature of the Self. Although distinct, these two phenomena share some phenomenological properties and several authors have been suggesting a deep relation between them.*
The article aim is twofold: in one hand to explore the nature of the aforementioned relation between FBI and OBE. On the other hand, to put at test my own proposed taxonomical distinction of the concept 'Core-Self' into two more fundamental categories (Core-Self Simpliciter' and 'Cognitive Core-Self') as previously suggested in other places.
The article starts by reassessing the proposed taxonomical distinction of Core-Self into Core-Self Simpliciter' (CSS) and 'Cognitive Core-Self' (CCS) and then proceeds by suggesting that FBIs and OBE are phenomenological consequences of CSS and CCS respectively.
In the second part of the article this hypothesis is put to test and its consequences both to the taxonomical proposed distinction of CSS and CCS and to the relation between FBI and OBE are discussed.

1. The 'Core-Self *Simpliciter*' /'Cognitive Core-Self' distinction

In this section, I start by briefly reassessing my own proposed taxonomical distinction regarding the concept 'Self' as suggested in previous talks and articles. I will consider first the now widely accepted division of the concept 'Self' into two sub-types: 'Core (or Minimal) Self' (corresponding to the most basic phenomenological sense of oneself) and 'Narrative Self' (regarding the proprieties of self-identification related to auto-biographical contents). My proposal was, and still is, to divide the very concept of 'Core-Self' *further into two sub-concepts*: what I call 'Core-Self *Simpliciter*' and 'Cognitive Core-Self'

This distinction was initially motivated when I was confronted with two distinct proposals regarding the *neural implementation* of Core-Self: Shaun Gallagher's cortical proposal (Motor, Pre-Motor and Prefrontal areas) and Björn Merker's sub-cortical one **(**Upper-brainstem complex: *superior culliculus – Pariequeductal Grey Area/hypothalamus - Substancia Nigra/Basal Ganglia*)[1]**.**

Without going further into details concerning each of these two proposals, the fact is that they seem, not only different but also *incompatible* answers to the question regarding the neural implementation of Core-Self. In fact, one proposal is cortical whereas the other is *sub*cortical. Thus, and *assuming* both proposals as empirically correct, we are faced with a sort of an explanatory dilemma: either 1. the two proposals select the *same concept* and the claim of *Multiple Realization* of 'Core-Self' is vindicated (i.e., it has several different neural implementations), or 2. the two proposals select *different and unrelated concepts*. The explanatory 'dilemma' (so to speak) consists of the fact that either of the two possibilities implies a fragmentation in our explanatory practices where some kind of unity would be preferable. My own suggestion in order overpass the dilemma was to notice that the two answers are, in fact, not incompatible. What happens is that Gallagher and Merker are addressing *two distinct although related* perspectives on the concept of 'Core-Self'. The relationship obtained between the two concepts is what allows for the avoidance of the dilemma.

To better understand what's really at stake here; take the analogy with the concept 'Vision'. Sure we can study vision at the level of the visual cortex (call it 'Cortical-Vision'). But we could also study vision at the more fundamental and evolutionary older level of 'photo-reception' (call it 'Vision *Simpliciter*'). These two perspectives on vision of course, are not incompatible and can and should be accommodated into a single *unified approach*. This is so because, among other things, 'Vision *Simpliciter*' (photoreception) is a necessary structural condition to *all* kinds of vision, including, of course, **'**Cortical-Vision'. In the same vein, my suggestion is that the concept 'Core-Self' could be fragmented/divided in two distinct concepts but related in a way similar to that of 'Vision *Simpliciter*' and **'**Cortical-Vision'. These two concepts will, of course, relate to Gallagher's and Merker's proposals. So, to Merker's proposal I suggest the designation Core-Self *Simpliciter* (CSS) whereas for Gallagher's I suggest Cognitive Core-Self (CCS).

Just like in the case of Vision, my proposal states that a unified approach to Core-Self is both possible and welcomed. This is so because just like 'Vision Simpliciter' is to 'Cortical Vision', CSS is a *necessary condition* to CCS.

This necessary condition relation is justified by CSS and CCS properties. I'll state, in an over-simplified way, the main *phenomenological* and *computational* proprieties

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

of each concept (to be developed in more detail in the next sections):

Core-Self *Simpliciter*: at the *Phenomenological level* it provides the basis for the emergence of the 'first person perspective' (1PP) and the basic contrast between Self and external world. This is achieved, at the *Computational level* by Pure On-line processing of sensorimotor integration through coordinate transformation in a common reference framework.

Cognitive Core-Self: At the *Phenomenological level* it adds cognitive complexity to *CS Simpliciter*, namely the capacity for imagery. This is achieved at the *Computational level* by processing off-line over the contents of *CS Simpliciter*, possibly trough the implementation of a Forward Model/Comparator (on the mechanisms responsible for CSS).

So, again, just like 'Cortical Vision' adds computational and neural complexity to the (already existent) mechanisms for 'Vision *Simpliciter*', CCS adds neural, computational and even phenomenological complexity to the more basic CSS.
For now, this brief and somewhat simplified portrayal of the distinction between CSS and CCS is sufficient for making the point of the remaining paper. It is time now to bring the notions of Full Body Illusions (FBI) and Out of Body Experiences (OBE) into the discussion. My main claim in this paper, and which I will try to argue for in the next sections is enounced in the following statement:

FBIs and OBEs are manifestations of, respectively, CSS and CCS.

Put simply, the idea is that FBI are a kind of sensorial illusion made possible by the CSS computational properties and the same for OBE in relation to CCS.

Supported in this hypothesis, my aim in this paper is twofold: in one hand I will try to suggest that FBI and OBE are intimately related (that relation being that of dependence of CCS from CSS). On the other hand, I hope to show some empirical and conceptual evidence in the process, in favor of my previous distinction between CSS and CCS.

2. A Charactarization of 'Core-Self *Simpliciter*'
This section aims to explore, in more detail, the computational and phenomenological properties of CSS and how those properties predict or explain the FBI phenomena. In order to achieve this end, I will address Bjorn Merker's (2005, 2007) own suggestion for the biological and evolutionary origins of what I call Core Self *Simpliciter*. In particular, the main question concerns the biological and evolutionary constraints in the emergence of this most basic form of (core) Self.

Merker departure is the so-called *Stability Problem*. This is basically the sensorial-motor integration problem faced by creatures with more complex bodies given the

several distinct sensorial modalities and the huge number of different possible body actions. In a nutshell, there are too many scattered variables at the level of sensory-motor integration, which prevents good sensorial-motor coordination in the environment. Merker (2005) provides a nice snapshot of the problem:

> The evolution of higher animals leads not only to increased complexity of single sensory and motor systems but produces a *diversification* of such systems in the equipments of a given species. Vision, hearing, touch, smell, taste, interoception, proprioception and vestibular systems are some of those on the sensory side, while a great variety of locomotors, orienting, grasping, and manipulation appendages – often paired in sets and with multiple, independently moving joints – proliferate on the motor side. Such diversity brings special problems in its train. On the sensory side, for example, the receptor arrays of different modalities are often disposed in different parts of the body. They are therefore, not affected in the same way by self-motion and so cannot be subject to the same compensatory remedies, nor be integrated directly. (Merker 2005, 92)

In order to achieve a satisfactory sensory-motor coordination to survive in the environment, 'complex' creatures thus need some sort of centralization/integration of the different sensorial and motor 'local solutions'. In a more formal terminology: we have several separate *frames of reference* for each sensorial and motor specific information and what is needed is to frame all this disparate information under a *single coordinate reference system*.

How can such different sensorial and motor frames of reference be unified? The answer is: trough *Coordinate transformations*.

Merker's own proposal is that nature's structural answer to this challenge relates to the phylogenetically very old and widely conserved structures corresponding to the basic Vertebrate Brain Plan, namely the roof/tectum of the midbrain/Superior Colliculus (S.C.). Merker stresses the peculiar anatomical and functional roles of this structure:

> [T]he spatial senses are topographically superposed in laminar fashion within a common, premotor, framework for multieffector control of orienting. Its functional role appears to center on convergent integration of diverse sources of information bearing on spatially triggered replacement of one behavioural target by another (Merker 2007, 67)

The S.C. laminar structure is such that some information registered in one layer (say a mapping of auditory space) translates directly to another layer (say, a mapping of occulomotor orientation) by means of *coordinate transformation*. This is why when we hear a sound our eyes 'know' immediately where to look in the search for the

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

source of the sound. What this means is that the S.C. laminar structure *in fact* brings the various local sensory-motor solutions under a *common referential system framework*.

Nevertheless, it seems that the SC is not the only place in the brain capable of implement sensorial-motor coordinate transformation. Several different places in the brain seem to implement this computational solution. As we going to give evidence later, one of those areas is the Posterior Parietal Cortex (PPC). That being so, and for the purposes of this article, my approach consists in *abstracting* the general computational formulation for coordinate transformation from its specific neural implementation (be it the SC, the PPC or any other). That is, all the specific brain structure responsible for sensorial-motor coordinate transformation can be described (at a very abstract and general level) in the same way. Therefore, *generally*, in neurological terms, this solution takes the form of an integrative brain structure where information is encoded in a specific neural population and pre-specified coordinate transformations *communicate synaptically* that information from one neural population to another.

My purpose now is to suggest how certain phenomenological features generally attributed to the *first person perspective* (1PP) can arise from this sensorial-motor coordinate transformation **computational** framework.

As suggested above in the description of the SC laminar function, coordinate transformations of *both* different sensorial sources (visual, tactile, audio, proprioceptive) and motor behaviours gives rise to a coherent unified perceptive experience. For instance: A visual space can be 'calibrated' to an audio space through coordinate transformation, such that we can visually identify the source of a sound. Suppose I see a glass being broken *and* hear a shattering sound at the same time. The coordinate transformations from one space to the other makes me associate the two, making me able to visually identify the source of the sound.

This example suggests how the *sensorial* side of experience can be unified. Now, to this *external* input of sensorial integration let's add *internal* information regarding the motor aspect of experience, i.e., body movements and position. Take the example of extra-occular proprioceptive information regarding the movements of the eyes of their eyeballs *relative* to the rest of the face. In this case, coordinate transformations provide a body/face-relative reference frame for eye position. The coordinate transformations from an extra-occular proprioceptive space to a 'face'-map space turns possible to detect *where* the eyes are in relation to that facial space; if they are looking down, up, left, right, etc, relative to that facial reference system.

Now, if we add to this the information of what the eyes *actually* see (i.e., the visual information space) what emerges is an eye orientation frame for that visual

information. As I move my eyes I experience –for example- *looking* right, left, up, down. A fundamental distinction arises in this context: the contrast between *my body* (my eyes in this case), which I move voluntarily, and the *external world*, which rests 'indifferent' (i.e., stable) to my volition. I voluntarily move my eyes to the right and can see what is positioned at *my right* in the external environment. Moreover, in geometrical terms, there must be a *point of origin* of this body frame of reference. This point of origin defines a phenomenological, subjective spatial perspective in the sensorial space, corresponding to what is known as *First Person Perspective* (1PP). For instance, Blanke and Metzinger define a weak 1PP as follows:

> A purely geometrical feature of an egocentric model of reality, most often referred to as a visuospatial model. [It] includes a spatial frame of reference, plus a global body representation, with a perspective originating *within* this body representation. (Blanke & Metzinger 2008, 7)

We should focus our attention on the two features identified by Metzinger and Blanke regarding 1PP, that is: 1) a spatial frame of reference and 2) a global body representation (with a perspective originating *within* this body representation).

Regarding the first feature, we have already shown evidence how information concerning the body (in our example, information regarding eye-position relative to the face) defines a frame of reference to orienting external sensorial information. Gallagher stresses this fundamental role of the body in providing a spatial frame of reference for sensorial input:

> [O]ne of the important functions of he body in the context of perception and action is to provide the basis for an egocentric (or body-centred) spatial frame of reference (…) this egocentric framework is required for the very possibility of action, and for the general structure of perceptual experience. The fact that perception is perspectivally spatial is a fact that depends precisely on an implicit reference to the spatiality of the perceiving body. (Gallagher 2005, 137)

Why so much dependence on the body? According to Gallagher, body awareness provides the only way to avoid the infinite regress in defining a perceptual-spatial, egocentric point of view:

> If one accepts the premise that sense perception of the world is spatially organized by an implicit reference to our bodily framework, the awareness that is the basis for that implicit reference cannot depend on perceptual awareness without the threat of infinite regress. To avoid the infinite regress one requires a pre-reflective bodily awareness that is built into he structures of perception and action but that is not itself egocentric. (Ibid, 137)

Proprioceptive information is particularly well suited to the task since:

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

> [P]roprioceptive awareness is not itself a perception of the body as an object; for if it were, it would require an ordering system, a spatial frame of reference that was independent of the body. Generally speaking, the proprioceptive spatiality of the body is not framed by anything other than the body itself. (Ibid, 137-138)

The above warning that proprioception is *not* itself a perception of the body *as an object* takes us to the second feature of 1PP, as proposed by Metzinger and Blanke, namely the requirement of a 'global body representation'.

At this point it is probably useful to introduce the notions of *Body-Schema* and *Body Image*[2]. Whereas a *Body-Schema* consists in a series of body structures and (mainly) unconscious contents *constitutive* of conscious experience itself, a *Body Image* corresponds to the actual conscious perception of ones body. It is more or less clear that, in the present context, the body frame of reference constitutive of 1PP relates to the notion of *Body-Schema* and the 'global body representation' requirement to *Body Image*. Regarding the latter, the (simple) idea is that, the perceptual experience of an external space includes several objects and events occurring in that external space, one of which is one's body itself. The main difference between *our* body as it appears in our perceptive field and other external objects, is the fact that we can recognize our body as *ours* whereas the other objects (other's bodies included) are experienced as independent from us (recall the above discussion considering the distinction between ones body and other objects in terms of conscious voluntary control of one rather than the other).

Accordingly, Metzinger and Blanke's requirement for a 'global body representation' as a necessary feature of 1PP can be rephrased as a requirement for a *Body Image*. This demand can be fulfilled in our coordinate transformation framework in the way loosely described above, that is: in one hand, the sensorial detection of the body in the *external* space and, in the other hand, the recognition of that body as belonging to us trough, among other things, its voluntary motor control.

Metzinger and Blanke make a further claim concerning this 'global body representation' (or *Body-Image*), namely that a (subjective) perspective should have its origin located *within* this Body Image. It is more or less easy to understand how this requirement is met within the coordinate transformation framework discussed so far. Take again the above-discussed case of the face as providing a coordinate reference for eye position. Loosely, at their 'resting position' (when just 'looking forward', not up, not down nor left or right) the eyes define the origin (i.e., the point zero) of this coordinate system (as the eyes move in a certain direction their position is defined as a certain point-vector in that system). Assuming this, one can conclude that this origin is implicitly «located behind the nose bridge inside our head. From

there we *appear* to confront the visible world directly through an empty and single cyclopean aperture in the front of our head» (Merker, 2006, 73, emphasis in the original). That is, the *Body-Schema* (namely trough proprioception information) defines the 1PP point of origin as being located within a certain specific place in the *Body Image*, more particularly, inside our head (as part of our Body-Image). One should not forget that this Body Image is not only visual but also a tactile and somesthetic one. Surely, we cannot see our own head (aside from mirror information) but other kinds of sensorial information other than vision fill the gap in the construction of our *Body Image*. As Merker stresses:

> The cyclopean aperture is a convenient neural fiction through which the distal visual world is "inserted" through a missing part of the proximal visual body, which is "without head" as it were or, more precisely, missing its upper face region. *Somesthesis by contrast maintains unbroken continuity across this region*. (Merker 2007 73, emphasis added)

So, summing up this section: through multisensory-motor coordinate transformation (as an evolutionary answer to the stability problem) we have the generation of: 1) a self-perspective first person position (1PP), 2) the postulation of an external independent world and 3) an image of body [a *Body Image*] where the origin of the subjective perspective is located. These are, in a nutshell, *Core-Self Simpliciter's* (CSS) proprieties.

3. 'Core-Self *Simpliciter*' and Full Body Illusions

In this section I propose to frame the Full Body Illusions (FBI) phenomena within the context of CSS properties as described in the last section. First, I describe succinctly the basics of FBI illusion, and then I propose to show how this phenomenon is (at least) partly explained in terms of CSS properties.

In their fundamentals, FBIs are very simply depicted: a subject is stroked in the back and sees, trough virtual-reality goggles, her own back a few steps ahead being *stroked in real time and in synchrony* with the felt strokes. After some time the subject reports feeling as being physically moved to the body they see in front of them. As an illustration of this, after blanking the image shown in the goggles and slightly displacing participants' position, they were asked to move to where they felt themselves being during the experiment (their original or previous position). As a result, subjects moved towards the supposed location of the body as indicated by the virtual-reality goggles and not their actual original position. This general effect is technically known as *proprioceptive drift* (Lenggenhager et al 2007).

How is this strange experience possible and what does it tells us about the self? My claim in this section is that the computational model for *CS Simpliciter* proposed in the previous section, in fact, **predicts** this phenomenon.

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

Of crucial importance here is the notion of 'pre-specified coordinate transformations'. Essentially, this is the idea according to which at least some of the multi-sensory coordinate transformations discussed above are determined by experience (or innately) resulting in enduring and specified space-to-space mappings. These cordite transformations are, so to speak, built-in in order to allow successful sensorial-motor coordination. This pre-specification can explain (and be clarified by) several illusions, FBI being one of them.

Take a first illusion as an illustration. In a well-known experiment (Stevens et al 1976), the muscles of subject's eyeballs were artificially paralysed by a pharmacological agent. When asked to look right, subjects reported watching the complete visual scene as jumping to their right. This is because the pre-specified coordinate transformation makes the (possibly efferent copy of) eye-motor command to have an effect on the orientation space in a pre-specified way (e.g. turn right in x degrees) whereas the visual scene stays the same since there is no actual eye movement. Therefore, the whole visual scene will be perceived as turning right in x degrees.

Before turning to FBIs, let's take another illusion that can be helpful in this context: the well-known 'Ventriloquist Illusion'. In this case the coordinate transformations establish a pre-specified correlation between sounds and their spatial source location by matching temporal synchrony between those sounds and visual events. This explains why we, despite knowing the contrary, tend to attribute the source of the speech to the dummy. The same rationale explains why when watching a movie we hear the speech as coming from the actors in the screen and not from the loudspeakers located in the movie-theatre.

The 'Ventriloquist Illusion' is of interest here since FBI can be explained in a similar way. In the former, pre-specified visual-audio coordinate transformations localize the source of the sound to where synchrony is found. *Mutatis mutandis*, in FBI pre-specified visual-tactile coordinate transformations spatially localize the body as standing in the position where synchrony between tactile and visual information is found. In the case of the experiment setting described above, this synchrony is found in the body that is located in front of the subjects. As a result, subjects report as being dislocated to that body, that is, they report a dissociation of their Image of the body (Body Image) from their 1PP. As we have seen, in normal circumstances the two features of 1PP, the Body Image and the Body-Schema, coincide, meaning that the geometrical origin of the orientation-in-space coordinate reference system (Body-Schema) is localized within a specific body location (Body Image) (inside the head just behind the nose bridge). It is this coincidence that breaks down in a FBI.

But this is just part of the explanation, since things turn to be slightly more complex than suggested. In fact, what seems to happen is that subjects *do not* experience disembodiment of the self (i.e. a separation of their subjective perspective from their body). Instead, subjects report a spatial misplacement of their Self *and* body (the experience is reported as "*I am* 'there'" and not as just "*My body* is 'there'") (Lenggenhager et al. 2007). There seems to be a Body-Schema and Body Image coincidence after all. Lenggenhager and colleagues give a general characterization of FBIs as follows:

> [A] n illusion during which healthy participants experienced a virtual body as if it were their own *and* localized their "selves" outside of their body borders at a different position in space. (Ibid, 1096, emphasis added)

So, what is possibly happening and what could explain the phenomenon?

Again, coordinate transformation suggests a possible explanation for this total displacement of self-position and not just body-position in FBIs. What possibly happens is that again in this case a *pre-specified* configuration in coordinate transformation is preserved. More specifically, and as already testified, in normal circumstances the geometrical origin of the body-schema (that is, the sense of 1PP) is experienced as being located within a certain specific part of the body-image: inside the head just behind the nose bridge. This strongly suggests that the coincidence between body and 1PP is preserved in a pre-specified way, in the sense used here. Either acquired or innate, there is probably a mechanism of compensation if, in abnormal circumstances, the 'normal' location of 1PP in the *Body Image* is disrupted. That is precisely what happens in FBI; visual-tactile synchrony is spatially located in a different position than the actual body, but given the pre-specified coincidence between 1PP and body-image, is not just the body-image that suffers displacement but 1PP as well. In other words, the 1PP-Body-Image coincidence is preserved, 'all things being equal', trough pre-specified coordinate transformation enabling the judgement: *'if my body is located there so is my Self/ 1PP'*.

Can we find any neuroscientific evidence supporting the hypothesis here suggested concerning the computational explanation for FBI phenomena? As far as I know, no Neuro-image studies have been undertaken concerning FBI experiences. Nevertheless, those studies have been conducted in the realm of so-called 'Rubber Hand Illusion' (RHI). RHIs were conceived and studied before but can be understood as FBIs 'sisters experience'. In a RHI setting, a subject rests her arm on a table while a screen is positioned to hide the arm from the subject's view. A real-sized rubber hand is placed on the table in front of the subject. Both the 'real' hidden hand and the rubber model are then stroked in synchrony with a small brush. After some minutes, subjects report as feeling the rubber hand as being their own hand.

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

Some Brain-Image studies suggested a correlation between the misattribution of the ownership of the rubber hand and the Posterior Parietal Cortex (PPC) (Ehrsson et al, 2004). The PPC is a brain region known for multi-modal sensorial- motor integration much in the way here proposed trough coordinate transformations (Anderson et al, 1997; Pouget & Sejnowski, 1997). Now, if we assume the claim according to which the FBI is conceived as a generalization of RHI to the totality of the body (see, for instance, Lenggenhager et al, 2007) then, the computational model for Core-Self *Simpliciter* I have been deploying, gains empirical credibility and the same is valid for the tentative explanation of FBIs in terms of pre-specified coordinate transformations.

It seems, therefore, that CSS's properties, achieved through multisensory-motor coordinate transformation (recall: 1) a self-perspective first person position (1PP), 2) the postulation of an external independent world and 3) an image of body [a *Body-Image*] where the origin of the subjective perspective is located) is capable of sustaining all the phenomenological features of FBI's experiences.

4. The CSS/CCS Distinction and Out-of-Body Experience

This section explores the hypothesis- according to which the set of phenomenological experiences known in the literature as 'Out-of-Body Experience' (OBE) - are 'built upon' and add complexity to the computational, phenomenological and neuronal proprieties of CSS and FBI mentioned in the previous section.

I start by briefly characterizing the main features of OBE and comparing them to the ones of Full-Body-Illusion (FBI).

OBE is a phenomenological experience where, in certain extreme situations an individual reports to experience her or himself as detached from its own body and being able to see its body from an external perspective.

Generally and succinctly, OBEs can be defined as satisfying the following phenomenological proprieties (Following Bünning and Blanke 2005):

1) Sense of disembodiment (Self outside the body)

2) Experience of the world from an extracorporeal but ego-centric visual-spatial perspective

3) Autoscopy: seeing one's body from this perspective

These proprieties mark some fundamental distinctions in relation to FBIs. Recall that in FBIs there is no sense of disembodiment: subjects report a sense of complete dislocation to a position different from where the actual body is (proprioceptive drift) but there is not a full-blown 'separation' from it. What seems to happen in

OBEs, in clear contrast to FBIs, is a *non-coincidence between 1PP and an image of the body*. What and how can such a difference be supported? According to the main hypothesis/claim suggested in this article and stated in the first section, OBEs are a consequence of the computational and phenomenological properties of CCS as much as FBIs are consequence of CSS's proprieties.

Recall that it was suggested that, at the *phenomenological level* the CCS adds cognitive complexity to CSS, namely the capacity for *imagery*. On the *computational level* this is achieved probably by *processing* the contents of CSS *off-line*.

In line with this hypothesis it makes sense to suppose that OBEs are the result of a capacity to provide a sort of meta-representation of instances and contents made possible by CSS. For instance, the capacity to *imagine* a first-person visual-spatial perspective as independent of the body could explain the disembodiment feeling typical of OBE. In fact, this hypothesis seems to be supported by a series of empirical findings regarding OBE.

I will briefly sort out some of those findings: In recent years, some neuroscientists have stressed the crucial role of the Temporal Parietal Junction (TPJ) on the occurrence of OBE. More particularly, this structure seems crucially responsible for **two** distinct cognitive capacities with an important role in the occurrence of OBE: 1) the generation of *mental imagery* of body location (Blanke 2005) and 2) as providing an internal multi-modal **off-line representation** of the body. Regarding the first capacity, in a study conducted by Blanke and colleagues it was shown that, in an *evoked potential study*, 10 out of 11 healthy subjects activated their TPJ when mentally imagining their visual-spatial point of view (that is, their first person perspective (1PP)) in a location different from the 'normal' one. They conclude that:

> "Our evoked potential data show (in healthy subjects who have never experienced an OBE) that mental own-body transformations, which mentally simulate the body position and visuospatial perspective that is experienced spontaneously during OBEs, selectively activate the TPJ at 330-400 ms. In addition, we show that this activation correlates with behavioural measures and phenomenological characteristics of OBEs (…)" (Blanke et al 2005, 555)

On what concerns the second cognitive capacity, a series of experiments undertaken by Tsakiris and colleagues evaluated the role of TPJ in the production and maintenance of an 'off-line internal body-model'.

> "Recent studies suggest that the right temporal and parietal lobes and in particular the right temporo-parietal junction (rTPJ) underpins an internal model of the body that could allow the brain to maintain a coherent representation of ones body" (Tsakiris et al 2008, 3014)

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

They devised an experimental protocol, which modifies the Rubber Hand experiment discussed in the previous section. This new setting was conceived to address the role of TPJ in the proprioceptive drift of the real hand to a fake rubber hand or, instead to a neutral object that doesn't look like a hand at all. They found that, by disrupting the TPJ trough transcranial magnetic stimulation, there was a much more higher probability of subjects to report a proprioceptive drift to a neutral object. Tsakiris and his team suggested that:

> "[T]he rTPJ is involved in updating and maintaining [a] body-model by testing which stimuli are relevant to one's own body. We suggest that the rTPJ is actively involved in testing the fit between current sensory input and a stored body-model (…) [t]his model describes the pre-existing visual, anatomical and structural features of the body". (Ibid, 3017)

These two cognitive capacities made possible by the TPJ are of crucial importance for the occurrence of OBEs. The first capacity, regarding the ability to manipulate mentally a dislocation of the first-person visual-spatial perspective, can support and explain phenomenological proprieties 1 and 2 of the above definition of OBE (the sense of disembodiment and the experience of the world from an non-corporeal visual-spatial perspective). The second capacity, the off-line representation of the body, can help in understanding phenomenological propriety 3 of OBEs (i.e., autoscopy).

Now, each of these two capacities seem to *assume* some other capacities and contents *already* being provided. The first capacity assumes a previous first person perspective (1PP) in order to be capable to mentally manipulate it. The second assumes an *on-line* image of the body that is capable of being modelled off-line. As it happens, and as we have seen, a 1PP and an (on-line) image of the body are precisely the two features attributed to CSS (and 'expressed' in FBIs). If so, then a first important consequence can be extracted, namely that that CCS (as defined in the first section) supports OBEs by providing the additional off-line and imagery processing to CSS contents and cognitive capacities responsible for FBIs. A second consequence is the backup to the idea that CSS is a *necessary condition* to CCS. That is, as stated in section 1, just like 'photo-reception' ('Vision *Simpliciter*') is structurally and functionally necessary for the more elaborate and complex *Cortical Vision* the same applies to CSS in relation to CCS. We have just testified this necessary relation more specifically at the functional and phenomenological levels. But this suggestion implies that a similar relationship between CSS and CCS should be expected to occur also at the structural level. Although to my knowledge no empirical confirmation to this claim seems to exist, important neural pathways seem to link the Posterial Parietal Cortex and the Temporo-Parietal-Junction (the areas suggested to correspond to CSS and CCS respectively). The structural

dependence relation between the two is an empirical consequence of my general hypothesis and suggests the need for further experimental inquiry. Thus, the idea is that whereas the PPC provides a multi-modal sensorial- motor integration through coordinate transformations which in turn is phenomenologically responsible for 1PP, the notion of an external world and an on-line image of the body, the TPJ adds complexity to these contents by computationally *re*-represent (or *meta*-represent) these contents turning phenomenolgically possible the capacity to imagine off-line the 1PP, the external world and the image of the body. *These 'extra' cognitive capacities, in turn, make possible the detachment of the 1PP from the image of the body by 'suppressing' the pre-specified coordinate transformation establishing the coincidence of these two features (still present in FBIs) and, therefore enabling the phenomenological feeling of disembodiment typical of OBEs*. If this is the case, then some strength is given to the notion that FBIs are a constitutive and necessary part of OBEs. As Lenggenhager et al say regarding their study in FBIs:

> "Because the present illusion was neither associated with overt disembodiment nor with a change in visuospatial perspective, we argue that we have induced only some aspects of out-of-body experiences (…). [T]he present data suggest that other mechanisms in addition to conflicting visual-somatosensory information (…) are involved in generating full-blown out-of-body experiences and a more complete transfer of selfhood to an illusory body." (Lenggenhager et al 2007, 1098)

My own conclusions and hypothesis seem to support this dependence/necessary relationship between Full-Body-Illusions and Out-of-Body-Experiences. On the other hand, taking both this relationship between FBIs and OBEs and the characteristics of both illusions, the very proposed taxonomical distinction between CSS and CCS gains empirical and conceptual strength.

Conclusion

At the end of the first section it was stated the main hypothesis that guided this article: that FBIs and OBEs are, respectively, manifestations of (malfunctioning) CSS and CCS.

The point of departure was a suggested taxonomical fragmentation of the concept 'Core-Self' into 'Core-Self *Simpliciter*' (CSS) and 'Cognitive Core-Self' (CCS), with a certain relation of dependence where CCS adds neural, computational and phenomenological complexity to the more basic CSS.

In sections 2 and 3, it was proposed the relation between CSS and FBIs. CSS was presented as the result of evolution's 'answer' to the Stability Problem in complex vertebrate bodies. This solution was identified as a multi-modal and motor integration through coordinate transformation. It was argued that, the result from

ASASE
Full-Body Illusions, Out-of-Body Experiences and what they reveal about the Nature of the Self:
some Functional, Computational and Neurological Considerations

this solution, was the emergence of a first person perspective, a notion of an 'external world' and a image of the body (a 'Body-Image'). The suggestion was that FBIs could be viewed as a result of a misjudged pre-specified coordinate transformations brought about by the illusory body visual-tactile synchronization.

Endnotes

1 As a matter of fact, Björn Merker never uses the phrases 'Core-Self' or 'Minimal-Self' in his writings. Nevertheless, it is quite clear that he is referring to the set of properties usually identified with such concepts.

2 This distinction is a contentious one in current Cognitive Science. Several authors use these notions in very different ways, sometimes inconsistently. Here I am using this distinction as suggested by Gallagher (see especially: Gallagher 2005, 17-39)

References

Anderson, R. et al. (1997). "Multimodal Representations of Space in the posterior parietal cortex and its use in planning movement." *Annu. Rev. Neurosci 20*, 303-330

Brünning, S. & Blanke, O. (2005). "The out-of-body experience: precipitating factors and neural correlates." *Prog. Brain Res. 150*, 331-350

Blanke, O., Metzinger, T. (2008). "Full Body Illusion and minimal phenomenal selfhood." *Trends in Cognitive Sciences 13 (1)*, 7-13

Blanke, O. et al (2005). "Linking Out of Body Experiences and Self Processing to Mental Own-Body Imagery at the Temporoparietal Junction." *The Journal of Neuroscience 25(3)*, 550-557

Ehrsson, H., et al (2004). "That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb." *Science 305*, 875-877

Gallagher, S. (2005). *How the body shapes the mind.* New York: Oxford University Press

Lenggenhager, B. et al (2007). "Video Ergo Sum: Manipulating Bodily Self-Consciousness." *Science 317*, 1096-1099

Merker, B. (2007). "Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. Target article, commentaries and author's response." *Behav. Brain Sci. 30*, 63–134

___ (2005). "The liabilities of mobility: a selection pressure for the transition to consciousness in animal evolution." *Conscious. Cogn. 14*, 89–114

Pouget, A. & Sejnowski, T. (1997). "Spatial Transformation in the Parietal Cortex Using Basis Functions", *Journal of cognitive Neuroscience, 9 (2)*, 222-237

Stevens, J. et al (1976). "Paralysis of the Awake Human: Visual Perceptions", *Vision Research 16*, 93-98

Tsakiris, M. et al (2008). "The role of the right temporo-parietal junction in maintaining a coherent sense of one's body." *Neuropsychologia 46(12)*, 3014-3018

...

8

The Implications of Anomalous Monism
for Intimate Selves

.

Gabriel Levy

Norwegian University of Science and Technology (NTNU) Trondheim, Norway

.

"To achieve intimacy, one must first be oriented to value and seek closeness. Second, one must be able to tolerate and even embrace the intense emotions that are inextricably part of close relationships and be able to share emotional experiences freely. Finally, one must be capable of self-disclosure, mutual reciprocity, sensitivity to the feelings of the other, and concern for the other's well-being" (Collins and Stroufe 1999, 127; cit. in: Bartz 2011, 716)

This chapter examines the relationship between intimacy and the self. I utilize one of the paradigmatic narratives about intimacy, an extremely liquid story from the Bible about the first human couple, Adam and Eve, to explore some of the theoretical baggage in the concept of intimacy. I present a way to make sense of intimacy from the perspective of Donald Davidson's *anomalous monism*. Davidson argued convincingly that, although the first-person has a certain kind of authority in ascribing mental states to him or herself, the concept of the self is irreducible because it is grounded in intersubjectivity (Davidson 2001). Since intimacy requires a sense of ownership of the subjective self, this can lead to problems within a Davidsonian framework, in which the intersubjective level is often prioritized. Taking the subjective perspective seriously from a Davidsonian point of view requires a *radical subjectivity* that even takes the conference ASASE where these ideas were presented into account, since the scholar writing this paper is undeniably also a subject. This was a conference I attended in Lisbon in May 2013 on the subject of the altered self. Academic conferences are places where ideas come together, but -at least as importantly- they are sites where intimacy can take place, where selves meet intimately. In this way, this chapter pursues the subject of intimate selves from subjective, intersubjective, and objective levels.

The quotation reprinted above is an example of a typical attempt to define what we need to achieve intimacy. It is perfectly adequate as far as a strategy. This chapter tries to delve a little deeper into the concept of intimacy, bringing in scholarship not only from the social sciences but also the humanities to help us make sense of it. We should keep in mind that intimacy is something that is better lived than described and analyzed. I am the last person who should be writing a chapter on intimacy— but perhaps this makes me just the right person to do so.

Recognition

If the intimate self is the core self, it is during intimate moments that we feel most authentically ourselves. *Why?* because we are *recognized* by another as ourselves (and we recognize ourselves). Since we are in a constant negotiation about what this particular self is, another way to put this might be that in the company of intimates, we prefer *that* performance of our self. It is not just a performance, since we are not pretending to be ourselves. We get to know ourselves better in the company of intimates because they reflect back a self that *lets things be-* a thought, I hope gets clear during this paper. Fear of intimacy involves a rejection of this dynamic on a number of different levels, but primarily concerns a self for whatever reason *in fear*—vulnerable, under stress, caught in negative paranoid mind-loops, depressed and looking inward at an unrealized inner core- fighting, flying or both.

If we follow the American philosopher of language and mind, Donald Davidson, we should say that the concept of the self is irreducible and has subjective, intersubjective, and objective layers. Davidson saw each of these layers as autonomous, but interacting (anomalously). As noted, he argued convincingly that the first-person has a certain authority in ascribing mental states to him or herself, but the intersubjective level is where these states begin to make sense (Davidson 1984 and 1987). The intersubjective layer is made up of agents in interaction. It involves a relation between two or more embodied agents (and the world). This is the space of second person address. Objectively speaking, we learn about ourselves partly by standing outside ourselves and observing ourselves. This is not something we do in real-time, but "off-line," when we sit back and reflect on our actions and speech. Intersubjectivity is the ground of objectivity, but *subjectivity* is equally as important. Though Davidson does not broach the subject, it may be lack of real engagement with the subjective that leads to self-alienation (see: Kyselo and Di Paolo 2013). We have unique access to our own mental states, our own feelings and emotions, even though we might not have external evidence of these states. These states are real, and they are not necessarily internalizations of something external - as some versions of externalism would argue. These states are recognized in different ways.

In this language, intimacy is a *condition of two selves owning their subjective states in dynamic interaction*. The core of the word *intimacy* involves a spatial metaphor of becoming closer within a space, deriving from the Latin word *intus*, meaning "within." In intimacy what becomes closer in space are these selves, paradoxical selves because they are subjective, intersubjective and objective at the same time. Davidson thought the concept of the self was irreducible for precisely this reason: it neither could be reduced to the subjective self, nor to an intersubjectivity relation, nor to an objectification.

One image that might represent this dynamic of multilayered intimacy is a *holograph*. Holographs can show two or more images depending what angle they are viewed. So more rightly, we can imagine human interaction as two holographic prints facing one another. The holographs would also have a reflective and refractive surface, one that integrates some perspective of the other. The holographs each have a physical makeup, they are materially embodied, they reflect light differently and may be made of different material, and the material is organized differently. These special holographs have an awareness of their material makeup. When the holographs face one another they move back and forth, sometimes confusingly, between seeing some sort of reflection of themselves in the other and seeing the other. From one perspective the holograph sees itself, from another it sees the other. The holographs in this sense could line up in "healthy" and "unhealthy" states.

Another way to understand intimate interacting selves comes from the field of psychoanalysis, which has at times argued for an interpersonal and dynamic concept of the self within the background of what we could describe as an intersubjective world of emotions. Davidson only addressed this to a limited degree in his musings about Freud, probably influenced by his third wife, philosopher and psychoanalyst Marcia Cavell, known for her groundbreaking book *The Psychoanalytic Mind* (Cavell 1993). Though psychoanalysis is more than a form of talk therapy, it reduces psychological problems to language. It assumes that with enough talk, in principle, we can fight our way through neuroses. However, talk is often not enough.

At least what we have in psychoanalysis would be a serious attempt to articulate the subjective self, a topic many disciplines refuse or fear to tread. In many ways the linguistic layers of psychoanalysis get in the way, rather than facilitate, realization and recognition of subjective states. For some people it may help offer language and words to feelings, but to others this language may get in the way, and the recourse to apply direct phenomenology of bodily existence may be a better way to realize and recognize subjectivity.

According to his small "intellectual autobiography" Davidson began his *romance* with Cavell *at a conference* dedicated to his work, at Rutgers (NJ) in April 1984 (Hahn 1999, 60). Davidson was then teaching at Berkeley and Cavell at Purchase, a college in the New York State University system. By June they spent a week together in Venice and by July they were married; what a whirlwind it must have been. Cavell eventually left her position at Purchase to move to Berkeley. It was in these years in the 1980s that Davidson's work shifts a bit; he begins his attempt to account for subjectivity within his own system, ultimately postulated three distinct and irreducible kinds of knowledge, three particular strange loops, necessary for human thought as we know it: "knowledge of the objective world… knowledge of the minds of others… and knowledge of the contents of my own mind." The last is what psychoanalysts like Cavell are especially interested.

Davidson has called his philosophical approach "spinozistic," because he considered himself an ontological monist, while at the same time maintaining conceptual dualism(s) (Hahn 1999, 63). As such, anomalous monism has been used as an antidote to phenomenology that posits naïve forms of dualism or a simplistic relation between knowledge and experience. There are two sides to this. Of course Davidson is a monist in the sense that he does not recognize any ontological distinction between the mind and body, the soul and the body. They are one and the same. But it is also a more thoroughgoing kind of monism that rejects the distinction between self and world. We can't stand outside ourselves to see the world: *we are embedded*. He did not advocate the self-certain subject so often criticized as a mirage of modernity by the likes of Foucault. This is a self of misrecognition, a self that cannot feel intimacy because it cannot realize the dynamic between two actively realized subjective selves.

Cavell uses Davidson as a starting point to try to rethink psychoanalysis from a monist position, one that does not bifurcate the soul or self into a rational self and an emotional self. These should be integrated. She argues, among other things, that though emotions differ from beliefs or desires, they are similarly propositional.

Davidson's musings on psychoanalysis often come in the context of his writings about "Irrationality" – paradoxes, incoherence, Self-deception. The self plays tricks on itself, often precisely because it is intersubjectively entangled. Demons come out, seek to deceive you and others. They do the talking for you. They pull one into ruins and negative feedback loops. Since the intimate self is often a hidden self – conspiracy is often seen as one source of intimacy – these forms of thinking put pressure on Davidson's monism.

Intimacy

I do not want to mislead you to think that intimacy is easy. It is a precarious thing because not only is it a constant effort to recognize and own the subjective self, intersubjective dynamics add an addition layer of complexity where selves inevitably filter through the intersubjective prism. Intimacy breaks down. But do not confuse negativity necessarily with a loss of intimacy. Intimacy requires a degree of negativity, discord, and pain. Without kindness, as we will see, intimacy can verge on torture, because the rawness of torture also involves a close dynamic of realized subjective selves. Negativity is something that must be moved through for a true intimacy to arise. So intimacy needs a form of kind negativity. After all, demons are fallen angels.

Along these lines, the best book we have on intimacy comes from Ziyad Marar, Global Publishing director of Sage publications (Marar 2012). He argues that intimacy is a subtle emergence. We only feel intimate in brief, fleeting, moments— it is a special thing. He enumerates four features, or what he calls "lenses," that he thinks characterize intimacy in its clearest cases. They are *reciprocity*, *conspiracy*, *emotion*, and *kindness*.

All four of these features are useful for my purposes. By *reciprocity* Marar has in mind something like the notion of intersubjectivity. For Marar, intimacy is a dynamic between two people, no more no less. Group solidarity has similarities to intimacy, but Marar thinks the most authentic form of intimacy is a form of self-discovery that takes place between two people.

By *conspiracy*, Marar means that true intimacy involves a form of keeping secrets. Intimates know things about one another that others don't. There is an inside and there is an outside, and those on the inside have privileged knowledge. In the Archaeological Museum in Naples they have the Gabinetto Segreto or *secret room*, which like the Secretum or secret annex at the British museum, houses all the artifacts deemed obscene by 19th century eyes. In other words, in ancient Pompeii where these artifacts were recovered from, these forms would not have been considered obscene. In this sense intimacy should be considered a Gabinetto Segreto, a secret room.

Marar has a whole chapter on "the kiss." Of course the kiss is intimate. Apart from sex, it is the closest thing to two selves experiencing the same thing. In my own field of religious studies dual nature of intimate knowledge comes through resolutely in the double meaning or ambiguity of the verb "to know" as in ancient Hebrew, which has both a cognitive and a sexual connotation. The first book of Genesis has an intense meditation on this connection.

<u>Gen. 4:1</u> וְהָאָדָ֗ם יָדַ֤ע אֶת־חַוָּ֣ה אִשְׁתּ֔וֹ וַתַּ֙הַר֙ וַתֵּ֣לֶד אֶת־קַ֔יִן וַתֹּ֕אמֶר קָנִ֥יתִי אִ֖ישׁ אֶת־יְהוָֽה׃

Gen. 4:1: *wᵉhāʾāḏām yāḏaʿ ʾeṭ-ḥawwā ʾištô wattahar wattēleḏ ʾeṭ-qayin wattōʾmer qānîṭî ʾîš ʾeṭ-yhwh*

Previous to this verse, we have the famous story about Man and Woman's expulsion from Eden because they ate from the "Tree of Knowledge of Good and Evil." The main issue after they ate, the main knowledge they appear to then have, was they could see one another's *nakedness*. What they knew, in other words, was *that* they were naked. God ends up punishing, Adam, Eve, and the serpent. He makes Adam and Eve some clothes before he kicks them out of Eden. It is right after this point we get the verse "The man knew Hava (Eve) his woman, so she 'conceived,'" another word with cognitive connotations, at least in English.

Scholars- I think- have had some trouble tying these two meanings, epistemological and sexual, together, but I think we are helped looking at it through this lens of intimacy. Knowledge involves a kind of intimacy about a subject or about a person. Sex is often an incomplete way to intimacy, and of course one can have sex without intimacy. Nevertheless sexual knowledge is often the most direct way to a sense of intimacy, partly because we usually see one another naked. A similar sense of intimacy can come, for example, with a personal physician who may examine our "private parts." This is someone who has seen your body exposed. Private parts get their name because they are meant to be private.

Along these lines, I would note that there is a deep connection between *privacy* and *intimacy*. Intimacy makes a barrier. This is part of the point made by Richard Schickel in his book *Intimate Stangers* about the intriguing place of celebrity in modern life. Celebrities are paradoxical because we feel like we might know them intimately, but they are in fact strangers. Voyeurism and reality TV are similarly ways to get at a certain form of intimacy without risking anything ourselves (Schickel 1985).

These forms of voyeurism and celebrity lack, among other things, Marar's third feature of intimacy, a mutual closeness and connection through *emotions*. Expanding from Marar, we can say that the intersubjective reciprocity of the first feature of intimacy must extend to the domain of emotions. At the same time emotions can only be recognized and owned from the subjective self.

All emotions do not have to be shared, but there has to be a kind of contract in place to accept each self owning its emotions without that expression refracting and reflecting in the other. *Somehow pause or slow down the holographic flicker*. There has

to be a place for the pure expression of subjective feeling without it being filtered through the prism of intersubjectivity[1].

Marar's fourth feature is *kindness*. As mentioned, without the last one, kindness, he thinks we are left with something like *torture*. He uses the example of O'Brian in Orwell's *1984*, the torturer of the main character Winston, with whom he had many of the trappings of intimacy without kindness.

These points are admittedly somewhat normative. I do not lay these features out in order to make a normative point, about what indeed "authentic" intimacy is and how we can authentically reach it, but to rather gives us some *tools* to think about intimacy and the self. I submit that the intimate self is a kind of altered form of the self, or perhaps it is the other way around. The self, realized in intimacy is a healthy self.

Intimacy is not about losing yourself in the other. There is a romantic conception of the relation between two people that perhaps we get from that original dual monad, Adam and Eve. Both versions of the Genesis story lead to the same problem. In the priestly version (Genesis 1) there is no sense that Eve comes from Adam's bone, but rather Adam and Eve are created in God's image. God is both male and female at once. In the more poetic version of the story (Genesis 3), Eve comes from the bone of Adam. These two are one "flesh." We repeat this idea at marriages.

As nice as it is to loose yourself in a kind of codependency with the other, this feeling should not be confused with intimacy. Initmacy in this respect is a lot closer to friendship than to romance. But of course, sex can often lead to intimacy, as can romance.

Romantic attachment involves a certain amnesia or cognitive dissonance. It leaves aside, at least temporarily, the corners and sharp edges of an individual personality. It does not seem to leave room for the other as truly other. It tries to incorporate him or her into the self. Intimacy feels instead like one is appreciated as a distinctive self at the same time that one is authentically (as much as possible) participating in an interaction.[2]

Intimacy feels like a balance between distinction and participation. According to Kyselo's work we see that the quotation in the epigraph of this text does not emphasize the distinctive self enough. It does not emphasize that intimacy is partly a project of self-realization and ownership.

Davidson's concept of communication fits well in this respect because it allows for the distinction between subjective, intersubjective, and objective layers. Perhaps the subjective self is the distinctive self. The intersubjective self is the participating self. The third type of level is a kind of online-offline perspective on the interaction itself,

the objective level, the third person. The objective self is perhaps the traditional notion of the self, the main character in the ongoing narrative about our lives, given content by the story we tell about ourselves. There has to be a place for the self, given in memory and narrative, not just the distinct and participating self.

Marar's book is based on some of the latest research about intimacy, most of which occurs in the context of social psychology. So far, what I can glean is that in the surveys conducted people say that *vulnerability* and *communication* are two keys to intimacy. And further, that men and women may have different *experience* and *expectations* about intimacy. This is of course a very coarse and general statement. There are also vast similarities. However, it seems clear that natural selection has led to some important sex differentiations given the different roles that men and women play in procreation.

It is a debated subject, but much of the research suggests that women have intimacy as a higher priority (See Laurenceau et al. 1998 and Reis, Clark and Holmes 2004). More precisely it seems women and men have different types of intimacy; we can think of these as two attractors or poles, not necessarily something essential about men and women. In other words there is the "male" mode and the "female" mode. The female mode has intimacy by connecting and working through emotions, talking about them, while men more often describe intimacy (with other men) as doing things together. Culture plays an important role here, but perhaps we can tentatively say that men gravitate to that form of intimacy because the emotional world is more confusing to them as the more autistic sex. Men are often seen as fearful of intimacy as it specifically connects to "responsiveness" – that is men are often perceived as not responsive to the emotional world. Men may fear conflict in emotional world, perceive it as a binary, as a kind of wall, a stop, not something to pass through to the other side. Men fear their emotions, they do not see emotions as a strange loop, but a line. Expectations seem to guide that fear in the sense that men are *expected* to act in certain ways. Control or the loss of it, more than anything guides the fear. In other words, men set up that wall as a way to control emotions. Passing through them involves a loosening of that control.

Intimate Triangulation

Just as Cavell appeared to influence Davidson's work, Davidson influenced Cavell's. There were two important ways in which this took place: first, Cavell began to consider the psychoanalytic session as an important case of triangulation and second, she began to consider the role of triangulation in development.

But by way of bringing this chapter to a close, I want to consider how we need to modify the triangle to account for intimacy. Clearly the notion of intersubjectivity is important. Intimacy involves someone sharing perspectives on one's self; in other

words, part of it includes discovering new things about one's self. But other factors are not addressed in the theory. For example, as noted, an important factor in intimacy is *responsiveness*. Responsiveness is a quality of the interaction, it is a judgment or appraisal of the communication partners reaction. This quality has the kind of nuance and granularity not found in Davidson's original idea. But what about the other aspects of intimacy?

Perspective and Temporality

Marar does not touch on two concepts that I think are essential for understanding intimacy. The first is *perspective*. As I have noted, intimacy differs depending on what perspective--first, second, third (or other?)--one takes on it. On the face of it, this is a statement of the obvious. When looking at an intimate relationship, everything hinges on whether it is I, you, he, she or their pluralized forms that take part in the dynamic.

With this point in mind, we can say that at its most basic level intimacy is about the relationship between YOU and I, the second and first person respectively. Thus the third person, the perspective of science, seem to miss something essential about intimacy. Historical narratives, such as the snapshot in Davidson's biography I presented in this chapter, can capture the first and second person perspectives by referencing documents written from those perspectives at the proper place and time. Similarly, autobiography and memoir obviously have the potential to capture intimacy because they are in the first person. Most historical narrative, however, is written in the third person. Perhaps this gap between first and second person and the third person conventions of academic conferences is what makes academic conferences often so impersonal and unempathetic.

These perspectives, but especially the relation between YOU and I are not static; they are essentially temporal. They involve an active engagement with *memories* about one's self and about the intimate interaction over time. The actors taking part in an intimate connection often approach the interaction at different speeds and reflect on the interaction, through various media, from different temporal perspectives; thus the dynamics of the interaction is some kind of compromise or balance of these speeds when easy, and temporal dissonance when strained. The third person, unlike the first and second, has the broadest perspective on temporality. This is possible by virtue of its narrative constitution. The third person perspective allows for a certain type of *memorializing* that can make or break intimacy.

Narrative Interude:

Billy's wife said she wanted him to have an extra-marital fling before she did, even though it was her idea to have an open marriage. He didn't realize how funny this was

until he talked to Mary about it. He met her in Amsterdam yesterday. She is coming to meet him now at a café. What a lovely comfy place this is, he thought. He immediately broached the subject of some weird, overly enthusiastic emails he sent her in the past month. Billy was bored, a bit paranoid, falling into old habits, they die hard. He really believed that story. He was looking through some old journals from college on his computer on the way to Amsterdam on the plane. He fell for some girls a bit too easily, started acting weird, especially when they started pulling away. Mary mentioned a guy who started to get too clingy to her, though Billy forgot most of the story by now, but she (and all) women don't find that at all attractive. At least not when it comes out of nowhere for no reason. Billy thinks some guys are actually programmed to rush things in that department. Women are the more selective sex.

Billy came to Amsterdam to work on this paper with her. He was also there because they seem to have a nice connection, and he is attracted to her. . . Billy wasn't sure what Mary wanted with him, but he knows he went too far with the romantic implications of their relationship a while back. She is not that interested in romance. She says she does not want to get caught in a symbiotic relationship. She is dating men like crazy. She says it's a way for her to feel like she is alive.

Mary and Billy met at an academic conference that his friend Josh, whom he met a year or two ago, organized. Billy gave a paper. Now he can sort of see that there was nothing unique or special about her coming on to him. This type of thing often happens at conferences. It was no big deal. Billy took it as one because he had been married for 7 years, and was such a sincere, nice guy.

When she arrived she immediately told him she had come to the decision that she was not going to have sex with him this weekend. Billy got the impression they would just have to see how it goes. Things are a bit hazy for Billy today, a day after, because after that point they had such a wonderful day together. He actually felt better after she said that because it took some pressure off him. He wasn't all that aware of the pressure he felt from his wife to have an extramarital fling (so she can).

They worked for a few hours on their paper topic. They are trying to write a paper together. This was the first time they actually tried to come up with something. Mary is brilliant, has a beautiful mind. So it was really nice to get going on the paper.

By the end of the night they were, after going to a few bars, having Sushi (…)

Billy noticed she brought up the subject of kissing a few times, and it was obvious he needed to kiss her at some point. The only question was when. It was a hard task for him because he hadn't kissed anyone aside from his wife for eleven years. They were at a bar here in Amsterdam called Minsk, a kind of Russian inspired theme, they were able to make his current favorite drink. Mary was drinking Mojitos.

Billy remembers saying something like, "I think I am supposed to kiss you." And she said something like "academics are always overthinking things – you need to act instead of think" – well, he didn't need to hear that twice. Billy remembers that a few seconds later he did act. He kissed her. Wow, was it nice to melt into her lips. To open the eyes and be able to stare into her beautiful green-brown beams, some kind of very earthy feeling, something so lovely. When the kissing got a little heavy, she would let out the most astonishingly lovely phrases, in her cute French accent, "Oh my God" she said a couple times. That was so cute.

So you get a sense now, I guess, of how the night went. They talked the whole time. Now he has a room at the Hilton. He is waiting for her to come over and they are going to do the same sort of thing again.

They spent the night together, but they didn't have sex. They were so close and they both wanted to, something stood in the way. Something went wrong by the end. . . She was working for a very long time, and they both realize afterwards that it was the first time they were really disconnected. How ironic that this thing they both really want would actually disconnect them and push them apart.

There was pain the day after. They spent the morning being together in bed, they couldn't get enough. Not having sex was a bit painful, but the pain was worth it. Saying goodbye was quick and chaotic. Billy jumped in a cab outside the hotel so he wouldn't be late for his plane. She was leaving later in the day for a conference where she was a keynote speaker. Billy just wanted to sit in silence, in the chaos, excitement and mystery of the wonderful connection they had made in those days. But the cab driver was an old guy who insisted on talking. Strangely and immediately Billy felt a sense of communion with this guy. Somehow Mary was in him. Billy could find her through him.

There was a sadness and frustration about the limits that his present life put on him (…) Guilt in that feeling of ambivalence. Of course he wanted to see his kids, but he wanted to stay with Mary more.

Adam, Eve and the Serpent:
The state I'm in is an altered state. I'm not the same self that I was when I started. The motions are dangerous, but at the same time I feel a sense of freedom. Monogamy and marriage traps selves in a certain type of relationship. The feelings I have, involve a kind of pull, we both have described it as a certain kind of pull that feels itself centered in the rib cage. Whatever the physiology of this feeling is, it is the likely source for seeing emotions centered in the heart, in the bodily center of the rib cage.

So the loop we're in is an emotional-cognitive one. It is dynamic and changing. It is intersubjective and embodied. I'm trying to figure out where the self fits in there. I've been saying that the *intimate self is a kind of altered self*. It is not the self of

everyday interaction. The intimate self is a vulnerable self, perhaps. It is *a self, taking part in an interaction*. People are closed off from intimacy for this reason. It takes a certain degree of risk to put oneself out there. But it's not the same as romantic attachment. It is not about losing oneself in the other. To the extent that we want to be dependent on the other, this seems to stand in the way of intimacy. There should be space for the other to be a distinctive self (Kyselo *fc.*).

Intimacy emerges out of a balance of interaction, so not all pairings have the potential for intimacy. These are part of the constraints and limitations placed on intimacy. Selves seem to need intimacy. Emotional and physical intimacy put together almost always veers toward romantic codependency. This is a biological attractor with deep roots in our evolution as a species. In some ways we could think that romantic codependency was a clarified form of intimacy. Some of the same elements are there.

The etymological roots of the Hebrew word for *heart*, Lev/Levav may go back to the idea of motion. In this sense it is tied directly to our word emotion. But what is the motion? I'm trying to capture for you that self in motion, not a static notion of the self. Certain motions are healthy- certain are unhealthy. The dynamics of intimacy is a certain type of healthy motion.

Two selves are structured much more in relation to the other, a type of codependency. It doesn't have to be that way, but it usually is. Males and female are attractors to different approaches to intimacy. In the poetic version of the Genesis story previously noted, after Yahweh Elohim forms Adam, he makes a garden, and plants two trees (among others) there, the Tree of Life and the Tree of Knowledge of Good and Evil. He says Adam can eat from any tree except the latter; if he does eat from that tree, he will definitely die.

Adam is lonely, and the animals in the garden don't make him feel any better, so Yahweh Elohim makes him a "helper" (ʿēzer kᵉnegdô - literally a helper fit or in front of him). Eve is created (or "built" - bānā) from Adam's rib or bone (ṣēlāʿ). While Adam (Man) is made from dust, Eve (Woman) is made from Adam. Adam is "formed" (yāṣar), while woman is "built". Then Adam riffs in verse about Eve being part of his flesh and bone and being made from him and for him. The narrator then explains that this is why a man leaves his parents and "cleaves" (dāḇaq) to his woman (or wife or Eve).

The pair was naked (ʿārôm) in the garden but they were not ashamed. The next section begins with a play on this word for naked. It points to the snake or serpent, who- it says- was more "crafty" (ʿārûm, crafty or wise when used positively) than all the other beasts. The word for snake, incidentally, is close to homophonous with

the word for magic (naḥaš). Is the author pointing to some connection between these concepts, between wisdom and nakedness (and magic)?

The serpent convinces the Woman to eat from the Tree of Knowledge of Good and Evil, and subsequently their "eyes are opened," the pair realize that they are naked and they cover themselves. Yahweh Elohim finds out about this when Adam lets it slip that they know they are naked. And in perhaps the greatest example ever of throwing someone under the bus, Adam blames Woman, who in turn blames the serpent. Yahweh Elohim curses all three and kicks the human pair out of the garden, after making (ʿāśā) them some clothes. The reason given for kicking them out is that Yahweh Elohim thinks Adam/Man (this part of the verse ignores the woman) is becoming too much like "one of us" (keʾaḥaḏ mimmennu), too much like a superhuman agent, and wants to avoid that he eats from the tree of life and thereby live forever. Upon leaving the garden the pair immediately start making babies and the genealogy of humanity begins in earnest.

While Adam is made from dust from the earth, Eve is made from Adam's rib, from his chest cavity, from the bone of his bones. This is all from the patriarchal biblical perspective: The idea that Woman was put on this earth to keep Man from being lonely, that Woman is part of Man; some form of ownership. She was basically a piece of him, literally taken out of him–his inner core, his substance.

I've been trying to capture the subjective component of intimate selves. This is as risky as any intimate encounter. This is anecdotal, narcissistic, but it makes me wonder why was Yahweh Elohim so upset with Adam and Eve?

Endnotes

[1] In this context more should be said about "dirty talk," the role of sexually explicit language in intimacy, but I will have to leave that to another time.

[2] I borrow these terms from Kyselo; see Kyselo and Tschacher 2014

References

Bartz, J. (2011). "Serotonin and the Neurochemistry of Intimacy." *Biological Psychiatry 69 (8)*, 716-717

Cavell, M. (1993). *The Psychoanalytic Mind*. Cambridge, MA: Harvard University Press

Collins W., Stroufe L. (1999). "Capacity for intimate relationships: A developmental construction." In: W. Furman et al (eds). *The Development of Romantic Relationships in Adolescence*. New York: Cambridge University Press, 125-147

Davidson, D. (2001). *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press

____ (1987). "Knowing One's Own Mind." In: *Proceedings and Addresses of the American Philosophical Association 61*, 441–58. Reprinted in: Davidson 2001

___(1984). "First-Person Authority," *Dialectica 38*, 101–112. Reprinted in: Davidson 2001

Hahn, L. (ed.) (1999). *The Philosophy of Donald Davidson*. Chicago: Open Court

Kyselo, M. (fc.). "The Body-Social Problem: Towards an Enactive Approach to the Self." *Frontiers in Psychology*

___ , Di Paolo, E. (2013). "Locked-in Syndrome: a Challenge for Embodied Cognitive Science." *Phenomenology and the Cognitive Sciences*, 1-26

Kyselo, M., Tschacher, W. (2014). "An Enactive and Dynamical Systems Theory Account of Dyadic Relationships," *Frontiers in Psychology 5*, 1-11

Laurenceau, J-P, et al (1998). "Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges." *Journal of personality and social psychology 74 (5)*, 1238-1251

Marar, Z. (2012). *Intimacy*. Durham: Acumen Publishing

Pagels, E. (1988). *Adam, Eve, and the Serpent: Sex and Politics in Early Christianity*. New York: Vintage

Reis, H. et al (2004). "Perceived partner responsiveness as an organizing construct in the study of intimacy and closeness." In: D. Mashek, and A. Aron (eds). *Handbook of closeness and intimacy*. Oxford: Taylor & Francis, 201-225

Schickel, R. (1985). *Intimate Strangers: The Culture of Celebrity*. New York: Doubleday

Attention and Redemption of the Self in Weil and Sartre

·

Bernardo Palmerim
Universidade de Lisboa, Portugal

·

Proceeding from Pierre Hadot's conception of ancient philosophy as therapy, William James' notion of the 'sick soul' and certain key concepts in Heidegger's *Being and Time*, we shall look at a case of existential suffering from *Nausea*, where Sartre depicts a redemptive, albeit temporary, mystical experience of alteration of the self, which is prompted by Roquentin's close attention to a piece of music. I will then compare Sartre's description to Simone Weil's autobiographical account of her mystical conversion. Roquentin's attention to an aesthetic object - which he equates with supernatural value as opposed to the nauseating weight of material existence - precipitates a sense of release of the self (non-self). We can find a parallel to his account, *mutatis mutandis*, in Weil's concepts of gravity and grace.

I will then contrast the notions of temporary and long-term alterations of self by comparing these concepts to William James' distinction between gradual and sudden spiritual unification (*lysis* and *crisis*). This difference in forms of spiritual healing will open the way to an explanation of two types of attention: aesthetic vs. religious. Roquentin's criteria for a redemptive work of art will provide us with some clues to the sort of thing artworks can do and how we relate to the world through forms of attention.

Keywords: Attention, Art, Religion, Therapy

Philosophical Sickness and Therapy
I have chosen *conversion* as a topic because it is by definition an experience that effects a significant alteration of self. Hadot says that

"conversion in its religious and philosophical acceptations (…) deals with a change of mental order, which may range from the simple modification of an opinion to the total transformation of the personality."(Hadot 2010)

This conception refers back to the Greek term of *metanoia*, which literally means 'a change of mind' or 'of understanding.' I am going to start by relating the need for *metanoia* to existential anxiety. Sartre's *Nausea* is a depiction of the role of philosophy in existential anguish, as Roquentin, the main character, gropes his way through the meaning of existence. Pierre Hadot tells us that in ancient Greece philosophy was "First and foremost (…) a therapeutic, intended to cure mankind's anguish."(Hadot 1995, 265-6). This is what Roquentin's quest amounts to. Sartre's book tells of a philosophical journey that starts with existence as a source of anguish and ends with what seems to be a release from it.

Conversion likewise presupposes a shift from a divided self, which perceives itself as essentially wrong ('sick') in its relation to the world, to its subsequent unification. This 'inner' struggle of the dramatic attempt to discover one's sense in the world is crucial to the Judeo-Christian tradition, as is patent in the myth of Original Sin which structures its theology, and where the process of conversion is regarded as a *re*unification, a going back to a prior state of ideal, moral health (where health is defined against sickness and is equated with the previous divine order, mythically expressed by the Garden). This split between the self and an idealized world can be taken to heart. As William James puts it, in his discussions of religious experience:

"unhappiness is apt to characterize the period of order-making and struggle(…) will take the form of moral remorse (…) The man's interior is a battle-ground for what he feels to be two deadly hostile selves, one actual, the other ideal."(James n.d., 169)

This difference between the actual and the ideal in respect of existence itself, Sartre makes salient in the figure of the Nausea.

Roquentin has, in his own words, grown sick. At the very beginning of his diary he writes that

"Something has happened to me, I can't doubt it any more. It came as an illness does (…) A little while ago, just as I was coming into my room, I stopped short because I felt in my hand a cold object, which held my attention through a sort of personality. I opened my hand, looked: I was simply holding the door-knob"(Sartre Nausea, 8).

Roquentin's Nausea is the experience of an overwhelming sense of existence – the perception of existence *per se*. With the Nausea, Sartre has created an image for what, in *Being and Time*, Heidegger calls 'objective presence', which Heidegger

claims lies at the base of the Western metaphysics' traditional dichotomy of subject and object (Cf. Heidegger 2010, 60). 'Dasein' is Heidegger's ontological appropriation of 'the self' as the *kind of being* the human being is in our world is hinged on our attention to phenomena and primordially united with the world.[1] (Cf. §12) As opposed to the traditional mind/body and self/world dualisms, "being-in designates a constitution of being of Dasein (…) But we cannot understand by this the objective presence of a material thing (the human body) "in" a being objectively present. Nor does the term being-in designate a spatial "in one another" of two things objectively present, any more than the word "in" primordially means a spatial relation of this kind. "In" stems from innan- to live, habitare, to dwell." (Heidegger 2010, 54) Roquentin is, from a Heideggerian perspective, striving to be-in the world, looking for his 'home' in the world, which he feels cut off from.[2]

Heidegger contrasts 'existence' in the sense of the Nausea (thingness, Cartesian materialism[3]) with what is disclosed in our everyday use of things (handiness). The being of the thing is disclosed in the sense that 'the less we just stare at the thing called hammer, the more we take hold of it and use it, the more original our relation to it becomes and the more undisguised it is encountered as what it is, as a useful thing.' (Heidegger 2010, 69) The being of other beings ('beings' comprising things, people, ideas, language) comes out in our interrelation with them, in our implicated perception of them, which is, after all, *our* normal way of being. In this description of our meeting with things, Heidegger uses the Greek term for 'things', *pragmata*, which is "that with which one has to do in taking care in dealings (*praxis*)." This care, a sense of relatedness, is the key element in *Being and Time* and is what structures Dasein. He calls this kind of seeing 'circumspection', meaning that things only make sense within a totality of reference.[4] Circumspection is a mode of attention to things[5], and indicates that useful things, when used and not simply 'stared at', reveal themselves as participating in a world.[6] This revelation of being-in the world is a disclosure of something that is already in place, namely our being in the world in modes of understanding and participation.

Roquentin's problem is that, in his growing obsession with the Nausea, he stares: as we have seen, he simply holds the doorknob. Heidegger's *pragmata*, on the other hand, are to be used: and before usage lies our already knowing our way about the world. In our everyday dealing with things, "in opening the door, I use the doorknob." (Heidegger 2010, 67) In using doorknobs we do not exclusively pay attention to *what* a doorknob is, as something occupying space in reality – we pay more attention to *how* it works. This is especially true if it fails to work (we might discover it turns the other way). We discover what a doorknob in how it is used. Accordingly, I am trying to suggest that 'in disclosing being-in-the-world, I use the poem.' Roquentin's failure to participate in the world is made salient through his

failure to simply open doors (his 'Nauseas'). Heidegger indirectly describes Roquentin's experience of the salience of everyday things in existence as conspicuousness.

"When we discover its unusability, the thing becomes conspicuous. *Conspicuousness* presents the thing at hand in unhandiness. But this implies that what is unusable just lies there, it shows itself as a thing of use which has this or that appearance and which is always also objectively present with this or that outward appearance in its handiness." (Heidegger 2010, 73)

Roquentin is himself conspicuous, obtruding in his surrounding world, unable to participate in the world by interacting with both people and things because he is stunned at the material *substantiality* of the world:

"The Self-Taught Man is babbling and his voice buzzes gently in my ears. But I don't know what he's talking about . . . I feel this black wooden handle. My hand holds it. My hand. Personally, I would rather let this knife alone: what good is it to be always touching something? Objects are not made to be touched. It is better to slip between them, avoiding them as much as possible.' (Ibid, 82)

 Heidegger warns that an 'analysis which starts with such beings [i.e. things] and goes on to inquire about being comes up with thingness and reality" (Ibid, 67). Roquentin is a philosophical parody of Cartesian materialism.

But Roquentin's problem is not simply one of sensorial disagreeability; he is grappling with a philosophical problem of the meaning of our being in the world, not just things. Roquentin is anxious about the world as such, not anything in particular. As Heidegger tells us, the "fact that what is threatening is *nowhere* characterizes what anxiety is about."(Heidegger 2010, 180) The leveled dreariness of everything around him (and of himself) makes the world salient: "innerworldly beings in themselves are so completely unimportant that, on the basis of this *insignificance* of what is innerworldly, the world in its worldliness is all that obtrudes itself."(Ibid, 181) The Nausea does not matter only as the thingness of things, but as the risk of not being. "What crowds in upon us is not this or that, nor is it everything objectively present together as a sum, but the *possibility* of things at hand in general, that is, the world itself." (Ibid) How to be or not to be in the face of the multiple possibilities of fate is also a source of anxiety for Roquentin.[7] It is nothing definite that assails us, it is the very way in which we do not belong as ourselves to the world, in participation: "what anxiety is anxious for is being-in-the-world itself." (Ibid, 187)

Roquentin's Nausea is a symptom of the philosophical disease of Cartesian dualism. Yet the dualism is itself is born out of philosophical inquiry. Moreover, in

the ancient perspective of philosophy brought out by Hadot, a counter-philosophy is needed to dispel these philosophical thoughts. This is, to a large extent, what Heidegger's 'destruktive' method is about. Yet theoretical discourse does not summon us to actually participate in the being of things, it only argues against other sets of arguments. Poetic language, Heidegger claims, is the right therapy: 'The communication of (…) the disclosing of existence can become the true aim of "poetic" speech.' (Ibid, 162) By focusing on expression over informative content – on *how* instead of *what* - poetic speech gives emphasis to the *way* things are, not only to substantiality, thus disclosing being-in.[8] In the same way that we instantly understand the 'signs' of the world – ""Initially" we never hear noises and complexes of sound, but the creaking wagon, the motorcycle" (Ibid, 163) –, neither do we generally fail to understand the being of things except through an abstractive effort of the imagination. The Nausea is not an insight into the nature of reality, but an abstraction of the actual, which already includes an understanding of our dealings in the world: 'It requires a very artificial and complicated attitude in order to "hear" a "pure noise."' (Ibid, 164) Thus philosophy invents distorted thought: "Gaining phenomenological access to the beings thus encountered consists rather in rejecting the interpretational tendencies crowding and accompanying us which cover over the phenomenon of "taking care" of things in general". (Ibid, 67) To find being is to let things be; and this might mean not thinking 'philosophically', or at least like Roquentin has been thinking.

Inspiration and Grace

When at the end of the book Roquentin is mystically uplifted by a jazz melody, the weight of existence is extinguished. This episode marks a redemptive experience of alteration of the self, prompted by Roquentin's close attention to a piece of music. We have just seen that paying close attention to things provoked existential anxiety in Roquentin, yet at this point attention to a thing made of language, a song, seems to be palliative. This is also framed by his attention. Notably, his decision to attend– 'I'm going to pay attention because, as Madeleine says, I'm hearing it for the last time' – is bound to an awareness of his existential condition. This is why the objective presence of the world becomes a source of anxiety; it is proof of our condition as mortal creatures, as perishable *things* among other things. At the same time, however, this very condition determines us as *Dasein*, the kind of being of human beings.

So Sartre's allusion to finality and death frames the attention he grants the tune, and this pulls him out - in *contrast* to his own sense of self as a mere thing - from the thingness of the world.

"A while ago I was certainly far from swimming in beatitudes (…) Thoughts of (…) my wasted life. And then, still further down, Nausea, timid as dawn. But there was

no music then (…) All the things around me were made of the same material as I, a sort of messy suffering (…) the very existence of the world so ugly that I felt comfortable, at home." (Sartre Nausea, 113)

Listening to the tune, Roquentin is absolved from his *thingness* – the beauty he describes is so essentially different from everyday thingness, that it provides Roquentin with a *counterpoint* for objective presence. His experience of self is altered through his understanding of the tune as a beautiful thing. Turning from the actual to the ideal, both he and the world have become less ugly. Where there was only existence, now there is also value. In this sense, art is the creation of value in existence.

The tune has triggered a change by having been able to sequester Roquentin from his daily routine: "like a scythe it has cut through the drab intimacy of the world (…) it has taken us unawares in the disorder, the day to day drift"(Sartre Nausea). It has forced him into a change of perspective, jolting him from his usual mood, and framed within the concept of conversion as Hadot understands it – philosophy, a form of adult education: 'philosophy itself has always remained essentially an act of conversion (…) philosophical conversion is an uprooting from and a break with the quotidian, the familiar, the falsely "natural" attitude of common sense' (Hadot 2010).

Recapitulating, the thought of death led Roquentin to really attend to the song; the song, in turn, enraptured him, seized his attention because it presents something completely out of the ordinary, and was thus able to produce a notable change in his sense of self: "And I am ashamed. A glorious little suffering has just been born, an exemplary suffering. Four notes on the saxophone. They come and go, they seem to say: You must be like us, suffer in rhythm." (Sartre Nausea, 113)

 It is his implicit comparison of his habitual (actual) perception of the world and self as fundamentally ugly to the beauty he perceives in the tune that generates a humble desire to imitate this (ideal) beauty. He falls in love with the tune because, as he puts it, it is *inexistent*: 'It does not exist because it has nothing superfluous: it is all the rest, which in relation to it is superfluous. It is. And I, too, wanted to be. That is all I wanted; this is the last word.' (Ibid, 114) But this ideal purity that stands out against Roquentin's usual existence by no means ceases to be *real*. There is a transition from a focus on existence to being.[9]

The melody can act as a counterweight to 'existence' conceived as the thingness or 'substance' of things and people in the world because the focus has shifted from *what* to *how* things are. The ('being of' this) song and Roquentin, the attentive listener, suffer in rhythm. That portion of time *is* in a particular way. In listening, he is 'inspired', 'possessed' by the song. When Roquentin asks Madeleine to play

the record again, he begins to muse about the creator of the song. 'I no longer think of myself.' In other words, he ceases to be the subject who is the object of his own thoughts. He no longer posits himself as a self-reflexive thinking subject, a *res cogitans* thinking *about* the world. This is, in turn, why his sense of redemption is linked to his (and the tune's) inexistence. Since the tune has already been described as inexistent, Roquentin's disappearance as a subject depicts his turning into the same mode of being as the tune: the separation between subject and object disappears through *participation*. "I am in the music." (Ibid, 18) In other words Roquentin, who had 'nauseatingly' equated his self with existence, 'is' by not being himself – i.e. by (phenomenally) being the tune. What he perceives, in the broader sense of understanding, during this time in which he is attentively listening, is the tune. During that time – the episode, which I had initially referred to as an "alteration of self" – the phenomenal fact is that Roquentin, is the tune. Truly paying attention to a thing is phenomenally becoming that thing. The picture of beauty as having the capacity alter one's essential self is an extension of the classical picture of inspiration, of Platonic origin and largely propounded by the Romantics. Beauty is captured by the seer and conveyed in poetic language, which is regarded as having an ability to communicate with our essence, our 'true self'. Objects of art are regarded as doorways to the ideal realm, as *icons* for participation. The Muses descend to inspire by taking the seer up to commune (*communicare*) in the idea of *Beauty*, who in turn embodies the idea in linguistic representation, or imitation (*mimesis*). The oracle is traditionally the person who can attend to the transcendent (or listen to the ineffable) and convey that message in language. Within this model of mediation, artworks are icons and artists are oracles. Artistic inspiration finds its analogue in the religious doctrine of grace. Simone Weil describes the same movements between realms, "Creation is composed of the descending movement of gravity, the ascending movement of grace and the descending movement of the second degree of grace', and 'Grace is the law of the descending movement." (Weil 2002, 4) Simone Weil's thought is fundamentally structured by a basic dichotomy of gravity and grace.

"All the *natural* movements of the soul are controlled by laws analogous to those of physical gravity. Grace is the only exception."(Ibid, 1).

The *grace of God* is the ideal counterpoint to the actual - the inspiration that comes down to redeem creatures. Curiously, the conceptual equivalence of the models of *conversion* and *inspiration* is visible in the *Nausea*, where the story of an atheist being inspired by a song bears the same structural similarities as that of a conversion, a *metanoia*.

Weil's mystical experience that resulted in her conversion to religion is also related to attention to an aesthetic object, namely the poem "Love (III)" by George

Herbert.[10] In a letter, she writes, 'I learned it by heart. Often, at the culminating point of a violent headache, I make myself say it over, concentrating all my attention upon it and clinging with all my soul to the tenderness it enshrines. I used to think I was merely reciting it as a beautiful poem, but without my knowing it the recitation had the virtue of a prayer. It was during one of these recitations that, as I told you, Christ himself came down and took possession of me.' (1951, p. 27) *Mutatis mutandis*, both experiences exhibit the same elements, framed by inspiration and a sense of spiritual healing: contemplation of ideal beauty, shameful humility, and a suffering that opens way to the presence of a joyful love. The major difference lies in what *follows* from the experience: Roquentin decides to write a book (he is 'inspired'); Weil becomes a Christian (she is inspired by the grace of Christ). Somehow both their lives gain sense, are reunited with existence through their respective experiences. But while Weil gains a belief in God, Roquentin does not. Weil's inspiration is explained by, and thenceforth secured by, grounded on, the example of the passion of 'Christ', which is expressed in the poem. She participates in a community of believers and enters into a way of life that is explained by 'Christ'. Roquentin's only wish and recourse, on the other hand, is the verb 'to be'. That is all he desires, 'to be'. Weil has a cause which provides her with a justification for existence, Roquentin has only effects to describe the world with, no concept to turn his mind to except the song. The song makes him feel right in the world.

Conversion
Both underwent a *change*, which presupposes a development in time. To explain the therapeutic unification of the self, James resorts to a medical analogy. 'The older medicine used to speak of two ways, *lysis* and *crisis*, one gradual, the other abrupt, in which one might recover from a bodily disease.' (James, 183) He pursues this analogy of conversion by relating it to two ways of searching (or desiring): by *volition* and by *self-surrender*. He gives the example of trying to remember a forgotten name: usually we make an effort to recollect, but if that fails sometimes it is best just to

"give up the effort entirely, think of something altogether different (…) some hidden process was started in you by the effort, which went on after the effort ceased, and made the result come as if it came spontaneously." (Ibid, 203)

As in our examples, repetition (*lysis*) leads to *crisis*: Roquentin heard the same record every evening, Weil had read the poem many times: only later did these objects present themselves and thus the world under a new light. But for a *crisis* to take place, there must first be a search, a desire-for, and so *lysis* always precedes *crisis*:

"there are two things in the mind of the candidate for conversion: first, the present incompleteness or wrongness, the "sin" which he is eager to escape from; and, second, the positive ideal which he longs to compass." (Ibid, 209)

The poem and the song provide a token of this ideal. Roquentin summarizes the three ingredients necessary for the prospective work of art that is the key to his redemption. "It would have to be *beautiful* and *hard* as steel and make people *ashamed of their existence*." (Sartre, Nausea, my italics) It also operates within the same mode of philosophical therapy: conversion as expressed in the *models of inspiration and grace*, where special forms of language and of attention – which we may call poetic language and accordingly poetic attention - are viewed as the healing means. I shall talk about (Roquentin's) song and (Weil's) poem as things of a particular kind. These things are no ordinary objects: they are overtly linguistic, and do not serve as a means for anything else than our participation in their being as instantiated by our attention. They are the product of a craft and require what is signified in 'attention'. The criterion of beauty expresses the object's capacity to mesmerize and continue to do so after Roquentin had conceded it his attention. In this sense, what we call 'beauty' is, more often than not, a *retrospective* description of how an object has effectively secured our attention.

The criterion that his story be, like the tune, as 'hard as steel' to withstand the test of time and stay 'the same, young and firm' is, from the point of view of a crafted thing, a property of recorded language. As Plato and Derrida have insisted, writing - as opposed to speech - can be repeated on different occasions.[11] The very *fact that* repeatability is a quality of 'written things' means that this kind of thing can also convey an *idea of universality*, which is what leads Roquentin to say "the melody is absolutely untouched by this tiny coughing of the needle on the record." (Satre, Nausea, 114) We can read the character of universality in the form of poems because they bear the logical form of tautology: the words remain the same and yet we interpret the same poem differently on different occasions. That is, just as in tautological statements different propositions can be replaced for others 'meaning' the same thing, the poem, in its fixed form, is open to different interpretations (propositions referring to the same subject), and a 'final meaning' that would break the circle is never supplied.

In trying to reach out to a poem, it just slips away and its indeterminacy and ambiguity keep calling on us to describe it. Its 'meaning' can never be supplied because it is writing, not speech: there is no one there to reply. This has the same communicative structure as prayer: the answer that one awaits *has* to be ineffable. Because God and poems do not speak (at least in any ordinary way), they make *us* speak. God and poems do not 'speak' because they do not referentially 'mean' – again, they do not offer a what but a how: a way of being. And no matter what we

say, our actual 'tiny coughing' does not touch the ideal melody. To call something 'a work of art' (we might also allegorically call a person or an action this) is to grant an insignia of perfection: it is 'hard' because it is *worthy* of repetition *because* there is nothing that we feel should be replaced in it ('ashamed of') – it is 'beautiful'. (So the three criteria hang together, in self-referential tautology). Art stands out on account of how it was done. And if nothing is out of place, then the world does not become conspicuous. When attention is circumspect there is being-in. In this sense, art makes us feel at home, because we are called upon to *make* a home, by granting care. But works of art are not always initially taken as expressions of home: more often than not their ambiguity means we have to interpret them, see how they reflect our own discourse about things. This is a sort of work and requires thinking: *lysis* always precedes *crisis*.

As we have seen, the ideal superlative that is as hard as steel offers a counterpoint in regards to which Roquentin feels belittled, in shame and humility.

"The Negress sings. Can you justify your existence then? Just a little? I feel extraordinarily intimidated." (Sartre Nausea, 115) –The ideal awakens his shame because it brings his existential condition into perspective: he is a creature born into a world where beauty exists. His shame is his sense of the distance between his actual self and his ideal self as participative "being". As Simone Weil puts it,

"The wrong humility leads us to believe that we are nothing in so far as we are ourselves – in so far as we are certain particular human beings. True humility is the knowledge that we are nothing in so far as we are human beings as such, and, more generally, in so far as we are creatures" (Weil 2002, 129).

Roquentin's shame is the result of the object's beauty opening his self to a humility that forces the disappearance of his ego, in what Roquentin feels as a desire "to drive existence out of me (…)to purify myself"(Sartre Nausea). But his desire to surrender his self is already the result of his heeding to the song.

Humility, Weil tells us, is what defines contemplative attention. 'In the intellectual order, the virtue of humility is nothing more nor less than the power of attention.' She also provides an example of music.

"When we listen to Bach or a Gregorian melody, all the faculties of the soul become tense and silent in order to apprehend this thing of perfect beauty (…) The mysteries of faith are degraded if they are made into an object of affirmation and negation, when in reality they should be an object of contemplation" (Weil 2002, 129).

As God for the believer, the mystery of beauty is contemplated, not explained. A major difference regarding our use of language and our sense of 'being' is at stake here, since propositional statements (the mode of explanation) do not enter into the

contemplative mode of attention. When we explain things we use the normal referential function of language where a subject is described by a proposition: x is y. But in contemplation our attention is focused solely on the thing or idea, not further descriptions of it, whereby the phenomenal occurrence is simply that x *is*. Roquentin's desire is this, just to be.

Yet trying to approach the ideal point ('to be', or in Weil's case, to be-in or be–with God) is not a product of effort – the rearrangement of dispositions of the self must be left to the quasi-automatic work, as it were, of the *psyche* itself. Weil also insists that if you strive too hard to understand, to find a solution, you will not find it:

"We do not obtain the most precious gifts by going in search of them but by waiting for them. Man cannot discover them by his own powers, and if he sets out to seek for them he will find in their place counterfeits of which he will be unable to discern the falsity." (Weil 1951, 62)

The effort lies precisely in being able to wait ('contemplate'), for things to make themselves evident. This seems to presuppose an attentional (negative) effort of silencing working in tandem with (positive) critical thinking. The silent part is attitudinal[12]; the discursive part is thinking through the solution of problems. If the doorknob does not work we try to find out what the problem is and fix it. The *attempt* to fix problems is interpretation. But solutions come of their own accord. The model of inspiration is, from this perspective, a causal picture of how thought arrives (out of nowhere). In the same tone as Weil but speaking of conversion, James says that

"Where, on the contrary, the subconscious forces take the lead, it is more probably the better self *in posse* which directs the operation." (James, 209-10). These examples help us see how both pictures of grace and inspiration, in what partakes to *crisis*, depend on a concept of self-surrender. Both descriptions presuppose that waiting in self-surrender will *automatically* direct us towards what we seek; in other words that the attitude of contemplative attention clears the way. For Weil, a believer, this kind of attention is equal to prayer (communion with god): "Absolutely unmixed attention is prayer" (Weil 2002, 117). We can also see how *lysis* and *crisis* as transformative processes of the self are ultimately grounded in two interspersed and interrelated ways of directing our attention: by searching and waiting.

Like James, Weil says that contemplation of the ideal of itself guarantees goodness: "It is impossible to do harm to others when we act in a state of prayer" (Ibid, 126). Conversion (the religious desire that requires a *moral* transformation of the self) entails a change of character, and as such the long-term fixedness of desire for ideal goodness within the self. Self-surrender is taking the self ('I') out of the contemplative process. For Weil, self-surrender occurs through humility, in what

she calls the transformative process of 'decreation'. James says the same thing when he relates conversion to self-surrender: "yielding (…) is giving one's self over to the new life, making it the centre of a new personality" (James, 210).With decreation Weil points to a use for that 'glorious suffering' which is imparted by perceiving an ideal that triggers shame and praise in us. *Crisis*, in these examples, is the result of having surrendered the self to the contemplation of beauty (or goodness or truth). For Weil decreation is a systematic method, a form of continuous moral training in humility to suppress egotistical impulses. But James and Weil are describing a religious change, which is pinned on morality, on rules regarding attention to others. This negative impulse, however, is the same that underpins Roquentin's sense of non-self: the experience, from the point of view of *perception*, is the same. As a whole, the allegory of the 'non-self' is essentially a description of contemplative attention: we are 'nothing' because we have turned our attention to a thing or another person, because we phenomenally participate in that being.

At stake are two kinds of person: Roquentin *desires to imitate beauty* in a work of art; Weil *desires to imitate Christ*. These, in turn, relate to two types of attention: aesthetic and moral; or in other words, attention to things vs. people. Roquentin's redemption hinges on what he may be able to achieve in his book:

"perhaps, because of it, I could remember my life without repugnance (…)And I might succeed —in the past, nothing but the past—in accepting myself" (Sartre Nausea).

He seems to view his book as a medal that may award him admiration and immortality ("'there would be people who would read this book and say: "Antoine Roquentin wrote it"'"). His redemption would be attained by perfecting his artistic merits; he would desire, as W. H. Auden says *In Memory of W. B. Yeats*, that *(b)y mourning tongues / The death of the poet was kept from his poems*. But there is no hint of a desire to be good, to change his personality or attitude towards people. Auden also says that *poetry makes nothing happen: it survives (…) A way of happening, a mouth.* With poetry nothing changes in the world, only in descriptions of it. The difference lies in our belief in descriptions, on what we take as valid descriptions of the world: but the difference is not that between *metaphor* and *fact*[13]. Conversion in the broad sense of *metanoia* simply indicates how we are always ultimately open to paradigmatic shifts in our understanding of the world.

Creative attention
Inversely, because her attention is extended to others, Weil sees no difference between bringing life ('being') to a book or to a person:

"true creation means self-loss. We do not perceive this truth, because fame confuses and covers with its glory achievements of the highest order (…) Love for our neighbor, being made of creative attention, is analogous to genius. Creative attention means really giving our attention to what does not exist. Humanity does not exist in the anonymous flesh lying inert by the roadside. The Samaritan who stops and looks gives his attention all the same to this absent humanity, and the actions which follow prove that it is a question of real attention" (Weil 1951, 92). Creation is not only the backbone of the poetic, but also the religious mode.[14] What is similar between Weil and Roquentin is that both have turned (*metanoia*) their attention from themselves to the 'outside', surrounding world (whence the self is self-evidently disclosed as always already having been in the world).

Yet they are different in that Weil turns to the grace of 'God', Roquentin to the inspiration of writing. Her turning to God, with its moral implications towards others, implies a change in conscience. Of itself, it is silent, for it is above all an attitudinal change. Religious change is an assent to an entire grammar that is theology; but this structures a way of life, which is based on the belief of a community. Here *communicare* is too based on communion and community. As a tradition and a faith, religious conversion offers a long-term structure that explicates the world. Aesthetic changes are more of a temporary kind since the change is not *necessarily* of the moral self but attains to the perception of particular things crafted out of language. The song has managed to take 'drab' out of the 'drab intimacy of the world.' Intimacy, closeness to the world, has been restored, cured, cared for. The song turned Roquentin's doorknob, as the poem did for Weil. Their turning around, their *metanoia*, meant that they have stepped into a new world, with new possibilities. For Weil, her new life makes sense within a certain set of retrospective descriptions. The cause for everything is God; and events are redescribable within that way of life. For Roquentin the future is unclear, for there is no background, universal structure to explain his existence. He just 'is' as opposed to just 'existing'. His contemplation fleets in time to different particulars, and cannot always come back to the "God is" of Weil. His future is now a sea of possibilities but which now offer hope and not anxiety. It lies open before him, but he may fail again, fall out of sync with the world, and there will be nothing behind him to secure his writing but the world he makes. But before these things opened their doors, they suffered in rhythm and felt humbly ashamed. We do not look at our better or true self as a source of true answers; but the best answers come from our looking *for* our authenticity. Art has an important role to play in calling upon us to turn towards the anxiety and shame of being-in-the-world (which we otherwise might allow to remain dormant in our everyday routines and familiarity with the world). Being lost and searching precedes the possibility of our finding ourselves again in different possibilities, in a fate all our own.

Endnotes

[1] Strictly speaking, we can no longer speak of self in Dasein, since the term also aims to both resist and integrate the concept.

[2] 'I suffer in my wounded flesh which turns, walks, I walk, I flee, I am a criminal with bleeding flesh, bleeding with existence to these walls. I am cold, I take a step, I am cold, a step, I turn left, he turns left, he thinks he turns left, mad, am I mad? He says he is afraid of going mad, existence, do you see into existence, he stops, the body stops, he thinks he stops, where does he come from? What is he doing?' (Sartre Nausea, 68)

[3] Sartre's reference to Descartes is explicit in spite (and as a source) of Roquentin's confusion: 'I am. I am, I exist, I think, therefore I am; I am because I think, why do I think? I don't want to think any more, I am because I think that I don't want to be, I think that I (...) because (. . .)' (Sartre Nausea, 67)

[4] So (the being of) a hammer initially and for the most part makes sense within (the being-in of) a workshop. 'The total relevance which, for example, constitutes the things at hand in a workshop in their handiness is "earlier" than any single useful thing, as is the farmstead with all its utensils and neighboring lands.' (Heidegger 2010, 84) This way of thinking turns the usual way of conceiving meaning around. Here 'hammer' only means because of a context of meanings: but these essentially originate in a world where we use hammers. Here we can clearly see the maxim of phenomenology, 'To the things themselves!' We can also recognize 'Wittgenstein's 'forms of life' in *Philosophical Investigations* are a close analogue, where the example is no longer things but people. Another Wittgensteinian metaphor for bringing out circumspection is 'language games', where words are things and grammar the totality of reference.

[5] 'Our absorption in taking care of things in the work-world nearest to us has the function of discovering (…) and with a varying attentive penetration.' (Heidegger 2010, 71)

[6] 'In the scope of its heedful absorption in useful things at hand, does not Da-sein have a possibility of being in which, together *with* the innerworldly beings taken care of, their worldliness becomes apparent to it in a certain way?' (Heidegger 2010, 72)

[7] 'I don't know what to do in front of the Passage Gillet. Isn't anyone waiting for me at the end of the passage? But there is also at the Place Ducoton at the end of the Rue Tournebride something, which needs me in order to come to life. I am full of anguish: the slightest movement irks me. I can't imagine what they want with me. Yet I must choose: I surrender the Passage Gillet, I shall never know what had been reserved for me.' (Sartre Nausea, 39)

[8] 'Being-in and its attunement are made known in discourse and indicated in language by intonation, modulation, in the tempo of talk, "in the way of speaking." (Heidegger 2010, 162)

[9] In the terminology of *Being and Time*, this means a transition from the ontic to the pre-ontological.

[10] Love bade me welcome, yet my soul drew back,
    Guilty of dust and sin.
But quick-ey'd Love, observing me grow slack
    From my first entrance in,
Drew nearer to me, sweetly questioning
    If I lack'd anything.

"A guest," I answer'd, "worthy to be here";
    Love said, "You shall be he."
"I, the unkind, the ungrateful? ah my dear,
    I cannot look on thee."
Love took my hand and smiling did reply,
    "Who made the eyes but I?"

"Truth, Lord, but I have marr'd them; let my shame
    Go where it doth deserve."
"And know you not," says Love, "who bore the blame?"
    "My dear, then I will serve."
"You must sit down," says Love, "and taste my meat."
    So I did sit and eat.

"Love (III)", by George Herbert (1633)

[11] I am including musical notation and recording as kinds of 'writing', focusing on the concepts of repeatability and inscription. On these concepts see Derrida's essay "Plato's Pharmacy" (1968), which offers a dialogue with Plato's Phaedrus.

[12] As Heidegger says, the call of conscience is silent. 'The call does not say anything . . . *Conscience speaks solely and constantly in the mode of silence*.' (Heidegger 2010, 273)

[13] To think of the world as groups of atoms would in this sense be no less of a myth, a picture, than to think of it as created by "God" or caused by Big Bang.

[14] '. . . createdness in the broadest sen se of something have been produced, is an essential structural moment of the ancient concept of being.' (Heidegger 2010, 24)

References

Hadot, P. (1995). *Philosophy as a Way of Life.* Malden: Blackwell Publishing

Hadot, P. (2010, May 17). *"Conversion," Translated by Andrew Irvine.* Retrieved from Words Pressed Home: http://aioz.wordpress.com/2010/05/17/pierre-hadot-conversion-translated-by-andrew-irvine/

Heidegger, M. (2010). *Being and Time.* Albany: State University of New York Press

James, W. (n.d.). *The Varieties of Religious Experience*. A Study in Human Nature.

J. Manis (ed). An electronic classics series. An Electronic Clasic Series. Hazleton, Hazleton, PA: The Pennsylvania State University

Sartre, J.-P. (n.d.). *Nausea.* Retrieved April 29, 2013, from Internet Archive: http://archive.org/details/Jean-paulSartre-Nausea1938

Weil, S. (1951). *Waiting for God.* New York: Perennial

Weil, S. (2002). *Gravity and Grace* (1st complete English language ed. ed.). London: Routledg

...

10

The Mutative Self and the Language of Insight in Henri Micheaux's Mescaline Experiments

.

Michaela Hulstyn
Stanford University, USA

.

Abstract: *In this paper I examine the work of Belgian-born poet Henri Michaux, who uses the concept of the mutative self to communicate his mescaline experiments of the 1950s and 1960s and reify his status as the ultimate witness figure. Lacking an operative metaphysical vocabulary for expressing the encounter with the beyond, Michaux develops a secular language for talking about the unselfing that occurs during mind-altering experiences. This self-experimentation allows him to proclaim that synchronic disunity is a natural (albeit highly uncomfortable) state. The experience of mutative unselfing alters Michaux's perception of time and space, leading him to discover unchartered expanses of the self.*

Keywords: self, self-experiments, altered-states, French literature, drug literature, unselfing

Introduction

Can narratives on drug experimentation shed light on philosophic and psychological accounts of the self? The questionable status of the self during altered states of consciousness was already present in John Locke's highly influential account of personal identity in the seventeenth century. In Locke's account, this problem is solved by the existence of an omniscient God that knows the entirety of ones identity, even the parts for which the subject is not fully conscious.[1] This importantly includes the moments of that individual's existence during which he is not "in his right mind," including the time he spends asleep, or drunk, or on the drug mescaline, as in the case of Belgian-born poet Henri Michaux. Nietzsche pronouncing God dead at the turn of the century has become the cultural marker for the shift away from theological-based understandings of the self that posited God in this objective third-person position. Lacking Locke's faith in a God that

could see our full identity, Michaux seeks to understand what happens to the metaphysical part of the self during altered states of consciousness.

Knowing that hallucinatory drugs had long played a role in religious ceremonies in various cultures across the world, Michaux, at the age of 54 began taking mescaline, a naturally occurring psychedelic derived from the peyote cactus, used for thousands of years in religious ceremonies by Native Americans in Mexico and in South America. In the first of his mescaline texts, *Misérable miracle* (1956), Michaux writes that while "les Huichols, les Tarahumaras, et bien d'autres Mexicains (…) allaient à un dieu en allant au Peyotl" [the Huichols, the Tarahumaras and plenty of other Mexicans (…) went to a god in going to the peyote] [2] other gods, like the god of the volcanoes or of fire for example, were not far. (Michaux 1972, 69) Primed for the experience of the supernatural, he writes that the Native Americans had to simply name his god, and he would appear to the drug-taker. Westerners however, undergo a different experience on the drug, and only see "infinite cascades of relativity." He explains,

"Quant à l'Occidental d'à présent, depuis longtemps incroyant aux dieux, et qui serait bien incapable d'imaginer une forme sous laquelle ils seraient susceptibles de lui apparaitre, ce que son esprit saisit, seul dieu qu'il aperçoive encore et qu'il serait vain d'adorer, c'est l'infinie relativité, la cascade qui n'a pas de terminaison, la cascade des causes et des effets, ou plutôt des précédents ou des suivants, où tout est roue entraînante et *roue entraînée.*" (Michaux 1972, 70)

[As for the Westerner today, who for so long has been an unbeliever in the gods, and would be incapable of imagining a form under which they could appear to him, what his mind grasps, the only god of which he is still aware and one that would be vain to worshop, is infinite relativity, the cascade without ending, the cascade of cause and effect, or rather the precedents or the followings, where everything is pulling wheel and pulled wheel.]

Underlining that his own interpretation of metaphysical experience is shaped by historical and cultural assumptions, Michaux compares the outcome of his experiments with the spiritual discourse of the Native Americans he mentions. His reflections are attempts at descriptions of the supernatural, without having recourse to a preexisting religious vocabulary. At certain points in these texts, Michaux acknowledges his fears that there may be nothing objective to be gained from these voyages into the beyond, since the human mind is highly suggestible. Peter Broome summarizes this tension in Michaux's work: "Whatever one's thought, one's visual image of the moment, one's current obsession, it will launch it on the road to the absolute. It has no objective truth of its own to propose, apart from that of a scientific process, and will lend accidentally to any vision" (Broome 1977, 91).

Despite these worries, however, Michaux ultimately posits that altered states are privileged moments of insight on the self. Lacking an operative metaphysical language for what I refer to as the "unselfing"[3] that occurs during mind-altering experience, Michaux develops a vocabulary for communicating this state of insight in a secular age. In the absence of a functional theologically based philosophical framework, Michaux attempts to track down the metaphysical piece of personal identity that persists through altered states. Given his understanding of the self as a witness figure living in an eternal succession of present moments, altered states of selfhood productively change his sense of time and space. The altered state is a monumental event that allows for the development of a narrative of before and after, one that opens up new expanses of the self.

While terms like "mind-altering" are commonplace in discussions of drug culture, especially as related to the psychedelic 1960s, few studies of altered states have considered a connection with philosophical discussions of the self, and what self-experimentation can tell us about the nature of the self during ordinary circumstances. In this paper I examine Michaux's take on the process of unselfing, as described during the self-induced drug experiments that began in 1956 and were carried out through the 1960s. I first consider Michaux's conception of the self during so-called normal states, using terminology from current research in psychology, philosophy and neuroscience to help understand the self-experience that Michaux describes in his work. It will then be possible to examine his take on altered states of selfhood in a synchronic sense, or, the way he describes the experience of disunity in the moment, as well as the effect of the experience in his long-term, or diachronic sense of the self. I also examine the implications of Michaux's work on the development of a metaphysical vocabulary for otherworldly experience in the absence of faith in an omnipotent third party witness.

Implicit in this study is the assumption that studies of writers like Michaux, and more generally speaking, studies of literary texts are valuable contributions to discussions of altered states of selfhood within empirical-based disciplines. It is my view that studies of literary accounts of altered states can contribute to philosophical discussions of the self. Especially given the fleeting nature of altered states, and the impossibility of pairing such highly unusual (and potentially unethical) self-experiments with technology in the hard sciences, the study of these accounts gives us precious insight into the phenomenology of experiences that might be impossible to study in the lab.

To review, in this paper I first explore the altered self in Michaux's work in a philosophical mode, arguing that Michaux experiences synchronic disunity at the moment of the experience, and later interprets unselfing in the diachronic sense as a fundamental mutation. Michaux's experience of mutative unselfing should be

distinguished from other narrative models that account for dramatic change in the self over time, such as the conversion narrative.[4] As opposed to Saul who becomes Paul, Michaux does not claim to become a different person entirely after his experience. Rather, mescaline allows for the illumination of formerly obscured contours of the self. I then situate Michaux in his literary and historical context in order to interpret the metaphysical language he develops for discussing altered states. At the end of this paper, I consider the value of Michaux's self-experiments for relations with the outer world, comparing his interpretations of unselfing with other Francophone writers of the period.

The self according to Michaux
Michaux offers a unique perspective on the question of the self in his writings that fall somewhere between autobiography and auto-fiction. Indeed, in reading Michaux's works one is made aware of the slippery boundaries between reality and fiction, as his narrators constantly wonder about the constructed nature of reality itself. Much like Paul Valéry, who created his *Monsieur Teste* character as a model for his theories on the self, Michaux over the course of his career created a pen name character, appropriately named "Plume" who serves as a testing ground for his ideas on consciousness and the self in writing. In turning to Michaux's texts on altered states of selfhood during drug experiments, we must keep in mind that before these experiences, Michaux was particularly attuned to the constructed nature of self-unity, and saw disunified states of the self as more natural states of being than the usual state of ego-dominion.

As such, before turning to Michaux's exploration of altered states of selfhood, it is important to first examine how Michaux believed the self to function under ordinary circumstances. This question is especially difficult in studying Michaux; he famously proclaimed the eternal escape of the self when faced with attempts to pin it down – in language, in ancestry, in literary influences and otherwise. Broome writes,

"There are enough warning-signs in Michaux's work to make any biographer uneasy. Not only has he always refused to entertain a facile relationship with his public, retreating behind the wall of imagination and slipping from the grasp in bristling metamorphoses, but he has never felt any solidarity with his own life" (Broome 1977, 1).

Indeed, there are many instances in Michaux's texts where he claims to "désolidarise" or "break away" from his life, not to mention the extra distance he often sets up between himself and the first-person-narrators of his texts. The theme of the absurd also figures in Michaux's works, and his texts on selfhood often testify to the divorce Camus once proclaimed between man and his life, the actor and his

décor.[5] In Michaux's work, the self is often described as an undifferentiated inner space without justification, identity, project or even reality. Because of this, he and his narrators are often frustrated by the lack of ensemble of life, the eternal documentation of the present.[6]

In his texts on his experience taking mescaline, Michaux stresses that disunity would be our natural mental state, if not for the dominion of the ego, which holds us under its tyrannical control. He writes, "La drogue, qu'on s'en souvienne, est plus révélatrice que créatrice" (Michaux 1966, 33). [The drug, let us remember, indicates, reveals, more than it creates" (Michaux 1974, 22)]. Of the return from one such mescaline trip he writes,

"Maintenant revient le pragmatique, l'utile, l'adapté, l'harmonieux, revient l'ego, ses bornes, son autorité, son annexionnisme, son gout des propriétés, des prises, son plaisir de s'imposer, de faire tenir ensemble, de forcer coûte que coûte. Et cela paraît naturel!"(Michaux 1966, 29) *[Now the pragmatic returns, the useful, the adapted, the harmonious; the ego returns, with its limits, its authority, its annexationism, is possessiveness, its grasping, its delight in imposing, in amalgamating, in forcing at all costs. And it all seems natural! (Michaux 1974, 20)]*

Given these observations, we should understand the self during ordinary circumstances in Michaux's work to mean a minimal degree of unification, as maintained by the ego. The process of unselfing, then, operates as a two-step process and undoes the ego's unifying power. During altered states of selfhood, Michaux documents the experience of unselfing in the synchronic sense, and later reflects on its diachronic implications. That is to say, Michaux first describes a loss of perceived unity at the moment of the experience, and later, in the process of writing, reflects on the meaning of the experience in a narrative of mutation. In the scope of this paper, I focus on the first and last of Michaux's texts on drug experimentation, *Misérable miracle* [Miserable Miracle] (1956) and *Les grandes épreuves de l'esprit* [The Major Ordeals of the Mind] (1966).

Altered states: the "core" self and the "extended" self
While documenting his altered self-experience, Michaux is most concerned with the effect of the substance on his basic or core self. The first level of the experience should therefore be understood as the effect of the experience on the most fundamental conception of mind. While disciplines such as neurobiology, psychology and philosophy account for a form of the core self, their definitions diverge. For instance, while Damasio (2010) writes that the "core self" is responsible for conscious attention to a given object, and emits "pulses" of provisional subjectivity in the field of neurobiology, phenomenologists Shaun Gallagher and Dan Zahavi (2008) argue that the most basic level of the self is a pre-reflective self-

consciousness. This discussion was contemporary to Michaux as well; In *L'être et le néant* (*Being and Nothingness*, 1943) Sartre argues for the existence of a pre-reflective, non-cognitive relation to the self.[7] In this paper, I assume the core self as examined by Michaux to be a reflective state. While Michaux seeks to consciously witness the experience, often drawing or writing notes while taking the drug, his texts are delayed reflections on these changing mental states. The focus on unselfing at the level of the core self should be understood as a reflection on the philosophical problem of synchronic identity, as these experiments temporarily destroy the subject's perceived experience of coherence or unity, and sometimes embodiment as well.

The difference between unselfing understood as a temporary state of disunity, and unselfing understood as total annihilation of the self is aptly illustrated by Michaux's description of his overdose experience. Whereas his other experiments with mescaline are characterized by an altered state that the observer is able to document, Michaux is not able to witness the experiment when he overdoses. During his fourth mescaline trip described in this work, Michaux writes that the overdose experience that caused him to fully "coincide with himself", as he puts it. He writes, "Qu'avais-je fait? Plongeant, je m'étais rejoint, je crois en mon fond, et coïncidais avec moi, non plus observateur-voyeur, mais moi revenue à moi, et là-dessus en plein sur nous, le typhon" (Michaux 1972, 124). *["What had I done? Plunging, I met myself, I believe in my foundation, and coincided with myself, no longer as an observer-voyeur, but me returned to me, and right on top of us, the typhoon."]* This is a different level of unselfing than Michaux had hitherto experienced. In a footnote to this passage, he writes, "Dans ma vie j'essaie (voulant observer) d'approcher le plus possible de moi, mais sans coïncider, sans me laisser aller, sans me donner" (Ibid). [In my life I try (wanting to observe) to approach myself as much as possible, but without coinciding, without letting myself go, without giving myself up.] Michaux attempts to distance himself from recreational drug use with this caveat, giving us an important feature of mutative unselfing: save for this exceptional overdose experience, the core self is never completely lost during these experiences. In this way, Michaux affirms the privileged status of the witness figure, the one who both lives and documents the altered-self experience.

The second level of the experience involves what is referred to, across disciplines as an extended sense of the self. Damasio (2010) calls this sense of selfhood the "autobiographical self" and Schechtman (1996) refers to the "narrative self" in her philosophical work. Both designate a sense of self that is extended through time and that is teleological in nature. The eventual reconstruction of the self after the witness event should be understood within the context of the diachronic personal identity problem. Given that the experience happened, how does Michaux think about

events past and future? In the work of Michaux, the construction of a narrative for this self-experience occurs *after* the witness event, during the writing process. Of course, the distinction between the two steps that make up mutative unselfing may be difficult to discern, as they are mutually dependent on one another. Michaux presents his experiments as productive mutations in his temporal and spatial self-experience.

According to Michaux this leads to a state of insight that does not disappear once the substance has physically left his system. For example, coming down from the drug, he writes, "Au sortir de la Mescaline on sait mieux qu'aucun bouddhiste que tout n'est qu'apparence. Ce qui était avant, n'était qu'illusion de la santé. Ce qui a été pendant était illusion de la drogue. On est converti" (Michaux 1972, 80). [Leaving Mescaline, one knows better than any Buddhist that there is nothing but appearance. What was *before*, was nothing but illusion of health. What was *during*, was illusion of the drug. One is converted.] The word "conversion" here is not used traditionally. Michaux does not seem to mean he is converted *to* a transcendent absolute, not even the Buddhist philosophy he references. Rather, he seems to mean there is no "going back" to one's former self once one has experienced the fragility of reality. Michaux insists that the self is changed by these drug experiments. This state of insight is "lightning on a dark night, all illumination" he writes, "but it is a lightning that lasts" (Michaux 1974, 166).

At the end of his last text on mescaline, Michaux asserts emphatically that an inner space is opened up within the self thanks to these experiments, and persists after the experiment is over. He writes,

"Mais dira quelqu'un, ces mondes sans objet, au-delà d'un objet, amour sans objet, contemplation sans objet, n'est-ce pas de la fume, dont il restera moins encore de la fumée? Une étendue nouvelle, un fond surcreusé, depuis en partie comblé, mais non annulé, subsiste après l'expérience et peut-être à jamais, non pas tout uniment non plus. " (Michaux 1966, 207)

[But, someone will say, aren't these worlds without object, beyond an object, love without object, contemplation without object- aren't such worlds so much smoke, leaving behind even less than smoke? A new expanse, a profoundly hollowed depth, which may afterward be partially filled, but not annulled, subsists after the experiment and perhaps forever, though not quite uniformly either. (Michaux 1974, 170)]

Michaux represents the experience as an illuminating rupture that opens up "new expanse" within. And yet, each moment of unselfing always involves its corresponding return. While the rare state that one reaches during the drug experiment may lead to valuable insight, it cannot be maintained. The return to the state of ego-dominion is inevitable, but perhaps even more importantly, ego-

dominion is more comfortable than the state of illumination. Michaux goes as far to compare his state of illumination to an "occupation." He writes,

"Fini! Les heures de l'occupation sont passées. A présent il est seul dans son cerveau. Admirable impression. Jouissance intime, de toutes peut-être la plus intime, si discrète pour être presque identique au 'moi' collant indissolublement à l'être en vie, et dont l'absence est une essentielle, indicible, incessante catastrophe" (Michaux 1966, 17). *[Finished! The hours of alien occupation are over. Now he is alone in his brain. A wonderful feeling. Inner enjoyment of rights, perhaps inmost of all, so private as to be virtually identical with the 'self' clinging inseparably to the living being, and whose absence is a basic, inexpressible, unremitting catastrophe" (1974, p. 9).]*

What this vocabulary suggests, despite Michaux's insistence on the ephemeral and provisional nature of selfhood in his work, is that even during extreme experiences of unselfing, someone is *there* to witness the unselfing event, to resist total self-abandonment or self-loss, someone to recount the experience upon his return. The metaphysical piece of the self – the witness position – is never fully lost during mutative unselfing. Moreover, the witness's viewpoint is improved by the experience, illuminated as by lightning on a dark night.

Altered states and metaphysical insight: Michaux in his literary-historical context
Much like his predecessors Baudelaire and Rimbaud, Michaux seeks to transcend everyday reality in order to access untapped expanses of human consciousness. With his psychedelic experiments and poetic writings thereupon, Michaux enters into a tradition of both Belgian and French avant-garde writers who positioned the poetic imagination as a site for revelation, naturally or unnaturally achieved. As early as 1836, the term *Supernaturalisme* appears in French when Nerval translated *Faust.* The term was then associated with German romanticism and referred to the power of revelation and the faculty of imagination, carrying the connotation of "above or beyond." In 1855, Baudelaire uses the word *surnaturalisme* to mean a heightened sensory experience and an internal state of revelation. Baudelaire's usage of the term marks a break with the traditional theological notion of supernaturalism, referring rather to an excess of experience. Influenced by De Quincey before him, Baudelaire was interested in the effects of wine, hashish and opium on the faculties of perception, writing on the subject in *Les paradis artificial* (1860). Like Baudelaire, Michaux would worry about the potential for solipsism – that the drugs represented a "cheap" route to revelation, and that what was revealed was only a distorted self-mirror, not a glimpse of the beyond. The Decadent movement further revived an interest in transcendence and the occult with Rimbaud embodying the figure of the poet-as-visionary. Of this influence on Michaux, Rigaud-Drayton writes, "Constructing the poet as a 'Seer', like Rimbaud, he conceived of poetry and painting as modalities of the visionary – and the

unlapsed. Nature and its poetic, visionary language was to be found locked inside the self, as much as outside it" (Rigaud-Drayton 2005, 5). While there are important differences between Rimbaud's poetic project and Michaux's mescaline experiments, (the centrality of aesthetics to Rimbaud and the supposed objectivity of Michaux's project) there are clear overlaps as well. Michaux's scientific/mystical explorations of the unknown recall Rimbaud's famous quest for insight via *dérèglement de tous les sens.*

In this lineage, the surrealist movement intended to transform modern consciousness in a practice or experience of the world leading to an altered state through art-as-knowledge, believing "that art arises from the imagination as a revelation: a form of insight-oriented knowledge." (Rabinovitch 2002, 8) While he shared their desire to reduce the control of reason as a means of achieving insight on the self and the world beyond phenomenal reality, Michaux refused to officially join any organized avant-garde collective of the period. He was influenced by figures like Lautréamont, who he discovered in 1922, a *poète maudit* claimed by the Belgian avant-garde over twenty years before the French.[8] Michaux had also read Aldous Huxley's *The Doors of Perception* (1954) but he thought Huxley's approach was too moderate – both in the dosage of mescaline that Huxley describes taking and in the level of alienation Huxley was willing to risk. (Martin, 2003, 516) Michaux's mescaline texts - *Misérable miracle* (1956), *L'Infini turbulent* (1957), *Connaissance par les gouffres* (1961), and *Les grandes épreuves de l'esprit* (1966) explicitly deal with the question of the realm beyond reality, probing for a metaphysical vocabulary for extreme experience in a secular age. What distinguished Michaux from Huxley, and from other experimental writing of the period was his commitment to the witness position and his (perhaps naïve) desire to document the experience without cultural or personal biases. Of *Misérable miracle,* Martin writes,

"Ce sera le seul livre de tremblement qui, par cet aspect documentaire, donnera à ce point l'impression d'une experience saisie sur le vif" (Martin 2003, 522). [This was to be the only trembling book, that by way of this documentary aspect would give the impression of an "on the spot" or unposed experience.]

Michaux wanted to be both the observed and the observer, which plays an essential role in his writings on these self-experiments.

"Je est un autre": Reverberations
One of the central products of Michaux's mescaline experiments is the experience of the self as other, recalling Rimbaud's famous *"Je est un autre"* formulation, which the poet used in a 1871 to express the strangeness of the self, and the need for the derangement of the senses to access the unknown. In the last mescaline text - *Les Grandes épreuves de l'esprit et les innombrables petites* [*Major Ordeals of the Mind and the*

Countless Minor Ones] – Michaux is disturbed by the experience of hearing another's voice as his own. He writes, "Une autre voix que la mienne se pousse dans ma voix, en traître. Je m'arrête. Étrange. Freinant. Décourageant. Avec un mot, qui n'était pas de moi, une voix étrangère aussi, a voulu intervenir " (Michaux 1966, 87). *[Another voice than mine makes its way into my voice, a traitor. I stop. Strange. Dismaying. Discouraging. With a word, which was not mine, a voice, also alien, has tried to intervene (Michaux 1974 68).]* While Michaux retains his position as the witness figure (who does not lose himself completely) in this example, he experiences the self as another's voice, pushing through his throat.

A decade later, this description becomes a marker for unselfing in a very different context. André Malraux, a contemporary to Michaux, describes an eerily similar experience of another's voice pushing through his throat in his description of his near-death experience during severe illness in a hospital in 1940. Malraux writes of what he calls the *je-sans-moi* [I-without-the-Ego] experience in a volume called *Lazare* (1974) [Lazarus]. During his illness, Malraux is reminded of the Russian front where he fought in 1916, and writes that the *je-sans-moi* leads to a radical feeling of communion amongst soldiers during war. He writes,

"Les hommes entendent toujours leur voix avec la gorge, et celle des autres avec les oreilles. Si nous entendions soudain une autre voix que la nôtre avec la gorge, nous serions terrifiés. J'avais écrit que tout homme entend *sa vie* avec la gorge, celle des autres, avec les oreilles, sauf dans la fraternité ou l'amour" (Malraux 1974, 238). *[Men always hear their own voice with their throat, and the voice of others with their ears. If we were to suddenly hear a voice other than our own through the throat, we would be terrified. I have written that every man hears his life with the throat, and the life of others with the ears, except in fraternity or in love.]*

Malraux affirms the value of unselfing in that it creates a radical feeling of communion amongst sufferers.

This description of the *je-sans-moi* experience would later lead Jean-François Lyotard to theorize what he calls "stridence" [stridulation], or the sound-feeling of one's own voice heard through the throat (as opposed to the voice of others, which one hears through the ears.) Lyotard writes that the abyss between "moi et je-sans-moi" [the ego and the I-without-the-ego] is unsurpassable by outer others – except during experiences of love (as Malraux already asserted) or dread. While in Malraux, and later in Lyotard, unselfing involves an important inter-subjective dimension, Michaux conclusions seem limited to the self. Indeed, at one point, Michaux goes as far to say that his experiments are "peu utiles" [of little use] for relations with others, because they are supposedly incommunicable (Michaux, 1966, 63). And yet, Michaux insists that the derangement of the senses is not about

escapism, and as we have seen, thinks proper self-experimentation demands a commitment to mental control throughout the process. In a long and tortured footnote to *Les Grandes épreuves* [*Major Ordeals*] Michaux writes,

> Que ceux qui prennent des produits pour s'adonner aux excitations collectives, trépignements, danses hystériques, bagarres ou viols, s'arrêtent et ne se mettent pas à croire qu'il y a quelque chose ici pour eux. On ne parle pas la même langue. On ne va pas aux mêmes effets. Celui qui est incapable de retenir les actes, incapable de garder tout *dans le mental* est complètement à côté. Le psychique contemplateur est 'retranché.'" (191)[*Those who take drugs in order to surrender themselves to the collective release and emotional abandon need not read further. There is nothing here that is meant for them. We do not speak in the same language. We do not look for the same effects. He who is incapable of keeping his actions under control, incapable of confining everything to the mind, has missed the point completely. The observer of psychic occurrences has to be "entrenched." (156)*]

Michaux's emphasis on "entrenchment" raises again the question of the communicability of drug experiments. Throughout the mescaline texts, Michaux is constantly commenting in the margins, adding footnotes and directly addressing the reader in the text as the experimental subject, and in the margins as the experimenter, never fully engaging with his passing observation on the uselessness of the experiments for relations with others. The presence of others in Michaux's mescaline texts is synecdochical: they appear in the form of hands, echoes, shadows, complicating his attempts to justify his drug trips with the claim that they are not motivated by escape of reality. The experience of the self-as-other, of another's voice pushing through his own throat, does not necessarily push Michaux towards connections with the outer world, as it does for Malraux. Although he conquers new inner territories, outer ones are but shadows in the mescaline texts. Michaux remains entrenched.

Endnotes

1 Shelley Weinberg (2011) writes, "Locke also seems to see his theory of personal identity as complementing his moral and theological views. 'Person', Locke says, 'is a Forensick Term appropriating Actions and their Merit', for which 'at the Great Day' we are judged by God and receive eternal reward or punishment (II.xxvii.26, I.iv.5). For further discussion of the "metaphysical criterion" in Locke's *Essay*, see "Locke on Personal Identity" Philosophy Compass, 398–407. For more on twentieth-century defenses and criticisms of Locke's theory on personal identity, see: Perry (2008)

2 Unless otherwise noted, translations are my own.

3 I am indebted to Joshua Landy (2012) for this neutral term that covers both self-transcendence and self-rupture, which he uses in his recent book, *How to Do Things with Fictions.* Oxford: Oxford University Press.

4 In my PhD thesis, I explore other narratives for "unselfing", such as destruction, multiplication and transformation.

5 See: Camus 1965, 101

6 Various Michaux critics have attributed to the writer an existential project, and on the whole have "tended to pay more attention to the existentialist undertones in his writings until the mid 1960's" (Rigaud-Drayton 2005, 6) at which point a more psychoanalytic perspective on his work comes to the fore.

7 For more on existentialism and the self, see: Guignon, C. (2004). *The Existentialists: Critical Essays on Kierkegaard, Nietzsche, Heidegger, and Sartre.* Lanham: Rowman & Littlefield Publishers

8 This would prove important for Michaux in his attempts to validate his native Belgian tradition at the same time that he sought to distance himself from it. See Rigaud-Drayton (2005, 30)

References

Broome, P. (1977) *Henri Michaux*. London: Athlone Press

Camus, A. (1965). *Essais.* Paris: Gallimard

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon Books

Gallagher, S., Zahavi, D., (2008). *The Phenomenological Mind*. London: Routledge

Locke, J. & Nidditch, P. (1975). *An Essay Concerning Human Understanding*. Oxford: Clarendon Press

Lyotard, J.-F. (2001). *Chambre sourde: Soundproof Room: Malraux's Anti-Aesthetics.* Stanford, CA: Stanford University Press

Malraux, A. (1974). *Lazare*. Paris: Gallimard

Martin, J. (2003). *Henri Michaux*. Paris: Gallimard

Michaux, H. (1966). *Les Grandes épreuves de l'esprit et les innombrables petites*. Paris: Gallimard

___ (1974). *Major Ordeals of the Mind and the Countless Minor Ones.* New York: Harcourt Brace Jovanovich

___ (1972). *Misérable miracle, la Mescaline.* Paris: Gallimard

Perry, J. (2008). *Personal identity*. 2nd ed. Berkeley: University of California Press

Rabinovitch, C. (2002). *Surrealism and the Sacred: Power, Eros, and the Occult in Modern Art*. Boulder, CO: Westview Press

Rigaud-Drayton, M. (2005). *Henri Michaux: Poetry, Painting and the Universal Sign.* Oxford: Claredon Press

Schechtman, M. (1996). *The Constitution of Selves*. Ithaca, NY: Cornell University Press

Mindfulness and Self-Experience:
No-Self Consciousness or Empty-Self Consciousness?

∎

Vera Pereira
Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

∎

Abstract: *Mindfulness training has been used for centuries in eastern traditions such as Buddhism and Yoga, with the goal of reaching altered states of consciousness capable of changing the experience and perception of self, for spiritual purposes. In the west, mindfulness-based psychotherapy programs have become popular in recent years and their efficacy is well established, being used as a treatment for depression and anxiety. In accordance with the original descriptions of the goals and effects of these techniques, neurophysiologic studies have shown that meditation affects self-referential processing and that from the early stages of practice there is significant neuroplasticity in self-referential and attentional networks in the brain. In this article, the phenomenological and neurophysiologic effects of mindfulness practice are explored, in an attempt to understand, from a scientific point of view, how altered self-consciousness experiences associated traditionally with meditation techniques are generated, and contribute, to induce drastic changes in the implicit and explicit conceptualization of the self of the practitioner. Taking in consideration available data about the effects of mindfulness training, I propose that continued and regular practice of meditation techniques can lead to a dissociation of narrative and momentary self-reference modes, and to a progressive de-identification with the narrative-self and an identification with the minimal-self, that is experienced and perceived as an "empty", non-personal psychological entity.*

Keywords: Mindfulness, meditation, altered states of consciousness, self, self-awareness, self-referential processing

Mindfulness and self-experience: no-self consciousness
or empty-self consciousness?

Mindfulness is an English word used to translate various Buddhist terms, mostly *Sati* (Pali), *Smrti* (Sanskrit) and *dran pa* (Tibetan). These terms refer originally to

"memory", "remembering" or "keeping in mind", and in the Buddhist tradition they are understood as referring to the quality of being aware and paying attention (Dreyfus 2011; Gethin 2011; Grossman & Van Dam 2011; Teasdale & Chaskalson 2011a). Mindfulness practice is dedicated to develop the ability of the mind holding "on to its object and allow for sustained attention regardless of whether the object of attention is present or not" (Dreyfus 2011, 42-43). In the Buddhist tradition, mindfulness is described as "involving attention and awareness of bodily sensations, and also awareness of more expansive mental contents and processes, such as emotions and an altered view of self" (Nanamoli & Bodhi 2001, cited by Grossman & Van Dam 2011).

Meditation techniques are central in mindfulness training, and can be found not only in Buddhism, but also in other traditions, such as Taoism, Yoga and Vedanta (Shear 2011). Meditation includes several types of techniques that are usually practiced in a specific context and that require a specific body posture, usually a seated posture in which the spine must be kept straight, although there are some forms of mindfulness meditation that don't require necessarily sitting positions, such as Yoga or walking mindfulness techniques (Deshimaru 2003-2005; Kabat-Zynn 1982; 1992; Lutz et al 2008).

Meditation techniques are central in mindfulness training, and can be found not only in Buddhism, but also in other traditions, such as Taoism, Yoga and Vedanta (Shear 2011). Meditation includes several types of techniques that are usually practiced in a specific context and that require a specific body posture, usually a seated posture in which the spine must be kept straight, although there are some forms of mindfulness meditation that don't require necessarily sitting positions, such as Yoga or walking mindfulness techniques (Deshimaru 2003-2005; Kabat-Zynn 1982, 1992; Lutz et al 2008).

Meditation techniques: brief presentation
Eastern meditation techniques can be classified in two main categories (Lutz et al 2008): a) focused attention techniques and b) open monitoring techniques.

Focused attention techniques These techniques involve directing and sustaining attention moment by moment on a chosen object, which can be one's own breath and bodily sensations accompanying respiration, a mantra or picture.

While trying to keep attention focused on the selected object, it is normal to notice that the focus has shifted from the object to another stimulus (internal or external, such as a thought about what we will do for dinner later, the work we have to finish tomorrow, or a sensation of pain on our backs). The meditator should notice the distraction and, after noticing it, he should refocus his attention on the selected object. With practice, the time that is necessary either to notice the distraction or

return to focus on the selected object gets smaller, until the sustained attention becomes an "effortless concentration".

Neurophysiologic studies show that *focused attention* meditation techniques can train attentional skills and induce long-term neurophysiologic changes in subcomponents of attention. Consistently with the phenomenological reports of expert meditators, studies have shown that activation in brain regions implicated in monitoring, engaging attention and attentional orienting is stronger in novice or recent expert mediators (19 000 hours of practice), than on very well trained experts (44 000 hour of practice) – with practice the exercise really seems to become effortless (Brefczynski-Levis et al 2007, cited by Lutz et al 2008).

Open monitoring techniques. Usually this type of meditation starts with *focused attention* mediation for calming the mind and foster concentration. The monitoring component of focused attention is used to progress for *open monitoring*, in which the goal is to remain only in the monitoring state, being attentive to whatever stimulus that might be detected, without a particular focus on a specific object. The practitioner should be in a nonreactive monitoring of whatever appears in attention. The instruction for novices is usually to label the experiences at each moment and then return to *open monitoring*. For example, if thoughts about the work one has to finish tomorrow are detected, they should be labeled as "worry", and so on. The continuity of this process leads to a nonreactive awareness of automatic cognitive and emotional interpretations of sensory, perceptual and endogenous stimuli.

The aim of this type of meditative practice is to "gain a clear reflexive awareness of the usually implicit features of one's mental life" (Lutz et al 2008, 164). This awareness is supposed to help change cognitive and emotional habits, and lead to an awareness of self less emotionally reactive and more focused on the present and less on the autobiographical past, present and imagined future.

Both styles presented are used in combined form. *Focused attention* can be seen as preparation for *open monitoring*. These two techniques are also used in western secular mindfulness-based psychotherapeutic interventions very popular today (Hofmann et al 2010), such as *Mindfulness-based stress reduction* (MBSR; Kabat-Zinn 1982) and *Mindfulness-based cognitive therapy* (MBCT; Segal et al 2002).

Travis & Shear (2010) proposed a third type of meditation techniques, named *automatic self-transcending*. These are supposed to involve the transcending of the procedures of the meditation, involving minimal cognitive control or effort. Practices such as *transcendental meditation*[1] are presented by the authors as an example of these techniques. Transcendental meditation (TM) involves the repetition of a mantra (word or series of syllables) and is sometimes considered as a *focused attention* form of meditation. For Travis & Shear (2010) this is not the case,

since TM is associated with different electroencephalography (EEG) frequency bands than focused attention and also because beginning practitioners of TM quickly develop a certain level of automaticity "in which the mantra becomes increasingly secondary in experience and ultimately disappears and self-awareness becomes more primary"(Travis & Shear 2010, 1116). What seems to be happening here is that the meditation technique used, which is *focused attention*, is transcended during practice, and this allows the meditator to have a different type of experience which induces more profound changes in self-awareness, associated with "self-transcendence" or "pure consciousness" states (Travis &Pearson 2000).

In any case, it should be noted that traditional *focused attention* techniques also become effortless with practice (Lutz et al 2008). The same seems to happen with *open monitoring* techniques such as *zazen*[2], the Japanese Zen Buddhist seated meditation, which can induce electrophysiological changes consistent with *automatic self-transcending* in experienced practitioners: high amplitude and low frequency alpha waves or low alpha waves (Deshimaru 2003-2005, 107).

Travis and Shear (2010) recognize that most meditative techniques, with practice, become effortless and automatic and can be used to achieve the same consciousness states as what they call "automatic self-transcending" techniques (1116). The specificity of *automatic self-transcending* techniques is their capacity to induce this automaticity at the outset and not only after extended practice as the other techniques.

"Pure consciousness" or "emptiness" states

With continued practice, most meditation techniques seem to be able to induce an altered state of consciousness that is conceptualized differently by Hindu (such as Yoga and Vedanta) and Buddhist traditions[3]. While Hindu traditions usually label this altered state of consciousness as "pure consciousness", "consciousness itself", "being" or "self" (ātman), the Buddhist traditions usually use expressions such as "emptiness", "non-being" or "no-self" (anātman) to describe it (Shear 2011; Travis & Pearson 2000).

This "pure consciousness" or "emptiness" experience is described as "free from the processes and contents of knowing", and its content is supposed to be "self-awareness itself" (Travis & Pearson 2000, 79). The "pure or empty ground of consciousness itself", is experienced when all contents of awareness are removed (Shear 2011). It is different from regular waking conscious experience in the sense that usually the contents processed in our minds are inner or outer objects, but not consciousness itself (Travis & Pearson 2000).

This state of consciousness is also presented as the natural or original state of consciousness. Reaching "emptiness" or "pure consciousness" states is compared to returning to the natural state of the mind:

> *Zazen* is not a special condition of the spirit. *Zazen* is not a teaching method of special mental conditions; it is not a method for creating a special personality: it is the return to the natural and tranquility. It is because of this reason that *zazen* can put an end to the disturbances of our complicated and fragile society and bring it back to the authentic conditions of the spirit[4]. (Deshimaru 2003-2005, 145; Zen monk of the *Soto* school).

> The time comes when no reflection appears at all. One comes to notice nothing, feel nothing, hear nothing, see nothing (…) But it is not a vacant emptiness. Rather it is the purest condition of our existence. (K.S, a Zen practitioner quoted in: Austin 1998, 463, cited by Shear 2011).

While Buddhist traditions usually conceptualize this experience as a state of consciousness without self, and consider that this experience leads to the insight that the phenomenological basic self-experience is not real and that there is "no-self", Hindu traditions usually see it as the experience of the true self, or *ātman*, the "witness-consciousness" that is always present in the background (in the case of Vedanta traditions, Hopkins 1992) or the "minimal agentive self" (in the case of the Nyāya tradition, Ram-Prasad 2011). In the particular case of Zen Buddhism, both types of interpretation can be found (Shear 2011).

Travis & Pearson (2000) asked meditators specialized in TM to describe their experiences during their practice and found that "pure consciousness" is described as "self-awareness isolated from the processes and objects of experience, and without a sense of time, space or body" (82).The authors wanted to verify the universality of this experience, looking for possible descriptions across cultures and time. They found that it is described since pre-history. They cite written descriptions of this experience in Egyptian hermetic writings, in the words of a Benedictine monk, a Japanese Zen Buddhist monk and even an English poet (William Wordsworth), considering that there is enough evidence to consider that "pure consciousness" is a naturally occurring experience in our species that is not dependent upon a specific cultural context.

This altered state of consciousness is, apparently, available to every human being after enough practice of any mindfulness meditation technique. It is a state in which self-experience is deeply altered, so much altered that some interpret it as a state of "no-self consciousness". It is also a state that is understood as a return to a more basic or original form of consciousness. Both Hindu and Buddhist traditions agree that this altered state reveals experientially to its owner, the real nature of

consciousness and self. There is a self that is seen as illusory in each case: while Hindu traditions consider that in "pure consciousness" states the true self is revealed, Buddhist traditions accentuate the fact that, in these states, the illusory sense of self is absent.

In an attempt to understand, from a scientific point of view, what these special types of altered self-awareness might be, and how they can be responsible for transforming the way the self is conceived, I will propose that we take into consideration dual-process theories of self-awareness and the neurophysiologic effects of meditation practice.

Dual-process theories of self-awareness
In western psychology and philosophy, self-awareness is frequently seen as a dual process that can occur in two basic modes. These two modes have been described and studied by several authors. The first mode has received several designations, such as:

I (James 1890)*; Subjective self-awareness, Subject-Self*, or *awareness of the self-as-subject* (Duval & Wicklund 1972; Sá-Nogueira Saraiva 2003; Damásio 2010, respectively)**;** *Pre-reflective self-awareness* (Zahavi 1999; Legrand 2007)**;** *Minimal self-awareness* (Gallagher 2000; Gallagher and Frith 2003; Fuchs 2010); and, *Core or Mental self-awareness* (Damásio 1999).

"Subjective pre-reflective" or "minimal self-awareness" corresponds to the most basic form of self-awareness that is present at birth, and allows human beings to distinguish their bodies and internal states from the world and other beings outside (Lewis 1991; Rochat 2003, 2011). Processing incoming information from the environment and responding to that information adaptively depends on some awareness of the self as an agent and receptor of information (Sá-Nogueira Saraiva 2010). In the beginning, this minimal self-awareness is mostly dependent on sensorimotor cues. States of minimal or subjective self-awareness take place when attention is directed away from the self and the person "experiences himself as the source of perception and action" (Duval & Wicklund 1972, 3).

This minimal form of self-awareness is somehow tacit, since it does not involve a reflective awareness of self, only an implicit one. "Any sensation, any perception or action directed towards an object implies a tacit self-awareness; it is given immediately, non-inferentially as mine" (Fuchs 2010, 549). This is the self-awareness associated with the pre-reflective sensation of being a subject or agent who experiences, processes and represents the external and internal world. It is focused on the present moment: the action being performed, the feelings felt or thoughts being processed.

The second mode of self-awareness has been named as: *Me* (James 1890); *Objective self-awareness*, *Object-self* or *awareness of the self-as-object* (Duval & Wicklund 1972; Sá-Nogueira Saraiva 2003, 2010; Damásio 2010); *Reflective self-awareness* (Zahavi 1999); *Narrative self-awareness* (Gallagher 2000; Fuchs 2010) or *autobiographical self-awareness* (Damásio 1999).

In "narrative" or "reflective self-awareness" states our attention is focused in our own selves and we can become aware of ourselves as an object in the world, creating a representation of our self that includes physical and psychological characteristics (Fuchs 2010). Because of this it is sometimes referred to as a *reflective* form of self-awareness. In these states of self-awareness we can focus our attention in our internal and private mental world and bodily sensations or on our more public characteristics that define as a social object: our body image, our behavior in certain situations, what others might think of us, our social and economic status, among other characteristics. The development of this type of self-awareness is dependent upon the acquisition of autobiographical memory, concepts and language (Rochat 2003; 2011; Fuchs 2010).

A representation of the *Object-self (*Sá-Nogueira Saraiva 2003, 2010) or *narrative-self* (Fuchs 2010) results from these states of self-awareness. It corresponds to an image of the self that takes into account how we are seen in other people's eyes, and the autobiographical history of the self – its past, present and future. Some authors see it also as a *symbolic-self* (Sedikides & Skowronski 1997) since it is constructed and represented in different ways depending on the value that each cultural context attributes to different characteristic and traits (Chentsova-Dutton & Tsai 2010; Markus & Kitayama 1991).

These two modes of self-awareness appear to be distinct psychological and neurophysiologic mechanisms or systems, since they are dependent upon different cerebral areas and have a different ontogenetic path: "pre-reflective minimal self-awareness" seems to depend on limbic system activity and is present since birth, while "narrative self-awareness" depends on frontal and parietal lobe activity (LeDoux 1989, 2002; Lou et al 2004; Pribram 1984; Stuss 1991; Weiskrantz 1986), appearing only around the 2nd year of age (Lewis 1991; Rochat 2003, 2011).

Self-referent processing in the brain and dual-process theories of self-awareness
Neuroimaging research has shown that there are two types of self-referent processing in the brain that are distinct but usually occur in an integrated form: *momentary* or *experiential self-reference* and *extended* or *narrative self-reference[5]* (Farb et al 2007; Mills 2011). These two processes appear to correspond to the two modes of self-awareness described in psychological and philosophical literatures. *Momentary self-reference* refers to one's momentary experience or psychological present, and

involves a right lateralized network involved in somatic and visceral sensation, engaging older neural regions. Moment-to-moment self-awareness arises from the integration of basic interoceptive and exteroceptive bodily sensation processes (Farb et al 2007; Mills 2011). This momentary mode of self-reference appears to correspond to the notion of a "pre-reflective" or "minimal" form or self-awareness.

Extended self-reference is involved with memory for self and other traits as well as the continuity of identity over time. It is associated with activation of the medial pre-frontal cortex (mPFC). It corresponds to the notion of a "narrative", "autobiographical self-awareness" (Farb et al 2007). It is know that, in the absence of the demand to attend to external stimuli, there is an automatic tendency to engage in *narrative self-referential processing*, which generates a constantly running stream of thoughts full of self-related material (mental images, beliefs, feelings, thoughts and desires) (McKiernan et al 2006, cited by Farb et al 2011, and Mills 2011). The cortical midline activity associated with *narrative self-awareness* is very similar to activity associated with the "default mode" of resting attention (Gusnard et al 2001; Raichle et al 2001, cited by Farb et al 2011). Since the *narrative self-awareness* system is phylogenetically and ontogenetically more recent that the *minimal self-awareness* system, the narrative mode has probably become the default mode through practice (Farb et al 2007).

When the attention is completely focused in outside stimuli or action being performed, such as in moments of new learning, intimacy or creative processing, the *narrative mode* is shut down and we engage in *experiential self-reference* mode, which keeps attention focused on the present (Mills 2011).

These two self-awareness systems are related, since *experiential/ momentary self-referential processing* activates *narrative/extended self-referential processing*: this is the reason why it is so easy to become distracted from the present moment and engage in self-related thoughts; attention to the present moment engages narrative self-talk (Farb et al 2007).

Meditation and self-referential processing
Neurophysiological and neuroimaging studies about the effects of meditation on brain functioning present some discrepancies concerning the areas activated and deactivated during mediation. These discrepancies can be explained by the diversity of techniques that were studied (Ives-Deliperi et al 2011; Lutz, Dunne, & Davidson 2007; Travis & Shear 2010), which might affect different brain areas. This means that more studies are needed before we can ascertain for sure what are the neurophysiologic and neuroimaging patterns related with each type of mediation techniques.

Despite the variety of findings, one pattern of results seems to be particularly relevant for clarifying what might be the psychological and neurophysiologic correlates of "pure consciousness" or "emptiness" states. These results are concerned with the effect of mindfulness meditation in *self-referent processing* (Berkovich-Ohana, Glicksohn & Goldstein, 2011; Farb et al 2007; Goldin et al 2009; Ives-Deliperi et al 2011; Mills 2011). Some studies appear to show that meditation practice helps to dissociate the *experiential* and *narrative* forms of self-reference, which are usually integrated. Mindfulness training such as the "Mindfulness-based stress reduction" program by Kabat- Zinn et al (1992) and "Music based mindfulness training", by Shinzen Young, increase the capacity to disconnect the experiential from the narrative mode, and engage in present-centered awareness of self, that is not affected by narrative intrusions, such as thoughts, feelings and desires (Farb et al 2007; Ives-Deliperi et al 2011; Goldin et al 2009; Mills 2011).

From a neurophysiologic point of view, this is associated with deactivations of the *midline pre-frontal cortex* (MPFC) region and an activation of *right lateralized*, *prefrontal cortical* and *paralimbic structures* (with the recruitment of *posterior viscerosomatic* representations). The novice practitioners also show different activation patterns, but they are not capable of de-activating the ventral and dorsal MPFC such as the trained practitioners (Farb et al 2007).

With practice, the increased ability to shut down the *narrative self-focus* and disconnect it from the *experiential self-focus*, could lead to the "pure consciousness" or "emptiness" states in which, at least for a few moments, there is just an experiential, momentary self-reference without narrative intrusions. The lack of narrative intrusions could be experienced as the sensation of "thinking about nothing" (K.S, a Zen practitioner, quoted in Austin, 1998, 463, cited by Shear, 2011), as if consciousness was "free from the processes and contents of knowing" (Travis & Pearson 2000, 79).

The capacity to progressively dissociate *narrative* and *minimal* self-processing appears also to introduce the possibility to start thinking about the self as something dissociated from specific mental contents such as thoughts, emotions, mental images, and representations, and as something not identical with the *autobiographical* or *narrative-self* (Ives-Deliperi 2011). This process is described by patients in "Mindfulness-based cognitive therapy" programs, which share with their therapists insights such as: "these thoughts and feelings aren't me" (Allen et al 2009, cited by Teasdale & Chaskalson 2011b).

The insights of patients in mindfulness-based therapeutic programs are in line with Buddhist and Hindu traditions: while Buddhist thinkers conclude that the self is not real because narrative self-talk is not identical or necessary for conscious

experience, Hindu thinkers appear to conclude that the experiential sense of self disconnected from time and specific contents of the mind is the "real self", and that the *narrative-self* is an illusion.

Continued meditative practice might have the following chain of effects: a) dissociation of *narrative* and *experiential self-referent processing*; b) identification of self with *minimal pre-reflective* or *experiential sense of self* and c) de-identification with the *narrative sense of self*. In this process, the *minimal pre-reflective self* appears to be experienced as an "empty-self", in the sense that it corresponds to an experience of the self as a source or point of awareness from which mental contents and external reality is observed, perceived and processed, that is different and independent from those same contents and which has no personal identity. This "empty-self" is identical to the Kantian notion of a "transcendental self-consciousness" (Keller 1998, 2-3, cited by Ram-Prasad 2011).

There are descriptions of this "empty-self" in both traditions. From the Hindu tradition: Ātman is "(…) the witness only, that which experiences, but its own nature lies beyond all suffering, pleasure or confusion" (Taylor 2000, 18); Ātman is " (…) the conscious being that underlies each person and remains unchanging in the midst of activity and change" (Hopkins 1992, 149). From the Japanese Buddhist *Soto Zen* tradition:

> What is the Self? My master Kodo Sawaki used to say: "We should know our pure and true self. We should become intimate with ourselves" (…) The real self can be named as the "pure subject" (…) Thanks to *zazen*, seated in tranquility, we can find our true nature, our superior pure self, avoiding to look for it consciously; what I mean is that we find it without "thinking" about it, stopping all mental activity directed at that search. (…) During daily activity we are not very sensitive to all of the varied impressions that appear and disappear in our spirit. (..) During *zazen,* we can objectively observe our mind. We perceive its movements, we observe them in freedom, we are intimate with them and have access to our own self, free and open (…). When we practice *zazen*, everything that is a part of the world around us, behind time and space, exists and does not exist objectively. Everything becomes one in our mind. That's the result of activity of our objective superior spirit[6]. (Deshimaru 2003-2005, 149-150, Zen monk of the *Soto* school).

"True self" versus "illusory self" appears to be more relevant for both traditions than "self" versus "no-self" distinction. The same was already noticed by Albahari (2011), which proposed that "the sense of "Who I am" versus "What I am not" is the true target of early Buddhist practices that seek to eliminate the sense of a self. What remains after the sense of self has dissolved is a unified perspectival "witness-

consciousness", that insofar as it lacks the illusion of a personal self, is intrinsically ownerless" (82).

In an attempt of conciliation between Buddhist and Hindu perspectives, I propose that we refer to "pure consciousness" or "emptiness" states as *empty-self consciousness* states. "No-self consciousness" is not descriptive of what is happening from a phenomenological, psychological and neurological point of view, since self-referential processing is still occurring and the first-person perspective is still present. "It is a state of "consciousness" in that the knower is conscious through the experience, and can, afterwards, describe it" (Travis & Pearson 2000). The self, as perceived during *empty-self consciousness* states is like an empty canvas, since it is not identical with mind contents and is not identical with the *autobiographical self*, but it isn't gone.

Empty-self consciousness states: between experience and reflection?
Although "pure consciousness" or "emptiness" states, from here on designated as *empty-self consciousness* states, are usually associated with non-reflection in the Buddhist tradition, which accentuates that nirvana (realization of no-self) is reached experientially (Albahari 2011), there are some reports of this altered state of consciousness that say that its content is "self-awareness itself" (Travis & Pearson 2000). If this is so, than *empty-self consciousness* might be or become at some point a meta-cognitive state in which there is the possibility to attend to the process of present-moment *minimal self-awareness*, which is not common during typical minimal self-awareness states. Such as Sá-Nogueira Saraiva (2010) noted in an evolutionary account of the *minimal* or *core-self*, which he calls *Internal Reference Point* (IRP): "the IRP's function is not to muse upon itself: it is a switchboard between the inside and outside [of the organism] and not a thing to be thought about" (955).

The insight that there is a "real self" and an "illusory self" could be a part of the experience itself, instead of being just the result of post-experience reflection. This would also mean that there is a narrative or reflective side to this experience, probably occurring after the brief periods of pure non-narrative *experiential self-processing*. Since pure *experiential self-reference* states not contaminated by *narrative self-reference* are difficult to maintain (Farb et al 2007), it is very probable that the experience is immediately translated into a narrative about the nature of the self being experienced during those states. This probably occurs also due to the meditators knowledge of Buddhist, Vedanta, Yoga or at least, New Age literature about meditation, which sets expectations and helps to frame the experience in an already know framework.

In any case, the realization that there is "no-self" or that there are two forms of self-consciousness and that one is more basic, vital or truthful than the other, does not correspond only to an experiential, momentary, and non-reflective state, it requires insight and reflection. More research about the phenomenology and neurophysiology of these altered states of consciousness might help to clarify the psychological process taking place and how much experiential and non-reflective is the experience of *empty-self consciousness* states.

As it happens with other religious and mystical experiences, the research of the phenomenology or psychology of mindfulness states faces the problem of distinguishing the experience itself from the meaning attributed to it, since mindfulness practice takes place in a particular cultural context, and even the novice practitioners are exposed to expectations and predictions of what will and should be happening during the experience. The ideal research situation to clearly identify and characterize this *empty-self consciousness* states would be to have naïve practitioners, exposed only to the techniques without any theoretical explanation about the intended psychological results.

Final remarks

The scientific literature available on the neurophysiological and neuroimaging correlates of meditation suggests that mindfulness techniques appear to be a specific kind of cognitive technology, which allows humans to enhance self-experience and learn with practice to engage in a consciousness state in which the mind is not dominated by *narrative* or *extended self-referential processing*. I have suggested that this dissociation from narrative self-focus during meditation can be responsible, or at least involved, in the altered self-experience designated as "pure consciousness" or "emptiness" in ancient eastern traditions. Additionally I proposed that experiencing these altered states of self-consciousness can change the way the self is conceptualized due to a progressive identification of self with the *minimal pre-reflective sense of self*, and de-identification with the *narrative* or *autobiographical sense of self*.

These assumptions should be seen as a modest attempt to link available data and theory in a coherent but provisory framework, since more research is necessary to understand clearly what types of meditation are linked with changes in self-referent processing and if these changes are associated with reports of *empty-self consciousness* states from a phenomenological point of view. The dissociation of *momentarily* and *extended self-referent processing* might be only the first step to enter in this altered state of consciousness, but not a sufficient condition to enable it. More neurophysiological changes might be responsible for, or at least correlated with *empty-self consciousness* states. Studies comparing self-referential processing and phenomenological reports from practitioners of different ancient traditions and

different mindfulness therapeutic programs, trained in several types of meditation techniques will be necessary to achieve a comprehensive vision of this altered self-experience. These studies should also aim to clarify the extent in which this experience is non-reflective and subjectively felt as non-corporeal and not time bounded.

Additionally to the lack of available data linking directly *empty-self consciousness* and self-referential processing modifications during meditation, another limitation can be addressed to the main ideas presented in this article. That limitation has to do with the fact that we have considered, in line with Shear (2011) and Travis & Pearson (2000), that "pure consciousness" and "emptiness" states are only one type of consciousness states, differently interpreted by two different traditions, when they could, in fact, be two different types of altered self-consciousness. The hypothesis that they are two interpretations of the same experience makes sense in light of the huge debate around the topic of "self" and "no-self" in the two traditions, which could probably have been settled easily, if there was a direct association of each interpretation with one particular kind of consciousness state. The present author has also considered the discrepancies found inside Buddhism itself, which has schools such as *Japanese Soto Zen* where we can find interpretations compatible with the idea of "pure consciousness" and "self", as an indication that similar mindfulness techniques, which produce identical results, can lead to different interpretations even in the same tradition.

Independently of what future research might unveil about the nature, quality and variety of mindfulness experience, one thing is certain: theories of self-awareness will certainly be affected by new and more precise knowledge of mindfulness meditation practice.

Endnotes

[1] Transcendental meditation is a meditation technique derived from the Vedic tradition of India that involves the repetition of a mantra (word or series of syllables), and that was popularized in the West by Maharishi Mahesh Yogi (1986, cited by Travis & Pearson, 2000).

[2] "Zazen" is the foundation of Zen practice and consists in a sitting meditation that is "based on a trinity: posture, breathing and meditation" (Yamahata 2002, 22). In Japanese, "za" means "to seat" and "zen" means "meditation" (Deshimaru 2003-2005, 54). "Zazen" is the name given to seated meditation in Japanese Zen schools.

[3] It should be noted that Hindu or Brahmanical traditions and Buddhist traditions include a huge variety of schools that have different positions about "self" (ātman) and "no-self" (anātman). Here we will consider a classical opposition between the two traditions about the nature of self. When referring to specific schools of both traditions, they are identified by their name and/or specific authors within them.

[4] Translation from the Portuguese edition by Assírio & Alvim, 2003-2005.

[5] Both expressions "momentary/experiential" and "extended/narrative" will be used interchangeably thorough the remaining article to designate the two types of self-referential processing.

[6] Translation from the Portuguese edition by Assírio & Alvim, 2003-2005.

References

Albahari, M. (2011). "Nirvana and ownerless consciousness." In: M. Siderits et al (eds). *Self, No Self? Perspectives from Analytical, Phenomenological and Indian traditions.* Oxford: Oxford University Press, 79 – 113

Berkovich-Ohana, A. et al (2012). "Mindfulness-induced changes in gamma band activity – Implications for the default mode network, self-reference and attention." *Clinical Neurophysiology, 123(4)*, 700-710

Chentsova-Dutton, Y. & Tsai, J. (2010). "Self-focused attention and emotional reactivity: The role of culture." *Journal of Personality and Social Psychology 98*, 507-519

Damásio, A. (2010). *O Livro da Consciência: A construção do cérebro consciente*. Lisboa: Círculo de Leitores

___ (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt

Deshimaru, T. (2003). *Verdadeiro Zen*. Lisboa: Assírio & Alvim

Dreyfus, G. (2011). "Is mindfulness present-centered and non-judgmental? A discussion of the cognitive dimensions of mindfulness." *Contemporary Buddhism 12(1)*, 41-54

Duval, T., & Wicklund, R. (1972). *A theory of objective self-awareness*. New York: Academic Press

Farb, N., et al (2007). "Attending to the present: mindfulness meditation reveals distinct neural modes of self-reference." *Social Cognitive and Affective Neuroscience 2(4)*, 313-322

Fuchs, T. (2010). Phenomenology and Psychopathology. In: S. Gallagher, D. Schmicking (eds.). *Handbook of phenomenology and the cognitive sciences.* Dordrecht: Springer, 547-573

Gallagher, S. (2000). "Philosophical conceptions of the self: implications for cognitive science." *Trends in Cognitive Sciences 4(1)*, 14-21

Gallagher, H., & Frith, C. (2003). "Functional imaging of "theory of mind". *Trends in Cognitive Sciences 7(2)*, 77-83

Gethin, R. (2011). "On some definitions of mindfulness." *Contemporary Buddhism, 12(1)*, 263-279

Goldin, P., et al (2009). "Mindfulness meditation training and self-referential processing in social anxiety disorders: behavioral and neural effects." *Journal of Cognitive Psychotherapy: An International Quarterly 23(3)*, 242-257

Grossman, P., & Van Dam, N. (2011). Mindfulness, by any other name…: trials and tribulations of sati in western psychology and science. *Contemporary Buddhism*, 12(1), 219-239

Hofmann, S. et al (2010). "The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review." *Journal of Consulting and Clinical Psychology 78(2)*, 169-183

Hopkins, T. (1992). "Hindu views of Death and Afterlife." In H. Obayashi (ed.), *Death and Afterlife: perspectives of world religions*. Westport: Praeger Publishers, 43-156

Ives-Deliperi, V. et al (2011). "The neural substrates of mindfulness: an fMRI investigation." *Social Neuroscience 6(3)*, 231-242

James, W. (1890). *The Principles of Psychology*. New York: Dover Press

Kabat-Zinn, J. (1982). "An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: theoretical considerations and preliminary results." *General Hospital Psychiatry 4(1)*, 33-47

___ (1992). "Mindfulness-based interventions in context: past, present and future." *Clinical Psychology: Science and practice 10(2)*, 144-156

LeDoux J. (1989). "Cognitive-emotional interactions in the brain." *Cognition and Emotion 3*, 267-289

___ (2002) *Synaptic Self: How Our Brains Become Who We Are*. New York: Viking

Legrand, D. (2007). "Pre-reflective self-as-subject from experiential and empirical perspectives. *Consciousness and Cognition 16*, 583-599

Lewis, M. (1991). "Ways of knowing: Objective self-awareness or consciousness. " *Developmental Review 11(3)*, 231–243

Lou, H. et al (2004). "Parietal cortex and representation of the mental Self." *Proceedings of the National Academy of Sciences (U S A) 101(17),* 6827–6832

Lutz, A. et al (2008). "Attention regulation and monitoring in meditation." *Trends in Cognitive Sciences* 12(4), 163-169

Markus, H. & Kitayama, S. (1991). "Culture and the self: Implications for cognition, emotion, and motivation." *Psychological Review 98*, 224-53

Mills, K. (2011). "Experiential self-referential processing & autobiographical memory retrieval: a preliminary look." *Portland State University McNair Scholars Online Journal* 5, 198-221

Pribram, K. (1984). "Emotion: A Neurobehavioral Analysis." In: K.Scherer & P. Ekman (eds.). *Approaches to emotion*. Hillsdale, NJ: Lawrence Erlbaum and Associates, 13-28

Ram-Prasad, C. (2011). "The phenomenal separateness of self: Udayana on body and agency." *Asian Philosophy 21(3)*, 323-340

Rochat, P. (2003). "Five levels of self-awareness as they unfold early in life." *Consciousness and Cognition 12(4)*, 717-731

___ (2011). "The self as phenotype." *Consciousness and Cognition* 20 (1), 109-110

Sá-Nogueira Saraiva, R. (2003). *Mundos Animais, Universos Humanos*. Lisboa: Fundação Calouste Gulbenkian

___ (2010). "Evolutionary psychology of the I and Me." *Revista Portuguesa de Filosofia 66*, 945-962

Sedikides, C., & Skowronski, J. J. (1997). "The symbolic self in evolutionary context." *Personality and Social Psychology Review 1(1)*, 80-102

Segal, Z., et al (2002). *Mindfullness-based cognitive therapy for depression: a new approach to preventing relapse*. NY: Guilford Press

Shear, J. (2011). "State-enlivening and practice-makes-perfect approaches to meditation." *Biofeedback 39(2),* 51-55

Stuss, D. (1991). "Self, awareness, and the frontal lobes: a neuropsychological perspective." In: J. Strauss, G. Goethals (eds). *The self: interdisciplinary approaches*. New York: Springer-Verlag, 255–78

Taylor, R. (2000). *Death and the afterlife: a cultural encyclopedia*. Santa Barbara, Calif.: Oxford: ABC-CLIO

Teasdale, J. & Chaskalson, M. (2011a). "How does mindfulness transform suffering? I: the nature and origins of dukkha." *Contemporary Buddhism 12(1)*, 89-102

____ (2011b). "How does mindfulness transform suffering? II: the transformation of dukkha." *Contemporary Buddhism 12(1)*, 103-124

Travis, F. & Pearson, C. (2000). "Pure consciousness: distinct phenomenological and physiological correlates of "consciousness itself". *International Journal of Neuroscience 100*, 77-89

Travis, F. & Shear, J. (2010). "Focused attention, open monitoring and automatic self-transcending: Categories to organize meditations from Vedic, Buddhist and Chinese traditions." *Consciousness and Cognition 19 (4)*, 1110-1118

Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford University Press

Yamahata, H. (2002). *Folhas caem, um novo rebento/ Falling leaves, a shooting sprout.* Lisboa: Assírio & Alvim

Zahavi, D. (1999). *Self-awareness and Alterity: A Phenomenological Investigation*. (=Studies in Phenomenology & Existential Philosophy). Evanston: Northwestern University Press

II

Altered Selves

…

12

How to define a Mental Disorder?

.

Jorge Gonçalves

Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

.

In this paper I intend to do a personal reflection on mental disturbance realism. The ideal of Psychiatry would be to define mental disturbance with the same degree of objectivity as physical illnesses. These exist regardless of our opinions about them. Genetic anomalies have existed before DNA had been discovered. In the same way, mental disturbance should be defined as existing in itself, regardless of the theories we have about it. However, Psychiatry, being closely related to specific values and social practices, has had difficulty in defining its object study in a "neutral and objective way", according to the scientific ideal, which is to be "a view from nowhere" in the words of Thomas Nagel.

In this paper I claim that mental disturbance is defined from a background of epistemic, social and ethic norms (Broome & Bortolloti 2009). This background is not always explicit. It is not my intention here to investigate the genesis and transformation of this background, only to defend that it is in relation to it that someone is labelled as mentally ill. Other alternative definitions have been put forward and I intend to support that these alternatives are no better than the definition from epistemic, social and ethic norms.

A first objection to my claim consists of the "evolutionist naturalist" model which seems to me to be very influential in modern day Psychiatry (Wakefield 1992). Here mental disturbance is conceptualized in a model analogous to the physical illness model. The nervous system, namely the brain, is constituted of structures that were formed evolutionarily. Each of these structures is related to certain functions (for example: Pio Abreu 2012). The malfunction of one or more of these structures has consequences in the mind and behaviour yielding pathology. This perspective has been dominant in theoretical terms, with research being oriented towards neuronal mechanisms that are supposed to underlie mental disturbances. These, however,

are not diagnosed directly from the analysis of neurological condition of the person. The diagnosis is established by a professional from a symptomatological phenomenology, grounded in standardized norms. These norms in turn have been grounded in social and ethical norms prevailing in a certain social context.

Take the well-known example of homosexuality. Until 1990 it was considered a mental disturbance by the World Health Organization. On 17 May 1990, the General Assembly of the World Health Organization (WHO) removed homosexuality from their list of mental disorders. After that date it was not considered a mental disturbance anymore. Progressively, homosexuality has been accepted as normal and today, in several countries, it is in equal footing with heterosexuality. However, the neuropsychological structure of the homosexual persons is the same. What has changed was the attitude of society towards this sexual orientation. Incidentally, its being disqualified as a mental disturbance was a very important factor in changing people's attitude towards homosexuality.

Could we generalize the case of homosexuality to other "mental disturbances"? Consider the example of schizophrenia for it does not seem to be a context-dependent disturbance: in any society a person who presents the schizophrenic symptoms will be declared mentally ill. I am not sure that such is exactly the case. There are cultures where hearing voices is positively valued (Bracken & Thomas 2010). However, let us admit that in all known cultures the symptoms associated with schizophrenia are recognized as a mental disturbance. Such a fact does not imply that schizophrenia was not defined by social criteria. What we could affirm is that schizophrenia, or the symptoms grouped by this designation, is negatively valued in all known social contexts. What I mean is that to classify an individual as "a schizophrenic" derives from his behavior in his social environment and not just from the direct analysis of the patient's brain.

Only in a later stage does one try to determine the neural correlates of this disorder. We must note that the fact that the mental disorder is defined in terms of epistemic, social and ethical norms does not prevent the research on its inscription in the brain. The point here is that the disorder's definition does not depend on the identification of its neural correlates. Let us suppose we identify in a person's brain a neural correlate normally associated with the presence of a mental disorder, for instance the aforementioned schizophrenia. Could this, by itself, be sufficient to consider this person as a mentally ill person? I agree when Pickard (Pickard, 2010) gives a negative answer for we would have to verify whether indeed the person has symptoms. If a person has a lesion in an area of the brain, frequently related with a certain function, we can affirm that she person potentially has a mental disorder, but she is not really a mentally ill person if she does not present symptoms. Therefore, that which determines the identification of a mental illness is its

symptoms and not the alterations in a brain state. To give an example, we identify a delusion by the subject's speech for it violates epistemic norms. We can identify the neural correlates of that delusion, but how can we know whether there are always the some neural correlates that unleash such delusions in any subject? Could it not be other correlates? Could a person have the some cerebral configuration and not suffer from delusions? The delusion is identified by the subject's speech; the research on its neural correlates comes after. We can think that in a very advanced phase of neuropsychology it would be enough to identify the neural correlates. However, the only proof we will have would be always the presence of utterances suggesting delirium. On the contrary, we can say at most that we are in the presence of a potential illness, but not in presence of an actual illness. The reference is always the speech and behaviours within the person's social context. Thus, even in an ideal situation where the neural correlates associated with a disorder are known, we are not allowed to define a disorder if we do not have a phenomenological description of the symptoms. Note, however, that we are far from this ideal situation

I do not believe it is possible to talk about mental disorders without using a language that does not include symptom descriptions. The description of lesions in a brain ideally separated from a social context seems per se insufficient to determine mental disorders. Another objection that can be made to the definition of mental disorders by epistemic, social and ethical norms is that from a pragmatic point of view what is at issue is the suffering, experienced by the actual person. Accordingly, suffering is not to be confused with labelling. It is not because a person is labelled as mentally ill that she suffers; on contrary, it is because she suffers that she is mentally ill. However, we must note that to suffer is not a disorder in itself. Not all sufferings are classified as disorders or illnesses. The concept of "illness" has its historical origin in the bodily illnesses. Certain lesions could cause suffering and we call them "illnesses". In the case of mental illnesses the situation is more complex for the decision to classify a mental state of suffering, as an illness seems to me to have a strong social and ethical component. Remember that according to Freud we are all slightly neurotic. The determination of the degree of acceptable suffering, to which we do not attribute the status of "disease" is defined by the normative social background.

But could we not say, if we take as reference the suffering that the subject is the one who knows better about her mental state, in that it is she who defines herself, so to speak, as a patient?

In a sense the subject is the one who knows more about herself, but it is evident that it is not she who determines her own diagnosis. The subject complains about something, which makes her suffer, and which is an obstacle in her life. But it is not she that determines whether or she is "ill". As proof, there are subjects that even

when diagnosed as suffering from a mental disorder do not, so to speak, "label themselves" as such. Some of them do not spontaneously consult professionals due to their being led by the influence of others. In this way it seems to me that the diagnosis is determined by professionals of a scientific community, from the exterior of the subject in a certain social context.

Mental disorder and psychological suffering cannot, therefore, be seen as one. To considerer a certain state of psychological distress as a mental disorder implies a certain ethical-social perspective that can change accordingly to the context.

If it is true that mental disorders are defined by social criteria, could Psychiatry preserve its status as an objective science, defining its object in a realist and transcultural way? Would a thesis that states that mental disorders are not objective facts but result from social conventions not be more accurate? We have already seen the above example of homosexuality that, if taken as a paradigm, can be extended to the whole field of mental illness.

George Graham (Graham 2010) holds that the conventionalist thesis is not valid and supports the psychological realism of a mental disorder, as mental, not excluding, however, the presence of neurological factors. On contrary, the causal-mechanist factors will be always present if there is a mental disorder. Graham argues that if it is true that mental disorders depend on social values, this is not a reason for us to reject them as "illnesses" in the realistic sense of the term. As mentioned, the concept "illness" would have initially originated in bodily illness and exported to mental illnesses (Szazs 1960). However, Graham defends that the concept of bodily illness depends on social values as well. Physical illnesses are caused by lesions, but the lesion by itself does not define the illness. What defines the illness is the attribution of a negative value to the consequences of the lesion. What happens is that these values are more universal and less controversial than the ones of physical illnesses and so they do not seem to have a historical-social origin, but they seem to be inscribed in the natural order itself. However, we must remember that adaptation in the case of a human being is to a social and symbolical environment, not to a natural one. Thus, for the author, the definition of a physical illness always implies an evaluative element and not only a descriptive one.

Thus the argument would be that since the concept of "illness" derives from physical illnesses and implies social values, the fact that mental illnesses/disorders also imply social values does not prevent them from being illnesses/disorders. However, the author acknowledges that if the concept of mental disorders is dependent on social relativism, psychiatry will find a realistic definition of mental illness hard to come by (Graham 2010, 108-113). If in a social context we consider homosexuality as a disorder and not in another social context, then the author

believes there can be no objective progress in psychiatry. This is so because we have not defined whether homosexuality is a disorder or only a way of living.

It is therefore necessary to identify the criteria for mental disorder with the highest possible degree of objectivity, even if it is known they will never lose their evaluative component. After eliminating several possible methods, Graham (Graham 2010, 139-142) decides to use a modified form of the thought experiment known as "ignorance veil", initially created by John Rawals in the context of Political Philosophy. Basically it consists of asking ourselves which basic psychological capacities we would like to have in any context unknown to us. The assumption is that we would rationally choose those capacities that would allow us to lead a "life worth living."

In this way he defined seven basic psychological capacities. It would be the impairment of one or more of them that would permit to define a kind of mental disorder (as he exemplifies in the last chapters of the book). These capacities are based on the assumption that the mind is defined by the concepts of "consciousness" and "intentionality". Consciousness is subjectivity or, grammatically, first person experiences. When we ask, "what is it like to be a schizophrenic?" we are asking ourselves about the conscious, subjective states of schizophrenia. Intentionality is the property of mental states to be about something. Beliefs, desires, fears, passions and other mental states are directed, in most cases, to objects, even if they are abstract objects. Graham, in line with John Searle and Donald Davidson, shares the thesis that intentionality is closely related with rationality (Graham 2010, 117). The latter does not mean the capacity to think logically in a narrow sense. Graham views rationality in a broad sense, which could imply emotions and the body. There is rationality, for example, in the way concepts interconnect forming a network. Another example of rationality is given by the interconnection between desires and beliefs. If I want to buy a book, I seek information where I can buy it according to my conveniences (price, distance and so on). Thus the idea is, assuming that the mind is defined by the concepts of consciousness and intentionality, to find the basic capacities, as aforementioned, in order to lead a life worth living. According to Graham (Idem, 147) these are the capacities: "bodily/spatial self-location", "historical/temporal self-location", "general self/world comprehension", "communication, "care, commitment and emotional engagement", "responsibility for self", "recognition of opportunities or "affordances"".

George Graham's idea deserves some criticism. First I am not so sure that we can equalize physical illnesses and mental disorders on the basis that both depend on values. It may be the case that both involve moral values, but the quantitative differences between these two types of illnesses seem to me so great that it is more

accurate to insert them into different categories. At one point in the evolution, consciousness and language appeared in our species. Culture developed with social and ethical norms, which put us in a meditated relation with Nature. The "functions", which also exist in other animals as we can verify, pass, so to speak, through a cultural "sieve" and new features arise. These norms and values also exist in the dysfunctions that appear in physical illnesses; however, here it seems to me that the maladaptation is much closer to natural evolution independent of culture. Graham claims that he does not see a great difference between a person who dies of heart attack and a person who kills himself due to depression. Both would depend on our positive valuation of life. In a society where life was not positively valued (if it is possible to conceive it) neither case would be classified as an illness. However, while in the case of the heart attack there is a direct relation between the organ lesion and its effect, in the case of depression it seems to me that this direct relation does not exist. A simple description of the functioning of the heart and its lesion will be enough to deduce the death of the patient. I think this is not true in the case of depression. The description of neural mechanisms will never be enough to deduce that someone would commit suicide, I think. Thus we can see there are significant differences between the two kinds of "illnesses". In the case of a heart attack the causes are mechanical and relatively controllable. Now, in the case of depression, the role of non-neurological factors (personal relations, individual ideas and so on) in the decision to commit suicide is always subject to speculation.

Thus, although in both cases social values are at issue as in a heart attack and suicide due to depression, in the case of a heart attack we can say with much more priority that there was a function determined by evolution, which was directly interrupted by a lesion. In the case of depression it does not seem to me that we can say the same for there are cultural factors in the determination of the illness/disorder. To consider that depression is a disorder means to appreciate life as something that is extremely positive and joy must be our normal condition. Could we not say the same in the case of the heart attack? Yes, but here there is not an actual evaluative option. It seems impossible to conceive a situation where a heart attack would be something physically positive since the function of pumping blood was selected by the evolution itself. The same could not be said in the case of depression for it could be partially the result of human thought about her own existence.

Nevertheless, despite large differences between physical and mental illnesses, I agree that both depend on values and it will not be for this reason that we cannot use the concept of "mental disorder". However, for this to have the status of objective reality it will be necessary to define it in a non-relativistic mode. This is related to one more problem that I see in Graham's proposal: the fact that he considers to have researched the basic psychological capacities necessary for any

human being to lead a dignified life, using the thought experiment known as "the veil of ignorance", but at the same time he affirms rationality is always linked to a social context (Idem, 196). To give an example, a *sadhu* (Indian ascetic) is positively valued in his culture, but a person who behaved as a sadhu in our society would be probably diagnosed as suffering from a mental disorder. The interpretation Graham would give, seems to me (the example is not his), is that in this case the sadhu follows a rationality that for us is not really rationality, because we do not put ourselves in his point of view. Let us suppose that according to religious Indian beliefs sadhus want to mortify the body and its needs for the purpose of letting other spiritual states flourish. Because for us these states do not exist, behaviour like that of the holy man would be classified as psychotic. But in this case we will fall into relativism.

One way to solve this contradiction which the author does not present (at least explicitly), will be to considerer that human reason progresses through time.

In this case we can keep a certain sociocultural relativism and at same time defend that there are societies, which are "more rational than others". Let us suppose there is a society, which considers homosexuality as a mental disorder and another, which does not. Do they have the same value? Would we say that it depends on context? It seems to me that we can claim it is "more rational" to not view homosexuality as an illness/a disorder given the knowledge we currently have in neurosciences and social sciences. Thus it would be more objective to exclude homosexuality from the labelling of mental illness/disorder. This position contains however a philosophical option, that there is a moral progress in human societies and that this progress is accompanied by scientific and technological progress.

One last criticism is related to Graham's thesis that if we are really dealing with a mental disorder then there will be necessary neurological alterations, which cause impairment of one or more basic capacities but without destroying them completely. If they were completely destroyed we would be in the presence of a physical illness, although the injured organ is the brain. This is the case of illnesses such as Alzheimer or Parkinson. In a mental disorder the basic psychological capacities continue to function but in an altered way thereby causing suffering to the subject. I think we can make a comparison with a broken CD that still lets you hear music, but defectively. However, Graham by stating that the impairment of the basic psychological capacities (and so rationality) is always related to the presence of brute causal-mechanical, neurological, factors, he implies that biology has a decisive weight in defining mental disorders. By itself this does not challenge the realism of mental disorders for we can always conceptualize it in terms of mental personality structures, which have an objective existence. Nevertheless,

these capacities always have a normative dimension, although they are not related to any particular context.

To conclude, the study of mental disorders cannot be limited to neurological research. One cannot avoid the use of mental predicates, meaning the use of the concepts "consciousness" and "intentionality" (as discussed above). Mental disorders are defined by a symptomatic phenomenology within a background of epistemic, social and ethical norms. These norms depend on specific cultural contexts. In this sense, there is a certain relativity in a mental disorder which was sometimes politically explored. Some advocated that the concept of mental disorder would only be a form of power, and mental disorders do not exist in themselves (Bracken & Thomas 2010). However, the presence of social and psychic suffering obliges us to consider the thesis of realism about mental disorders.

In order to make the concept of mental disorders objective it is necessary to release it from contingent social customs and try to define it in a transcultural way, searching to define the conditions of its determination in any social context. This operation does not preclude an evaluative dimension; it is not something neutral and purely empirical. It implies determining the basic capacities that allow leading a life worthy of being lived. However, this project for Psychiatry only seems possible to me if we admit that rationality progresses and spreads throughout the human species. Otherwise, the norms from which we determine the concepts of "normal" and "pathological" will always be linked with particular socio-historical configurations.

References

Pio Abreu, J. (2012). *Elementos de Psicopatologia Evolutiva*. Lisboa: Fundação Calouste Gulbenkian

Bracken, P. & Thomas, P. (2010). "From Szasz to Foucault: On the Role of Critical Psychiatry". *Philosophy, Psychiatry, and Psychology 17 (3)*, 219-228Broome, M. & Bortolotti, L. (2009). "Mental illness as mental: a defence of psychological realism." *Humana.Mente 11*, 25-44

Graham, G. (2010). *The Disordered Mind*. New York: Routledge.

Pickard, H. (2010). "Schizophrenia and the epistemology of self-knowledge". *The European Journal of Analytic Philosophy: Special Edition in the Philosophy of Psychiatry, 6 (1)*, 55-

...

13

Self Awareness and the Self-Presenting Character of Abnormal Conscious Experience

▪

Pablo López-Silva

University of Manchester, UK

▪

Abstract: *Some philosophers suggest that a minimal form of self-awareness is an integral element of the way in which all experiences are given (SPC: self-presenting claim). The main argument for this is that the phenomenological quality of 'mineness' of the experience reveals the self as a part of all experiences. Since the sense of mineness is taken as intrinsic to the givenness of the experience, it counts as an argument for the SPC. In this essay, I assess this claim and its main argument. After describing the phenomenological approach to self-awareness that grounds the discussion, I comment on some pathological cases that challenge the SPC. After this, I examine the standard reply in defence of the SPC and I focus the discussion on cases of thought insertion. I conclude that although the standard reply adds interesting elements to the general discussion, it becomes philosophically problematic especially when it tries to deal with cases of thought insertion.*

Introduction

Some psychiatrists and philosophers have claimed that conscious experience has a 'self-presenting' character. This self-presenting claim can be stated as follows:

SPC: *A minimal form of self-awareness is always given as an integral element of the intrinsic intentional structure of all experiences. Phenomenal consciousness always involves self-awareness.*

One of the main arguments for the SPC is that the phenomenological quality of 'being mine' of the experience reveals the subject of that experience as an integral part of its givenness (Zahavi & Parnas 1998; Zahavi 2005; Gallagher 2012)[1]. Since the sense of mineness is taken as a constant feature of this givenness, it counts as an

argument for the SPC. In this context, no experience lacks the sense of mineness and all experiences are self-presenting. In this chapter, I shall assess these ideas. After providing some background for the discussion, I discuss some pathological cases that challenge the SPC. I then examine the standard defence of the SPC and I focus the discussion on cases of thought insertion. I conclude that although the reply adds interesting elements to the discussion, it is philosophically problematic and, for this reason, the main argument for the SPC is not compelling.

Zahavi's Phenomenological Approach to Self-Awareness

Dan Zahavi has become one of the main defenders of the SPC. He suggests that to have a conscious experience of, let's say, X is also to have a degree of awareness of one's self as the subject experiencing X (Zahavi & Parnas 1998, 691; Zahavi 2005, Chapter 5). For instance, while listening to a song, I am also aware that I am the one listening to that song. This does not seem controversial. However, Zahavi adds:

"The 'I', the subject of self-awareness is simply a feature of the givenness of the experiencer" (Zahavi 2002, 697)[2].

Here, Zahavi suggests that the subject of the experience is integral to the way in which experience are subjectively given, i.e. the subject of the experience is an inherent and primitive element of the experiential stream and its natural appearance (Zahavi 2000; 2012)[3]. Zahavi suggests that one arrives at this conception of self-awareness simply by inspecting the structure of conscious experiences (Zahavi 2005, Chapter 1). Quite interestingly, according to Zahavi, this primitive self-awareness is 'pre-reflective', i.e. prior to any reflection on the experience and existing as a 'background presence' in the givenness of all experiences (Zahavi 2005, 115-146; Gallagher 2012, 127-130). This pre-reflective self-awareness constitutes the foundation for any type of reflective self-awareness (Gallagher and Zahavi 2010, 4). In this context, Zahavi seems to use the term 'pre-reflective' to describe a type of 'unattended' awareness of one's self that is part of the structure of the way in which experience appears to us. At least, this is the impression he gives when – following Husserl – he states that:

 "Self-consciousness, rather than being something that occurs during exceptional circumstances, namely whenever we pay attention to our conscious life, is a feature characterizing subjectivity as such, no matter what worldly entities it might otherwise be conscious of and occupied with" (Zahavi 2005, 11)[4].

In his work, Zahavi does not deny the existence of a 'higher order' self-awareness. However, he suggests that the pre-reflective type is a necessary condition for any other type of higher order self-awareness that Zahavi calls 'reflective self-awareness'. Let's contrast these two types of self-awareness:

(a) On the one hand, the so-called 'reflective self-awareness' is the product of directing our attention to the structure of our own occurrent mental states. This would imply a certain degree of cognitive and representational processing (Zahavi 2005). However, according to Zahavi, this takes us away from the most fundamental nature of self-awareness, namely, its pre-reflective givenness. Zahavi seems to understand the term 'reflective' in a Lockean sense i.e. as the mind's ability to turn its gaze inward upon itself (self-directedness), becoming the object of its contemplation (Locke 1975, 107). As a result of directing our attention to a certain mental state, we can reflectively be aware that we are in that mental state. An example of this type of self-awareness would be the so-called 'narrative self-awareness' (Gallagher 2000; Zahavi 2007).

(b) On the other hand, the 'pre-reflective' type is understood as an 'unattended' awareness of one's self that is 'always present' as an element of the common way in which experiences are given before directing our attention to them. Zahavi understands this type of self-awareness as part of the givenness of experience. The term 'pre-reflective' takes two meanings in this context: '(1) it is an [unattended] awareness we have before we do any reflecting on our experience; (2) it is an implicit and first-order awareness rather than an explicit or higher-order form of self-consciousness' (Zahavi & Gallagher 2010, 1). According to Zahavi, we do not need to do any reflection on the structure of experience in order to be pre-reflectively aware that we are the ones undergoing our experiences. Zahavi claims that the 'reflective self-awareness' just thematizes an awareness that belongs to the very same structure in which experience intentionally is, and hence: "an explicit reflective self-consciousness is possible only because there is a pre-reflective self-awareness that is an ongoing and more primary self-consciousness […]" (Ibid). Finally, Zahavi states the relation between reflective and pre-reflective self-awareness as follows: "Rather, the point is that this reflective self-awareness is derivative, and that it always presupposes the existence of a prior unthematic, non-objectifying, pre-reflective self-awareness as its condition of possibility" (Zahavi 2002, 17).

Some philosophers will deny the notion of 'pre-reflective self-awareness' (see Metzinger 2003; Rosenthal 1997). But we might all agree that for any conscious experience it is possible to reflect on it and to become aware of it as our own experience. However, this is not what Zahavi has in mind. Zahavi's claim is more controversial; he suggests that the reason why this is true is because of the existence of a pre-reflective self-awareness that is part of the structure of the givenness of the experience. However, it is reasonable to say that this idea requires further considerations. For instance, Thomas Metzinger (2003) does not deny the existence of a reflective self-awareness when we reflect on our experiences. However, he says

that this self-awareness has nothing to do with pre-reflective preconditions since it is an illusion created by a self-system. The main idea is that experience itself does not have a self-presenting structure and therefore, we cannot attribute to it any type of self-awareness as integral to its givenness. Let's now examine Zahavi's main argument for his claim.

Self-Awareness and Sense of Mineness

Zahavi proposes that in order to be reflectively aware of ourselves, we need to be familiarized with ourselves in a more fundamental way, namely, with our own ongoing experiencing. In Zahavi's view, the SPC is revealed through the analysis of the concept of pre-reflective self-awareness. According to Zahavi, the way in which experiences are given contains certain basic phenomenological features. These features are 'pre-reflective' in the sense that they are prior to any reflection on the experience, and the most important of which is the so-called 'sense of mineness'. This will be Zahavi's main argument for the SPC.

Zahavi (2005) suggests that the different modes of givenness of the experience (imaginative, perceptual, cognitive, etc.) share a common feature. All of them are marked as seeming to be my experiences; all of them have a 'sense of ownership or mineness'. Gallagher & Zahavi (2010) claim that: "the experience [always] happens in an immediate way and it is implicitly marked as my experiencer" (Gallagher & Zahavi 2010,1)[5]. Zahavi writes:

"Thus, whenever I experience something, my self […] is present and is, so to speak, therefore implicated. The experiences do not simply pass me by, as if they were foreign entities, but rather they are exactly mine" (Zahavi 2005, 81).

Zahavi claims that the sense of mineness is intrinsic to all experiences in the sense that: 'whether a certain experience is experienced as mine or not, however, depends not on something apart from the experience, but precisely on the givenness of the experience' (Ibid, 124). Importantly, Zahavi suggests that the sense of mineness of the experiences captures an immediate and non-observational access to oneself. Since this property is pre-reflective and reveals the self as a part of the structure of any experience, the sense of mineness is taken as an argument to say that conscious experience is self-presenting.

In sum, the sense of mineness is taken as intrinsic to the structure of the givenness of all experience (all experiences are given as mine) and it needs to be taken as an argument for the claim that suggests that a minimal self-awareness is revealed as a structural constant of the givenness of experience. This is because the sense of mineness presents the experiences as being mine, i.e. presenting me as the subject of those experiences.

Disrupted Sense of Mineness and the SPT

It is hard to imagine what abnormal mental phenomena feel like. However, it is widely agreed that abnormal experiences constitute a valuable contribution to understating the most essential features of phenomenal consciousness and self-awareness (Parnas & Zahavi 2003). These reports have been taken as a plausible way of testing theories in phenomenology, philosophy of mind, neuroscience and cognitive science (Gallagher 2000; Parnas & Zahavi 2000; Metzinger 2003, 439). Conversely, a theory which fails to accommodate pathological cases may be problematic.

In the context of our discussion, pathological cases confront us with situations that challenge the argument from the sense of mineness. Since our discussion of the SPC concerns the way in which experiences are given to us, it is extremely important to pay careful attention to the phenomenology of these cases in order to raise plausible conclusions on this debate. In this section, I shall examine the claim that no experience lacks the 'sense of ownership' that underlies the SPC. There seem to exist an important number of empirical cases that suggest that it is possible to have experiences with no sense of mineness and therefore, that the argument from the sense of mineness is false. The most challenging cases come from people suffering from thought insertion in schizophrenia. For this reason, let me start with the weaker cases to finally focus our discussion on cases of disrupted experiences of thinking.

Disrupted Bodily awareness

Bodily awareness denotes a certain type of experience that presents our body from the first person perspective, including proprioception, kinaesthetic experiences, bodily sensations and the sense of balance (Martin 1995). Usually, we experience a sense of mineness with respect to the experience of a body part, i.e. in normal cases the experience of our body part is marked as mine. However, there are cases in which this condition is not met, which suggests that this property would not be constant in the givenness of experience, contrary to Zahavi's proposal. For instance, people suffering from somatoparaphrenic delusion deny the sense of mineness of the experience of a certain body part. With the lack of this experiential property, patients attribute the ownership to someone else or they personify it. In these cases, patients receive sensory stimulation from the alien body part; however, the experience of that body part lacks the sense of mineness, i.e. the experience of the alien limb is not marked as their own. Phenomenologically, patients suffering from this delusion experience some body part as not theirs. We might suggest that the lack of the experiential quality of mineness plus a delusional general state makes the patient claim that the owner of that certain body part is someone other than the patient. A similar case – and perhaps clearer – is given by the alien hand syndrome

in which patients although feeling the sensational properties of a body part, do not experience the body part as their own[6]. Patients suffering from this condition deny the sense of ownership for the experience of a body part.

These cases give us prima facie evidence to suppose that at the phenomenological level, the sense of mineness can be somehow dissociated from a certain type of experience. This last distinction is important because Zahavi seems build his whole account specifically at the phenomenological level. If we take the reports of these patients at face value, we might suggest that: 'such people experience bodily sensation in a limb that they do not experience as their own. That is, they posses bodily-awareness without a sense of ownership' (Smith 2013, 10). Although these patients have sensorial inputs, these inputs do not imply the patients having the experience of the body part with a sense of mineness.

Disrupted Action Awareness

In normal cases, we feel a sense of mineness attached to our conscious actions. Usually, from the first person perspective, the awareness of an action reveals the action as mine. However, there are cases in which this experience is disrupted. Biran et al present the following case:

"JC [56-year-old right-handed man] attributed wishes and plans to his right limb. These attributions were evident in statements such as "It has a mind of its own", "Wants to be the boss", "Its own way", "Wants to dust the way it wants", "It won't go the way I want", "Controls the towel while I dry myself"." (Biran et al 2006, 567)

In the so-called anarchic hand syndrome, patients are aware of their actions but they do not have the experience of those actions as their own (Marchetti & Della Sala 1998). In these cases, patients receive sensory information given by the movements of the hand; however, the experience of that movement is not phenomenologically given as 'their' experience. From these cases we can raise a similar conclusion to the one offered in the last section. If we take the patients' reports at face value, we can suggest that they experience action awareness with no sense of ownership (De vignemont 2007).

Disrupted Experience of Conscious Thinking

The phenomenon called 'thought Insertion' – TI henceforth – is widely recognized as one of the most significant symptoms of schizophrenia (first-rank symptoms; Mellor 1970)[7]. As a matter of fact, TI on its own is enough for a diagnosis of schizophrenia (F20, ICD-10; World Health Organization 1992). Smith (2013) refers to these cases as a disruption of the sense of ownership of intellectual experience i.e. 'the experience of consciously thinking'. Smith suggests that:

[In normal situations] when I think through a problem I am aware of the problem itself, but it is also manifest to me that it is me that is thinking about it. That is, there is a sense of ownership over my episodes of conscious thinking (Smith 2013, 11)

In contrast with normal cases, in TI, subject experiences thoughts which are not his own intruding into his mind (Mullins & Spence 2003). As Sims writes:

"[In TI, a patient] experiences thoughts that do not have the feeling of familiarity, of being his own, but he feels that they have been put in his mind without his volition, from outside himself" (Sims 2003, 168)

In these cases, patients do not experience a sense of mineness attached to the episode of conscious thinking. In TI, the 'conscious thinking' is not given as having the phenomenal property of mineness (Metzinger 2003). Quite often, patients suggest that the real owner has inserted the thought into their minds (Mellor 1970). The following reports are instances of TI:

"I have never read nor heard them; they come unasked; I do not dare to think I am the source but I am happy to know of them without thinking them. They come at any moment like a gift and I do not dare to impart them as if they were my own."
(Jaspers 1963, 123)

"I look out of the window and I think the garden looks nice and the grass looks cool. But the thoughts of Eammon Andrews come to my mind. There are no other thoughts there, only his. He treats my mind like a screen and flashes his thoughts onto it like you flash a picture"
(Mellor 1970, 17)

"Thoughts are put into my mind like 'Kill God'. It's just like my mind working, but it isn't. They come from this chap, Chris. They're his thoughts."
(Frith 1992, 66)

"Sometimes it seemed to be her own thought '(…)but I don't get the feeling that it is.' She said her 'own thought might say the same thing (…)But the feeling it isn't the same . . . the feeling is that it is somebody else's (…)"
(Hoerl 2001, 190)

"I didn't hear these words as literal sounds, as though the houses were talking and I were hearing them; instead, the words just came into my head—they were ideas I was having. Yet I instinctively knew they were not my ideas. They belonged to the houses, and the houses had put them in my head"
(Saks 2007, 29)

Taking the patients' reports at face value, we can prima facie suggest that patients suffering from TI do not experience their thoughts as their own and therefore, from the first person perspective, they seem to lack a sense of mineness attached to intellectual experience[8]. In light of this, there are good reasons to think that the described cases provide plausible evidence to suppose that the 'givenness' of some experiences is not accompanied by a sense of mineness (as the SPC presupposes). We do not deny that there is 'something it is like' to be in the described states. It is quite natural and plausible to suggest that there is something it is like to experience an anarchic hand episode or an inserted thought that is qualitatively different from experiencing a full owned action or a full owned thought. However, the described cases show that the phenomenological sense of mineness can be dissociated from certain experiential modalities and therefore, this show that this property would not be integral to the 'givenness' of experience, as Zahavi proposes.

The described cases show a disassociation between the metaphysics and the phenomenology of experience. The hand that a patient suffering from somatoparaphrenic delusion is aware of is necessarily the patient's hand, however, the experience of it is not given as having a sense of mineness. The hand is not experienced as the patient's hand. There seems to be a dissociation that occurs at the phenomenological level. Hence, taking at face value the reports, we cannot be completely sure that all experiences reveal a minimal pre-reflective form of self-awareness (minimal self) as a constant structural element (SPC). This is because, on Zahavi's view, the presence of a sense of mineness associated to each token experience is a condition for the self to be revealed in the experience. As we have seen to some patients certain experiences are not phenomenologically given as their own.

Similar interpretations of these cases are found in different authors. For instance, Metzinger (2003) has suggested that cases of TI present patients alienated from their own thoughts for which they experience no sense of ownership or agency[9]. Explicitly referring to Zahavi's view, Metzinger (2006) proposes that these types of cases demonstrate that the sense of mineness is not 'an invariant dimension of the first-personal givenness' as Zahavi claims (cf. Zahavi 2005, 143). In cases of thought insertion, patients might have introspective access to their own mental states, but still experience these states not only as being influenced or controlled by someone else, but as alien, as belonging to another. In line with this, Smith (2013) concludes from the revision of similar pathological cases that: 'it seems reasonable to say that, in some sense at least, people can become alienated from their own bodily experience, action, memory and conscious thought' (Smith 2013, 12). Based on this, he concludes that some pathological cases support the claim that some experiences can lack a sense of ownership. In conclusion, it seems plausible to say that it is

possible to have experiences with no sense of mineness attached to them, for this reason the main argument for the SPT is not compelling.

Are cases of 'disrupted sense of mineness' an argument against the argument from the sense of mineness? Zahavi's Reply

Zahavi (2005) is well aware of the existence of pathological cases that challenge the argument from the sense of mineness. In this section I shall examine some of the main aspects of the reply offered by Zahavi and the problems that each aspect faces when it tries to deal with cases of disrupted episodes of conscious thinking.

Zahavi claims that the most challenging pathological cases for his view come from the discussion of thought insertion in schizophrenia (Zahavi 2005, 3-4). However, although challenging, the author concludes that they do not show the disruption of the sense of mineness, but rather the disruption of the 'sense of agency'[10] (Zahavi 2005, 143-144). Gallagher (2012, 132) distinguishes these two senses as follows:

Sense of agency: The pre-reflective experience that I am the one who is causing or generating a movement or action or thought process.

Sense of ownership: The pre-reflective experience that I am the one who is moving or undergoing an experience.

Zahavi claims that in normal cases sense of agency and sense of ownership occur simultaneously, but in abnormal cases such the described ones, they can come apart. Zahavi claims:

"Whereas the sense of agency refers to the sense of being the initiator or source of an action or thought, the sense of ownership refers to the sense that it is my body that is moving, that the experience I am living through are given as mine […] In cases of involuntary action [as in anarchic hand syndrome], the two [senses] can come apart. If I am pushed or if I am undergoing spasms, I will experience ownership of the movement(s)-I rather than somebody else, am the one moving-but I will lack a sense of agency; I will lack the experience of being the agent or initiator of the movement" (Zahavi 2005, 143-144)

This reply is based on the standard approach to thought insertion proposed by Stephens & Graham (2000) and Gallagher (2000a; 2000b)[11]. It is important to note that the distinction between agency and ownership was originally introduced by Stephens & Graham (2000) at the level of attributions and for this reason, in their discussion the term 'sense of agency' is treated as a result of a (reflective) disposition to make judgments about experience and not as an intrinsic (pre-reflective) property of the way in which experiences are given. However, Gallagher introduces this distinction at the experiential level by suggesting that the 'sense of agency'

belongs to the intrinsic –normal- appearance of experiences, i.e. pre-reflective[12]. Therefore, it is conceptually important to distinguish between the experience of mineness and the judgment (or attribution) of mineness in this discussion (De Hann & De Bruin 2010). Nevertheless, although this distinction introduces several interesting insights to the discussion, there are several reasons why it is problematic:

First, the distinction looks plausible in cases of action awareness in which one can clearly phenomenologically distinguish between voluntary and involuntary movement, namely, as one being the author or having the intention of performing such actions. In fact, it is widely agreed that the standard approach to thought insertion establishes a parallelism between the loss of agency in cases of action with the loss of agency in thinking (Stephens & Graham 2000; Gallagher 2012). However, this is problematic. As in the case of actions, this model presuppose an 'intention to think' in any case in which the agent thinks (Mullins & Spence, 2003; Cermolacce, Naudin & Parnas, 2007). In cases of thought insertion, this 'intention' (T) would be disrupted resulting in a loss of the sense of agency. However, this leads to an infinite regress: The 'intention' (T) itself needs to be a type of thought that accompanies a thought (T*), and this type of thought (T*) needs to have a previous intention that is a type of thought (T**) and so on ad infinitum. Indeed, if every act of thinking is explained by a prior intention to think, the model leads to an infinite regress whenever and agent thinks.

Secondly, the distinction cannot offer an account of the unique phenomenology of thought insertion, or at least: 'it makes thought insertion lose all of its puzzling features' (Bortolotti & Broome 2009, 217). The most puzzling aspect of the phenomenology of thought insertion is the 'alien' character of the thought that is experienced by the patient. In fact, this is the main aspect that distinguishes thought insertion from obsessive thought, thought withdrawal and influenced thinking (Mullens & Spence 2003). However, by explaining thought insertion on the basis of a disrupted sense of agency and a retained sense of mineness, Zahavi's reply cannot account for the alien character of the phenomenon. In fact, this account cannot discriminate between normal cases of spontaneous thoughts, obsessive thoughts and cases of thought insertion. In everyday life, many thoughts pop into our mind without being preceded by specific intentions –as Zahavi's distinction presupposes-. However, this does not lead us to experience these thoughts as 'not being mine' i.e. with an alien character. Similarly, Jaspers (1963) claims that obsessive thoughts arise against the patient's will. This leads the patient to perform different obsessive rituals as attempts to suppress the disrupting obsessive thought. Obsessions have been explained on the basis of a disrupted sense of agency and a retained sense of mineness. However, obsessive thoughts lack the 'alien' quality that inserted

thoughts present. Based on these two cases, it seems plausible to say that the distinction fails to offer a distinct account of the phenomenology of thought insertion different from intrusive thoughts and obsessions. This distinction cannot capture the characteristic conflict that a patient suffering from thought insertion faces.

Trying to reply to these objections, Zahavi is forced to make an additional move to deal with the no-sense-of-mineness cases. This move is given by the author's way of understanding experience. Let me explain this point. Following Williams James and against transparency philosophers, Zahavi claims that experiences are always something given to somebody; there are no 'free-floating' experiences[13] (Gallagher 2012). In experience –through them-, objects are given as having certain properties, 'as appearing in a certain way' (Zahavi 2005, 121). There is therefore, a double dimension to the structure of the givenness of all experiences. According to Zahavi, there is an important difference between asking about the properties that a certain object appears as having (or what the object looks like to the observer) and asking about the properties of the experience of that certain object (or what the perceiving feels like to the perceiver) (Ibid 2005, 121). The former pertains to the properties of the perceived object -which Zahavi calls sensed-, while the latter pertains to the properties of the act of perceiving –which Zahavi calls sensing-. From this, Zahavi claims that: 'We cannot be conscious of an object [a song] unless we are aware of the experience through which this object is made to appears [my listening]' (Ibid, 121) and exemplifies his point by claiming that: 'The taste of the lemon is a qualitative feature of the lemon and must be distinguished from whatever qualities my tasting of the lemon has' (Ibid 2005, 119). Finally, Zahavi suggests that although these two aspects –sensing and sensed- can be conceptually distinguished: 'they cannot be detached and encountered in isolation from one another' (Ibid, 123). Taking this into consideration, Zahavi concludes that in cases of disrupted sense of mineness in action awareness and experiences of conscious thinking (TI), the act of perceiving, acting or thinking (different modes of sensing) is always given as mine, i.e. I am the one undergoing those modes of sensing. In cases of thought insertion, the author concludes that the thought is not given as mine in the sense that I am not the source of that thought, but I am the one experiencing the thought as not being mine (Ibid, 144).

Zahavi's model of experience looks plausible in certain case, such as cases of visual perception. One can clearly distinguish between the way in which, say the chair is given (sensed) and the way in which my experience of the chair is given (sensing). The same applies in cases of bodily perception where one can distinguish, say my hand as having certain properties (sensed) and my experiencing (sensing) [14]. Therefore, it seems reasonable to say that this model deals in a more or less plausible

way with cases of disrupted bodily awareness. The model struggles to account for cases of somatoparaphrenic delusion, but for the sake of argument we can ignore this point here. However, this model leads us to some unavoidable problems when it is applied to cases of conscious thinking.

First, as has already been pointed out by Metzinger (2006), a quite general worry is that Zahavi does not explain how the conceptual distinction between sensed and sensing grounds the distinction between the sense of agency and mineness. He explains neither why nor how the former supports the latter.

Secondly and more relevant to our discussion, it is not clear if we can distinguish between the 'givenness of a thought' and the 'givenness of my thinking the thought'. The distinction is equally problematic in cases of inserted emotions, where it is difficult to distinguish between 'an emotion' and 'my experiencing the emotion'. In fact, it is completely plausible that no separation between a 'thought' and 'thinking the thought' can be made. As Cermolacce, Naudin & Parnas (2007) suggest: 'there is no separation here between thinking and a familiarity with what one is thinking' (p. 709). This is because the thinking is the thought itself and –in cases of emotions- the experiencing is the emotion. Therefore, when a patient reports an experience of inserted thought it is not conceptually accurate to propose that the patient preserves a sense of mineness for her thinking. This is because in cases of thought insertion, it is the act of thinking (the thought) that is not given as the patient's act of thinking (thought).

Finally and more importantly, this model of experience leads us to (yet) another infinite regress. Let's understand 'thinking' as a case of 'intellectual experience', just as Zahavi seems to do (Zahavi 2005, 124). Based on Zahavi's model, we need to distinguish between the way in which a thought is given (T: sensed) and the way in which my thinking a thought is given (TT: sensing). In cases of thought insertion (T) would not be given as mine, but (TT) still given as mine. This distinction becomes problematic because 'thinking the thought' (TT) is a type of experience and by definition, we need to distinguish between the givenness of (TT) and the givenness of my thinking (TT) i.e. (TT*) and so on ad infinitum. Therefore, if we understand conscious thinking as a case of intellectual experience, Zahavi's distinction leads to an infinite regress whenever I have an intellectual experience.

Conclusion

Dan Zahavi develops an interesting and philosophically detailed phenomenological approach to the relationship between phenomenal consciousness and self-awareness. In doing so, he suggests that conscious experience has a self-presenting character (SPC) via a sense of mineness. This essay has offered a philosophical assessment of his main argument, focusing on the claim

that no experience lacks a sense of mineness. Roughly speaking, Zahavi proposes that all experiences include a sense of mineness as a part of the way in which they are given. However, reports of people suffering from different pathologies challenge this idea. From the revision of cases of somatoparaphrenic delusions, alien hand syndrome and thought insertion, among others, we concluded some experiences can lack sense of mineness. Zahavi is well aware of the existence of these cases and he elaborates a reply based on the standard approach to thought insertion. However, the author's reply presents several philosophical problems, especially in relation to cases of thought insertion. In analysing cases of thought insertion, Zahavi makes a distinction between the sense of agency and the sense of mineness, which establishes an implausible parallelism between the phenomenology of action and conscious thinking, eventually leading to an infinite regress. In addition, his distinction cannot discriminate between the phenomenology of obsession and everyday disruptive thoughts from cases of disruption of conscious thinking thought. The distinction explains neither the uniqueness of the phenomenology of thought insertion nor the conflict that patients feel when they experience this abnormal experience. Trying to save his argument from these kinds of objections, Zahavi makes a problematic distinction between 'the givenness of an object' (sensed) and the 'givenness of my experience of the object' (sensing). This distinction is plausible in cases of perceptual experience, but it is conceptually inaccurate when applied to cases of conscious thinking and leads to another infinite regress. Finally, based on the philosophical examination offered in this essay, I conclude that although the standard reply adds interesting elements to the discussion, it is philosophically problematic in cases of thought insertion. For this reason, we may conclude that the main argument given by Zahavi in support of the SPC is implausible. This does not refute the SPC, but shows that stronger arguments are needed to defend it.

Endnotes

1   I will follow Dan Zahavi – the main defender of SPC – and refer to this property as the sense of *mineness* and the *sense of ownership* interchangeably

2   This must not be taken as a metaphysical claim. In his project, Zahavi (2005) takes a Husserlian phenomenological position towards the concepts he wants to explain. This means that Zahavi tries to avoid any theoretical and metaphysical presuppositions regarding the structure of conscious experience, making claims only from the way they appear

3   This raises further questions about a possible metaphysical identification of the experiential stream with a 'minimal self' (Zahavi 2011). However, our discussion will be focused on the phenomenological discussion. See note 2

4   The author uses the terms 'self-awareness' and 'self-consciousness' interchangeably (see: Zahavi 2005, 225)

5   It is important to note that here 'implicit' means 'pre-reflective'

6   Moro, Zampini & Aglioti 2004, 440 write about cases of alien hand syndrome: "Stimuli were detected in all trials even though the hand was *still felt as belonging to another person*. Specifically asked about how it was possible to perceive stimuli delivered to another's hand, the patient reported that 'many strange things can happen in life.' "

7   The history of the term is full of disagreement and philosophical discussion. For a nice summary of this history (see: Mullins & Spence 2003)

8   A similar case can be found in cases of inserted emotions in which patients refer to the way that certain emotional states are given to them not as their own states: I cry, tears roll down my cheeks and I look unhappy, but I have a cold anger because they're using me in this way, and it's not me who's unhappy, but they're projecting unhappiness onto my brain. They project upon me laughter, for no reason, and you have no idea how terrible it is to laugh and look happy and know it's not you, but their emotions'. (Mellor 1970, 17)

9   The sense of agency is defined as: "The pre-reflective experience that I am the one who is causing or generating a movement or action or thought process" (Gallagher 2012, 132).

10 For a specific discussion about the definitions of this term, see: De Haan & Leon de Bruin (2009)

11 For critical discussion of this approach, see: Bortolotti & Broome (2009) and Martin & Pacherie (2013)

12 In his reply, Zahavi explicitly adopts Gallagher's distinction (See: Zahavi 2005, 143-144)

13 James (1980 I, 225) claims that: 'Every thought is part of a personal consciousness […] the universal fact is not 'feelings and thoughts exists', but 'I think' and 'I feel', [yet] to give an accurate account of [this] is the most difficult of philosophical tasks'

14 People like De Haan & De Bruin (2009) will criticize this idea

References

Bayne, T. (2006). "The Feeling of Doing: Deconstructing the Phenomenology of Agency." In: N. Sebanz &. Prinz (eds.). *Disorders of Volition*. Cambridge, MA: MIT Press, 49-68

Billon, A. (2011). "Does consciousness entail subjectivity? The puzzle of thought insertion." *Philosophical Psychology 26(2)*, 291-314

Block, N. (1995). "On a confusion about a function of consciousness." *Behavioral and Brain Sciences 18*, 227-247

Bisiach, E. et al (1991). "Remission of somatoparaphrenic delusion through

vestibular stimulation." *Neuropsychologia 29*, 1029–1031

Biran, I. et al (2006). "The alien hand syndrome: What makes the alien hand alien?" *Cognitive Neuropsychology 23 (4)*, 563-582

Bortolotti, L. & Broome, M. (2009). "A role for ownership and authorship in the analysis of thought insertion." *Phenomenology and the Cognitive Sciences 8*, 205–224

Carruthers, P. (1996). "Language, Thoughts and Consciousness. An Essay in Philosophical Psychology". Cambridge: Cambridge University Press

Chalmers, D. (1996). *The conscious Mind: In search of a fundamental theory*. New York: OUP

Cermolacce, M. et al (2007). "The ''minimal self'' in psychopathology: Re-examining the self-disorders in the schizophrenia spectrum." *Consciousness and Cognition 16*, 703–714

Della Sala, S. et al (1994). "The Anarchic hand: a fronto-mesial sign." In: *Handbook of Neuropsychology 9*, 233-255

De Haan, S. & De Bruin, L. (2009). "Reconstructing the minimal self, or how to make sense of agency and ownership." *Phenomenology and the Cognitive Sciences, 9*, 373–396

De Vignemont, F. (2007). "Habeas Corpus: The Sense of Ownership of One's Own Body." *Mind & Language 22 (4)*, 427–449

Flanagan, O. (1992). *Consciousness Reconsidered*. Cambridge, MA: MIT Press

Frank, M. (1984). *Was ist Neostrukturalismus?* Frankfurt am Main: Suhrkamp

Frith, C. (1992). *The cognitive neuropsychology of schizophrenia.* Hove: Lawrence Erlbaum.

Gallagher, S., Zahavi, D. (2010). "Phenomenological Approaches to Self-Consciousness", In: Edward Zalta (ed.). *The Stanford Encyclopedia of Philosophy (*Winter 2010 Edition), URL = <http://plato.stanford.edu/archives/win2010/entries/self-consciousness-phenomenological/>

Gallagher, S. (2012). *Phenomenology*. U.K.: Macmillan

___ (2004). "Neurocognitive models of schizophrenia: A Neurophenomenological critique." *Psychopathology 37(1)*, 8-19

___ (2000), "Philosophical Conceptions of the Self: Implications for Cognitive Science", *Trends in Cognitive Sciences 4*, 14–21

Gallagher, S., Zahavi, D. (2008). *The Phenomenological Mind.* London: Routledge

Guidano, V. (1994). *El sí mismo en proceso.* Barcelona: Paidós

Hidalgo-Borrajo, P. et al (2009). "Síndrome de la mano ajena: revisión de la bibliografía." *Revista de Neurología 48 (10)*, 534-539

Hoerl, C. (2001). "On thought insertion." *Philosophy, Psychiatry & Psychology 8*, 189-200

Howell, J. (2010). "Subjectivity and the elusiveness of the self." *Canadian Journal of Philosophy 40 (3)*, 459-483

Hume, D. (1888). *A Treatise of Human Nature*. Oxford. Clarendon Press

Jaspers, K. (1963). *General Psychopathology* .7th edition. Manchester: Manchester University Press

Jankovic, J. (1997). "Tourette syndrome. Phenomenology and classification of tics." *Neurologic Clinics 15(2)*,267-275

Krueger, J. (2011). "The who and the how of experience." In: Zahavi, D. et al (eds.). *Self, No-Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*. Oxford: Oxford University Press, 27-55

Locke, J. (1975). *An Essay concerning Human Understanding*. Oxford: Clarendon Press

Martin, M. (1995). "Bodily awareness: a sense of ownership." In: J. Bermúdez, et al (eds.). *The Body and the Self*. Cambridge, MA: MIT Press

Martin, J-M. & Pacherie, E. (2012). "Out of Nowhere: Thought Insertion, Owenership and Context-Integration." *Consciousness and Cognition* http://dx.doi.org/10.1016/j.concog.2012.11.012 [Author Manuscript]

Marchetti, C. & Della Salla, S. (1998). "Disentangling the alien and the anarchic hand." *Cognitive Neuropsychiatry 3*, 191–207

Mellor, C. (1970). "First Rank symptoms of schizophrenia." *The British journal of Psychiatr 117*, 15-23

Metzinger, T. (2006). "Reply to Zahavi: The Value of Historical Scholarship." *Psyche 12(2)*, 1-4

___(2003). Being no one. Cambridge: MIT Press

Moro, V. et al (2004). "Changes in spatial position of hands modify tactile extinction but not disownership of contralesional hand in two right brain-damaged patients." *Neurocase 10*, 437–443

Mullins, S. & Spence, S. (2003). "Re-examining thought insertion." *British journal of psychiatry 182*, 293-298

Nagel, T. (1974). "What is it like to be a bat?"In: *Philosophical Review 83*, 435-450

Parnas, J. & P. Handest (2003). "Phenomenology of Anomalous Self-Experience in Early Schizophrenia."*Comprehensive Psychiatry 44 (2)*, 121-134

Parnas, J. & Sass, L. (2001). "Self, Solipsism, and Schizophrenic Delusions." *Philosophy, Psychiatry, & Psychology 8 (2/3)*, 101-120

Parnas, J. (2003). Self and schizophrenia: A phenomenological perspective. In T. Kircher & A. David (eds.). *The self in Neuroscience and Psychiatry*. Cambridge: Cambridge University Press, 217–241

Rosenthal, D. (1997). "A Theory of Consciousness." In: N. Block et al (eds). *The Nature of Consciousness*. Cambridge, MA: MIT Press, 729–753

Saks, E.R. (2007). *The centre cannot hold. My journey through madness*. New York: Hyperion

Sartre, J.-P. (1956[1943]). *L'Etre et le néant.* Paris: Tel Gallimard; English translation: H. E. Barnes. Being and Nothingness, New York: Philosophical Library

___ (1957[1936]). "La transcendance de l'ego." Paris: Vrin; English translation: F. Williams and R. Kirkpatrick. *The Transcendence of the Ego*, New York: The Noonday Press,

Searle, J.R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press

Schneider K. (1959). *Clinical Psychopathology*. Transl. M.W. Hamilton. New York, NY: Grune & Stratton, 1959

Sollgerber, M. (fc.). "Making sense of an endorsement model of thought-insertion." Mind & Language

Smith, J. (2003). "Alienation and self-presentation." Consciousness online. Unpublished manuscript. Available from: http://consciousnessonline.files.wordpress.com/2013/02/smith-co5.pdf

Strawson, G. (2009). *Selves*. Oxford: OUP

Zahavi, D. & Parnas, J. (1998). "Phenomenal consciousness and self-awareness: a phenomenological critique of representational theory." *Journal of Consciousness Studies 5 (5–6)*, 687–705

Zahavi, D. (2008). *Subjectivity and Selfhood: Investigating the first-person perspective.* Cambridge, MA: The MIT Press

___ (2007). "Self and other: The limits of narrative understanding." In D. Hutto (ed.): *Narrative and Understanding Persons.* Royal Institute of Philosophy Supplement 60. Cambridge: Cambridge University Press, 179-201

____ (2002). "First-person thoughts and embodied self-awareness: Some reflections on the relation between recent analytical philosophy and phenomenology." *Phenomenology and the Cognitive Sciences 1*, 7–2

Schizophrenia is the Collapse of Self-Recognition.
What Self?

.

José Luís Pio Abreu
CFCUL, Universidade de Lisboa, Portugal

.

Trying to understand schizophrenic patients is a challenge for clinical psychiatrists. How is it possible for a person to write something and be convinced that another person has written the text? Or have the conviction that his thinking is shared and commented on by everyone? Or believe that the secret movement of his internal organs is controlled and manipulated by other people? In addition of being a source of disturbance, these questions threaten our common sense about the self and its stability. The self is central to our cognitive activities and vision of the world, and several authors have been looking for this elusive concept. Analyzing what happens in schizophrenic patients may be a way of discovering how it is constructed. The common idea of the self may be more complex than the primitive and autobiographical selves that have been considered until now. We need perhaps to envisage the existence of an intentional self.

1. The neurobiology of schizophrenia

The predisposition to schizophrenia is now understood as difficulties in having problems with brain connectivity by way of long fibers, which link distant zones of the neo-telencephalon (Fornito et al 2011). These difficulties are due to several genetic variations and early environmental aggressions that disturb the development and organization of long fibers in the telencephalon. Some structural brain anomalies have also been reported, such as large brain ventricles, persistence of the cavum septi pellucidi, which is a sign of immaturity, the presence of ectopic gray matter, and partial agenesis of the corpus callosum (Degreef et al 1992; Sherton et al 1997; Andraesen 2001). Schizophrenic patients and their families also have more symmetrical activity while performing a simple task, where normal subjects have more asymmetrical activity (Spence et al. 2000). With all these difficulties, pre-schizophrenics and schizophrenic patients can achieve the same results in complex

tasks as normal controls, but with more chaotic, disorganized and laborious brain activity (Kim et al 2003). In the same task, healthy subjects only resort to the dorso-lateral pre-frontal cortex and the inferior parietal zone.

In daily life we perform very complex tasks: for instance, when evaluating reality, we are constantly alternating between our own intentions and the intentions of other people. Pre-schizophrenic subjects may have some difficulties with these tasks, in spite of giving correct responses after laborious brain activity. However, in certain phases of life, the pressure to perform complex tasks may increase, and patients give up performing them. Something in their brain may change, and the result may be psychosis.

2. Functions of the Telencephalon

Before analyzing these events, we should examine the role of those brain structures which are disrupted in pre-schizophrenic and schizophrenic subjects. The phylogenetic and ontogenetic origin of the telencephalon is the hippocampus (Ariens-Kappers 1929). In humans, this small structure is hidden by the temporal lobe and large brain hemispheres, but it is dominant in primitive animals (fish and reptiles), where it is involved in recognition of territory, since it is linked to the olfactory pathways. Deprived of olfactory signals, primitive animals, such as salmon, are unable to find their territory (the source of the river where they were born) to spawn (Hassler & Larsen 1955). The hippocampus degenerates in human Alzheimer Disorder, in which patients lose memory and become disorientated in place and time. Place cells, which respond to specific places in the environment, were first discovered in the hippocampus. In Mammals, visual and auditive information, which was previously processed at lower brain levels (the mesencephalon) has risen to the telencephalon and also is at the service of territory recognition. But the hipocampus, which is at the base of memory organization, maintains its importance in recognition of places.

However, there is a threefold increase in volume of the telencephalon from evolved primates to man. We can now ask why we need such a large telencephalon. My provisional response is that we need it

(1) to handle signs – words, drawings and manufactured objects - which permits us to extend our territory, since signs are essential for all cognitive activities;

(2) to recognize other people and ourselves, since human territory is composed of other people. Humans have spread all over the world in families and groups, and they are now settled in large communities where they need to reconcile their own expectations with the expectations of other people. Finally, we need it

(3) to evaluate reality while shifting between our intentions and the intentions of others. Schizophrenic patients have specific difficulties in every one of these three domains.

3. Handling signs

The problem of signs will firstly be discussed. Several mammals, and dogs above all, use their urine and droppings to mark territory. They leave signs on their territory, which they then recognize by smell. Perhaps humans also deal with olfactory signs. However, we have learned to deal with spoken and written words.

Spoken and written signs presuppose the functional and anatomical asymmetry of the brain. This asymmetry arose with the asymmetrical activity of the limbs while Homo sapiens manufactured objects, made drawings and wrote. Representing the right arm, the left hemisphere processes words and signs in general. An injury in the left hemisphere produces aphasia: the patient can't name objects or words. He can't relate objects with words, nor words with other words (Cutting 1997). He can't relate signs with signs. However he recognizes the meaning: he may not name a pen, but he can use it properly or say that it is used for writing. On the contrary, if a patient has an injury in the right hemisphere, he can describe a pen, but may use it to brush his hair. He can't recognize the meaning of signs.

Thus, we can say that the left hemisphere processes the Saussurean signifier of the signs. Words, drawings or objects are signifiers which represent the world (Eco 1990). In its turn, the right hemisphere processes the signified, or the meaning, if we accept, as Wittgenstein (1995) and others proposed, that the meaning is the way we use the signs. In fact, a chair can be very similar to a stool, but what defines the different concept and name of each is how each is used. While perceiving or looking for a chair, we know that a chair is something where we can sit and lean back on, whereas we can't lean back on a stool. Thus, its concept and name must be different. We can also say that when we perceive a chair, its meaning is the possibility of sitting down and leaning back, and when we look for a chair, we have the intention of finding something to sit on and lean back on.

Timothy Crow (1998; 1999; 2004; Mitchell & Crow 2005), has developed arguments suggesting that meaning and intention have the same nature and are both processed in the right hemisphere. Difficulties in lateralization in schizophrenics disturb intentional activity and the ability to attribute meaning to signs. The most striking symptoms of schizophrenia, delusional mood (the sensation that all meaning in the world has changed) (Jaspers 1963) and delusional perception (attribution of an arbitrary and delusional meaning to a common perception), reveal difficulties in assigning meaning. The written and spoken narratives of schizophrenic patients are sometimes full of neologisms and idiosyncratic meanings. All these symptoms

reveal that schizophrenic patients have obvious difficulties while attributing meaning, i.e. difficulties in linking Saussurean signifiers and signified (Crow 2004).

4. Recognition of self and others

The recognition of self and others, which is another complex task, will now be examined. If a child is in front of a mirror, and has a visible but unexpected mark on her nose, she goes directly to her face to remove the mark. This shows that the child recognizes herself in the mirror. However, younger children (under two years of age) try to remove the mark from the image in the mirror (Gallup et al 2003), meaning that they do not recognize themselves. The mirror experiment has been tested with other animals. Some evolved primates also recognize themselves in the mirror, but only when they are adults. This shows that self-recognition is a learned task.

Children recognize themselves only after a long period of imitation activities and pretend play, during which they first recognize the permanence of objects and specifically significant other people. With these activities, children develop a folk physics and a folk psychology which is now called Theory of Mind or "Mentalization" (Wellman & Lagattuta 2000). Later, they can change roles with other people and continue to organize their theory of mind, including the development of inter-subjective views and moral decision.

The universal hide-and-seek game, where a child tries to deceive playmates, is a way of learning that other people may have different mental states from herself, and that her own mental states belong only to herself. The sense of self-ownership may then emerge. Other components of the self may develop later. When adolescents play computer games, they develop the sense of self-agency. Having the intention of moving the ball, they can verify their success when the results (or meaning) match the intention. Sometimes they have to decide if they are the real agent, or if the movement of the ball is being directed by the intention of another playmate. This occurs when the results do not match their own intention and, in this case, they have to study and, so to speak, guess the player's intention in order to win the game. Schizophrenics have difficulties with the sense of self-agency (Hur et al 2013).

5. Evaluating Reality

The ability to shift between our own intention and the intention of others supports the daily exercise of evaluating reality, and this is where schizophrenics may have difficulties. Klaus Konrad (1959) describes two common situations to explain this exercise.

The first concerns being on a train at the station. When the train begins to move, we have the sensation that we are static, and it is the station and people on the platform that are moving. We correct this view by putting ourselves in the shoes of the people on the platform, then realizing that, in fact, we are the ones who are moving.

The second situation is closer to the schizophrenic experience. If we are sitting at home, near a window, and hear someone whistling outside, we immediately assume it is intended to attract our attention, perhaps calling us as if we were a dog. This is the experience of self-reference common in schizophrenics. However, we can correct this idea if we look through the window and see that the man is behind a dog. Putting ourselves in the shoes of the man, we realize that his intention is to call the dog and not us. In fact, we don't need to look through the window: we do this exercise in our imagination. Schizophrenics have difficulty in alternating between intentions and finding alternative meanings. Thus, they feel that everything is related to them (self-reference) and they have difficulties in evaluating reality objectively (Raballo et al 2006). A common example of alternating between intentions is the presentation of ambiguous pictures. A drawing can represent simultaneously a duck and a rabbit, and we can see one or the other, depending on our intention. Ambiguous pictures are now used as tools to help schizophrenic patients to develop their cognition (Moritz & Woodward 2007).

6. The psychotic breakdown

Thus, the main problem of the schizophrenic patient is alternating between his and others' intentions, and assuming alternative meanings. All through life, there are situations where self-recognition and recognition of self-agency is decisive but laborious. These are the situations that might precipitate a schizophrenic episode.

Epidemiological studies reveal that expressed emotions in the family can precipitate schizophrenic episodes (Butzlaff & Hooley 1998). These are situations where ones own and others' interests, feelings and expectations are evaluated and sometimes negotiated, leading to a great effort being made to understand other people and alternate between intentions of the self and others.

The first episode generally occurs in young adults, when they search for new intimate relationships and break the boundaries of personal spaces, leading to personal fusion. Sometimes, a conflictual affair, where the mind must alternate between the self's and the other's expectations, leads directly to a psychotic episode. Schizophrenia is more prevalent in immigrant populations (Cantor-Graae & Selton 2005), and moving to a new community is another factor, which can precipitate psychosis. Predisposed young men joining the army have a psychotic breakdown when they arrive and not during or at the end of the mission, notwithstanding war stress (Steinberg & Durell 1968).

As schizophrenics make an additional effort in these complex tasks, something occurs in their brains. Cortical zones which process complex tasks can show a comparatively greater activity at first, but this declines when the difficulty increases (Jansma et al 2004). Thus, instead of cortical processing, the brain may look for a more economical (authomatized) mode of processing information (Broome et al 2005). It is not very clear what happens next, but the basal ganglia are involved in this change, leading to increased and deregulated dopamine activity with the creation of dopaminergic "aberrant saliences" (Kappur 2003). These events correlate with, and are used to explain, the beginning of the psychotic episode when the more visible symptoms of schizophrenia occur. However, and this is the main point in schizophrenia, these visible symptoms reflect the breakdown of the self. They were called the first-rank symptoms by Kurt Schneider (1966). Thus, the patient is thinking, but is listening to his thoughts as if they were spoken from outside. He believes that his thoughts are withdrawn from his head and widely broadcast by an outside agency. Sometimes he believes that someone inserts thoughts into his head. Frequently, he hears voices of other people arguing about him or commenting on his activity in the third person. These hallucinations have the same properties as inner language, and are produced in the same brain zones. (Jardri et al 2011), but they are attributed to other people instead of himself. Thus, schizophrenics have lost the ownership of some of their thoughts. In the same way, the patient complains that some feelings and impulses and his will are produced outside of himself, and his organs, body and facial expressions can equally be moved from outside. As with the feeling of ownership, the feeling of self-agency is also disturbed (Hur et al 2013).

7. The self in schizophrenia
In psychopathology, we usually say that schizophrenic symptoms disrupt the unity of the self (Pio Abreu 2008), as opposed to identity, which is disturbed in dissociative pathologies and is concerned with the temporal continuity of the self (or the autobiographical self). There are different pathologies that disturb other properties of the self.

Recently, the self has been studied in relation to the default network. The default network corresponds to the brain zones that become more active when the subject is not involved in external tasks. In this "stand-by" mode, some middle areas of the brain, such as the medial pre-frontal cortex, anterior cingulum, precuneus and posterior cingulum, become more active. Some lateral zones of the convexity, such as the parietal inferior zone, also belong to the default network.

Activities related to the self and inner activities, such as thinking about autobiographical memory, envisioning the future, theory of mind (empathizing) and moral decision-making, also activate zones belonging to the default network (Buckner et al. 2008). Damasio (2010) considers this zone, and especially the

posterior medial zone, as the substrate of the autobiographical self. However, the autobiographical self is a human development of the nuclear self, related to the midbrain, where Panksepp (2003) also localizes his primordial self. None of these selves are primarily affected by schizophrenia. Thus, we have to look for the self, which is disturbed in schizophrenia in upper areas of the brain.

These areas may be organized by learning, since the recognition of the self is also a learned task. In fact, some areas of the default network, namely the convex zones and the inferior parietal zone, are not present in young children (Fair et al 2008). The inferior parietal zone, especially in the right hemisphere, seems important, because it is very active when the subject has to decide between self-agency or an outside agency (Farrer et al 2003). We have seen this when speaking about moving balls in computer games, but another interesting situation involving tickling can be told.

People do not feel ticklish when they tickle themselves. Why not? It is now known that when I have the intention of doing something and I really do something that I am going to perceive, such as seeing a moving ball or feeling ticklish, an efferent copy of the command order runs through long fibers to inhibit the posterior sensitive zone. Thus, if I have the intention of tickling myself, I do not feel ticklish because the efferent copy inhibits my sensorial perception. There is an experiment where subjects can tickle themselves with a tool, which may then be manipulated by the researcher in order to tickle in a slightly different way to the intended movement (Blakemore & Frith 2003). In this case, normal people feel ticklish, but schizophrenic patients have some difficulty in feeling the difference, perhaps because they cannot organize the efferent copies through long fibers in the telencephalon very well.

Frith et al (2000) have proposed a neurocognitive model, which can explain these events. Once an action is intended, an efferent copy is sent to the parietal inferior zone in order to compare the intention with the perceived result and decide on self-agency. Furthermore, while the subject is concerned with the results, muscle or proprioceptive sensibilities and their re-afference copies also inform the parietal inferior zone automatically. If the results match the intention, everything proceeds with the sense of self-agency and ownership. However, if the comparator does not work, and it does not work in schizophrenia because long fibers are damaged, the inferior parietal zone signals an alarm demanding attribution to an outside agency. Thus, the patient becomes concerned with his own moving arm as if it was an object, and loses implicit knowledge. If the proprioceptive feedback is also difficult, he can lose the feeling of ownership too (Gallagher 2000). Losing ownership and self-agency properties of the self, schizophrenics additionally lose the common concept of the self, namely its unity.

231

To conclude, if we look at our current concept of the self, which is challenged when schizophrenic symptoms are experienced, we have to consider an upper intentional self, constructed by learning, and processed in the more recent areas of the brain, above the autobiographical (or narrative) self, the nuclear self and the primordial self.

References

Andraesen N. (2001). *Brave New World: Conquering Mental Ilness in the Era of the Genome*. Oxford: Oxford University Press

Ariens-Kappers C. (1929). *The Evolution of the Nervous System in Invertebrates, Vertebrates and Man*. De Erver F. Bohn

Broome, M. et al (2005). "What causes the onset of psychosis?" *Schizophr Res, 79 (1)*, 23-34

Buckner R. et al (2008). "The Brain's Default Network. Anatomy, Function, and Relevance to Disease." *Annals of the New York Academy of Sciences 1124, The Year in Cognitive Neuroscience 2008*, 1–38

Butzlaff, R. L., Hooley, J. M. (1998). "Expressed Emotion and Psychiatric Relapse. A Meta-analysis." *Arch. Gen. Psychiatry. 55(6)*, 547-552; doi:10.1001/archpsyc.55.6.547

Cantor-Graae, E., Selten, J. (2005). "Schizophrenia and migration: a meta-analysis and review." *Am. J. Psychiatry 162*, 12– 24

Conrad, K. (1958). *Die Beginnende Schizophrenie. Versuch einer Gestaltanalyse des Wahns*, Georg Thieme Verlag, Stuttgart. Trad. esp.: La Esquizofrenia Incipiente. Intento de un Analisis de la Forma del Delirio, Ed. Alhambra, Madrid-Mexico.

Crow, T. (1998). "Nuclear schizophrenic symptoms as a window on the relationship between thought and speech." *Brit. J. Psychiatry 173*, 303-309

___ (1999). "Commentary on Annett, Yeo et al., Klar, Saugstad and Orr: Cerebral asymmetry, language and psychosis — the case for a Homo sapiens-specific sex-linked gene for brain growth." *Schizophrenia Research 39*, 219–231

___ (2004). "Auditory hallucinations as primary disorders of syntax: An evolutionary theory of the origins of language." *Cognitive Neuropsychiatry, 9 (1/2)*, 125-145

Cutting, J. (1997). *Principles of Psychopathology: Two Worls, Two Minds, Two Hemispheres*. Oxford: Oxford University Press, 183-189

Damasio, A. (2010). *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon Books, 263-290

Degreef, G. et all (1992). "Abnormalities of the septum pellucidum on MR scans in first-episode schizophrenic patients." *AJNR 13*, 835-840

Eco, U. (1990[1973]). O Signo. Lisboa: Ed. Presença

Fair, D. et al (2008). "The maturing architecture of the brain's default network." *PNAS 105 (10),* 3659-3660; doi:10.1073/iti1008105.

Fornito, F. et al (2011). "General and Specific Functional Connectivity Disturbances in First-Episode Schizophrenia During Cognitive Control Performance." *Biol Psychiatry. 2011 Jul 1;70(1),* 64-72. doi: 10.1016/j.biopsych.2011.02.019. Epub 2011 Apr 22.

Frith, C. et al (2000). "Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action." *Brain Research Reviews 31,* 357–363

Gallagher, S. (2000). "Philosophical conceptions of the self: implications for cognitive science." *Trends in Cognitive Sciences 4 (1),* 14-21

Gallup Jr, G., et al (2003). "Self- Awaedness, Social Intelligence and Schizophrenia." In: T. Kircher and A. David (Eds). *The Self in Neuroscience and Psychiatry.* Cambridge: Cambridge University Press, 147-165

Hassler A., Larsen, J. (1955). "The homing salmon." *Scientific American 163,* 72-75

Hur, J. et al (2013). "The crisis of minimal self-awareness in schizophrenia: A meta-analytic review." *Schizophr. Res. (2013)* http://dx.doi.org/10.1016/j.schres.2013.08.042

Jansma, J. et al (2004). "Working memory capacity in schizophrenia: a parametric fMRI study." *Schizophr. Res. 68,* 159– 171

Jaspers, K. (1963[1946]). *General Psychopathology.* Manchester: Manchester University Press, 98-103

Kim J-J et al (2003). "Functional Disconnection Between the Prefrontal and Parietal Cortices During Working Memory Processing in Schizophrenia: A [15O]H2O PET Study." *Am. J. Psychiatry 160,* 919–923

Mitchell, R., Crow, T. (2005). "Right hemisphere language functions and schizophrenia: the forgotten hemisphere?" *Brain 128,* 963–978

Moritz, S. & Woodward, T. (2007). "Metacognitive Training for Schizophrenia Patients (MCT): A Pilot Study on Feasibility, Treatment. Adherence, and Subjective Efficacy." *German J. Psychiatry 10,* 69–78

Panksepp, J. (2003). "The Neural Nature of the Core SELF: Implications for Understanding Schizophrenia." In: T. Kircher, A. David (Eds). *The Self in Neuroscience and Psychiatry.* Cambridge University Press, 197-213

Pio Abreu, J. (2008). *Introdução à Psicopatologia Compreensiva* (5th edition). Lisboa: Fundação Calouste Gulbenkian, 88-96

Raballo, A. et al (2006). "Self-Centrality, Psychosis and Schizotaxia: A Conceptual Review." *Acta Biomed 77,* 33-39

Schneider, K. (1966). *Psicopatologia Geral.* São Paulo: Mestre Jou

Shenton, M. et al (1997). "A review of magnetic resonance imaging studies of brain abnormalities in schizophrenia." K. Krishnan & P. Doraiswamy (eds). *Brain Imaging in Clinical Psychiatry.* New York: Marcel Dekker, 297–380

Spence, S. et al (2000). "Functional anatomy of verbal fluency in people with schizophrenia and those at genetic risk: Focal dysfunction and distributed disconnectivity reappraised." *Brit. J. Psychiatry 176,* 52-60

Steinberg, H. & Durell, J. (1968). "A stressful situation as a precipitant of schizophrenic symptoms: An epidemiological study." *Brit. J. Psychiatry 114,* 1097–1105

Welman, H. & Lagattuta, K. (2000). "Developing Understanding of Minds." In: S. Baron-

Cohen, et al (Eds.). *Understanding Other Minds.* (2nd.edition). Oxford: Oxford University Press, 21-49

Witgenstein, L. (1995). *Tratado Lógico-Filosófico e Investigações Filosóficas,* 2ª. ed. Lisboa: Fundação Calouste Gulbenkian

15

How is our Self altered in Psychiatric Disorders: Neuropsychopathological Hypotheses

▪

Georg Northoff

Mind, Brain Imaging and Neuroethics Research Unit
Institute of Mental Health Research
Royal Ottawa Mental Health Centre,
University of Ottawa, Canada

▪

Special Issue

▪

Abstract: *The self is central in our experience and has often assumed to be necessary for any kind of consciousness in philosophy. Recent investigations in neuroscience demonstrate a particular set of regions like the cortical midline regions to be associated with the processing of stimuli specifically related to the self as distinguished from those remaining unrelated to the self. Furthermore, findings show close overlap between self-related activity and high levels of resting state activity in especially anterior midline regions. Interestingly, recent findings in psychiatric disorders like depression and schizophrenia show resting state abnormalities in exactly these regions that is, the cortical midline structures. Based on the psychopathology of the self and the recent neural findings of the resting state, I here develop specific neuropsychopathological hypotheses about the self in depression and schizophrenia that are amenable to future experimental testing.*

Introduction

You read these lines. You find them boring and your experience is thus signified by boredom. Who experiences this boredom? You. You are the subject of the experience of boredom. Without you as subject of this experience, you could not experience anything at all, not even boredom. This subject of experience has been described as the 'self'. It is your 'self' that makes it possible for you to experience things. The self is a necessary condition for the possible constitution of experience

and thus also consciousness. It is clear therefore that there is much at stake when it comes to the self. We thus need to discuss how to characterize and define the concept of self. Why is the self so important? Because we usually assume that somebody must have consciousness. Somebody speaks a language. And somebody acquires a second language when coming for instance to a new country. Without somebody we may remain unable to do all these things. Who though is this somebody? This is what is traditionally called self. Hence the self is of central relevance. Who though is this self and how is it altered in the psychopathology of psychiatric disorders like depression and schizophrenia?

This is the topic of the present contribution. Rather than going into much conceptual detail about the self [1-3], I here start right away with recent neuroscientific findings that demonstrate neural overlap between resting state activity and self-specific activity in especially anterior cortical midline regions. This is the first part of this paper. That is followed by a second part where I will focus on the self in depression where, after describing psychopathological and neural findings, I will develop what I describe as neuropsychopathological hypotheses. The same will then be done for schizophrenia in the third part of this paper. The focus on depression and schizophrenia allows me to develop specific neuropsychopathological hypotheses that are open to future experimental testing. However, due to the reasons of space, this will prevent me from giving a more general overview about the psychopathology of the self in other psychiatric disorders like in addiction (see for instance [4, 5]) and personality disorders (see [6]).

1. Self and Resting State Activity

Neuronal Findings Ia: Anatomical Rings, Midline Regions and Self-Specificity

Neuroscience in general and functional brain imaging in particular detected the self in the last decade. In the last 10 years several imaging studies on the brain's recruitment of neural activity during personally or self-specific stimuli were detected. Subjects viewed and judged words (or other stimuli like pictures or sounds) that were closely related to themselves as, for instance, the term "piano" for concert pianists. These self-specific words were then compared to other words unrelated and thus non-self-specific to the person (see [7-11] for reviews)

Interestingly, most of these studies observed strong activity in the anterior and posterior cortical midline regions like the perigenual anterior cingulate cortex (PACC), the supragenual anterior cingulate cortex (SACC), the ventro- and dorsomedial prefrontal cortex (VMPFC, DMPFC), the posterior cingulate cortex (PCC), the precuneus, and the retrosplenium (see [9-12] for reviews and meta-

analyses). These findings have led to the question of whether the cortical midline structures specifically process the stimuli's degree of self-specificity and can thus be considered a network specific to the self.

However, the assumption of self-specificity of these regions has been put into doubt because tasks and stimuli other than those focusing on the self also recruit these regions [7, 8, 12]. This includes various cognitive functions like mindreading and decision making, and social functions like empathy, reward, and emotional-affective functions.

Another observation is that self-specificity does not only recruit cortical midline structures. If self-specific stimuli are presented independent of any associated cognitive tasks (like judgment), they also elicit neural activity changes in subcortical midline regions like the dorsomedial thalamus, the ventral striatum, the tectum, the periaqueductal gray, and the colliculi (see [13-16]).

Is there thus a subcortical-cortical midline system mediating the self? There is indeed some neuroanatomimcal ground justifying the assumption of such subcortical-cortical midline system (see though the critical stance of [7, 8, 12], who consider the midline regions to be too unspecific while they presuppose a different more cognitive notion of the self as presupposed here). This is well documented in the threefold radial-concentric anatomical organization, with inner, middle, and outer rings, that spans from the subcortical to cortical regions (see [1, 2, 17] as well as [1, 2]).

On the subcortical level the three rings group around the third ventricle. Most adjacent to it is the inner ring, the median part that extends to the aforementioned regions on the cortical level. This is followed by the paramedian and lateral core as middle ring and finally by the most outer lateral part, with both finding their continuation onto the cortical level. The threefold anatomical organization thus suggests integrated subcortical-cortical systems.

Simply put, the inner ring includes all cortical and subcortical regions directly adjacent to the ventricles like the anterior and posterior cingulate on the cortical level, while the outer ring contains all regions that are visible on the outer surface of the brain like the lateral prefrontal and parietal cortex and the sensory and motor cortices. The middle ring is sandwiched in between and contains on the cortical level the VMPFC, the DMPFC, and the precuneus.

Neuronal Findings Ib: Anatomical Rings and Self-Specificity

How is such threefold subcortical-cortical neuroanatomical distinction related to the self? Does it map self-specificity better and more congruently than the traditional dichotomous medial-lateral distinction?

For answers to these questions, neurologist Todd Feinberg relies, much like traditional neurologists such as Charles Sherrington, on the careful study of neurological patients. These patients suffer from lesions in particular regions of the brain and often experience bizarre changes in their phenomenal consciousness, including their sense of self. Feinberg attempts to explain his patients' unusual experience of the self with the concept of the threefold anatomical organization (see [18-20]).

Feinberg [18, 20]now assumes that the inner ring reflects the bodily or "intero-self," while the outer ring may be rather related to the environmental self or the "extero-self." The middle ring is more related to the integration between both and thus *the self proper*, the "integrative self," than spans across intero- and exteroceptive stimuli and thus body and environment.

Can we gain any empirical support in favor of the different rings' association with self-specificity? Pengmin Qin from my group [10, 21]conducted a meta-analysis of all imaging studies on the self using self-specific, familiar, and nonfamiliar stimuli. The self-specific condition yielded activity changes in the inner ring regions like the PACC, the insula, and the PCC as well as in the middle ring, that is, VMPFC and DMPFC.

The familiarity condition, in contrast, did not yield any signal changes in the inner ring and its anterior regions, that is, insula and PACC. The familiarity condition did induce signal changes in the middle ring, VMPFC and DMPFC, as well as in the posterior regions of the inner ring, the PCC. Finally, the posterior regions of the inner ring like the PCC were also recruited during the nonfamiliar condition. Unlike familiarity, the non-familiar condition did not recruit any other midline regions in the anterior parts but rather the temporo-parietal junction and the temporal pole.

Taken together, these findings suggest an inner-to-outer gradient coupled with an anterior-to-posterior gradient in the brain with regard to self-specificity. The more inner and anterior regions like the PACC and insula allow for the neural processing of high degrees of self-specificity, while the converse holds for the more outer (like lateral regions) and posterior regions (like PCC and precuneus) that are more associated with low degrees of self-specificity. Simply put, the own self is associated with anterior and inner parts in the brain, while others' selves are more related to posterior and outer parts.

We have to be careful, however. The aforementioned findings described concern mainly on the cortical level, while there are much less data about the subcortical regions. Hence, future investigation is needed to reveal whether, for instance, the inner-to-outer gradient also applies to the subcortical level.

One step in this direction has been taken by Hans Lou from Denmark. Having investigated the self in a series of excellent studies (see [22-27]), he assumes a subcortical-cortical paralimbic network to be central in mediating self and consciousness. This includes the thalamus (pulvinar), the subgenual and pregenual anterior cingulate cortex, the medial prefrontal cortex (VMPFC, DMPFC), the striatum, and the posterior cingulate cortex/precuneus. That network can be characterized as a self-reference network that as such may balance other networks associated with reward, emotion, and executive-cognitive functions. As such, Lou considers it central for consciousness (see also the later discussion of this network in further detail).

Moreover, we have to distinguish between the self-specificity of internal and external contents and their recruitment of midline regions. Internal contents concern the own thoughts and the own body, whereas external contents are related to objects and events in the environment. Future studies are needed to show, as I assume, that both internal and external contents with high degrees of self-specificity are processed in the subcortical-cortical midline regions.

Neuronal Findings IIa: Rest and Self-Specificity—Anterior Midline Regions

The cortical midline structures are core regions of the default-mode network that shows particularly high neural activity in the resting state (see[28, 29]). Since the midline regions have been shown to be implicated in mediating self-specificity, neural activity during self-specificity may strongly overlap with the high resting-state activity in the very same regions. This is indeed the case, as several studies could show.

D'Argembeau [30] conducted an H20 positron emission tomography (PET) investigation. Subjects underwent four conditions, thinking/reflection about one's own personality traits, thinking/reflection about another person's personality traits, thinking/reflection on social issues, and a pure rest condition where subjects could relax. This allowed them to compare self- and non-self conditions as well as to investigate the relation between self-conditions and the resting state.

What about their results? The VMPFC showed significant increases in regional cerebral blood flow (rCBF) during the self condition when compared to the other and the social condition. In addition, they compared all three task-related conditions, that is, thinking/reflection about the own personality traits,

thinking/reflection about another person's personality traits, and thinking/reflection on social issues, against a rest condition. This yielded increased rCBF in the DMPFC and the temporal regions, while no differences were observed in the VMPFC. Conversely, the rest condition showed rCBF increases in a large medial fronto-parietal and posterior medial network with no differences in the VMPFC.

The separate account of self and rest allowed the authors to directly compare both conditions with each other. This yielded strong overlap in the VMPFC between both conditions, rest and self-specificity that showed similar rCBF increases. In contrast, other and social conditions induced rCBF decreases in the same region.

Postscanning subjective measures demonstrated that self-referential thoughts were most abundant in the self condition while being more diminished in the other three conditions. The authors therefore correlated the postscanning measures of self-referential thinking with the rCBF changes. This yielded a positive relationship in the VMPFC. The higher the rCBF in the VMPFC, the higher the degree of self-referentiality in the thoughts subjects reported themselves (across all four conditions as described earlier).

This strong association, that is, regional overlap, between self and rest in especially the regions of the inner ring was further confirmed by recent meta-analyses by Pengmin Qin from my group [10, 31] as I will describe in the following.

Pengmin Qin conducted a meta-analysis of human imaging studies on the self when compared to *nonself*. Most important, he also included studies on the resting state to compare their neural activity pattern to the one during self- and non-self-specific stimuli. This allowed him to, for instance, directly compare resting state in the default-mode network with the regions recruited during self- and non-self-specific stimuli.

What are his results? They confirm the ones described earlier showing regional overlap between self and rest. More specifically, the regional activities during self-specific stimuli and the ones during resting state overlapped especially in the PACC extending to the VMPFC, while no such regional overlap with the resting state was observed in the other conditions, that is, familiarity and nonfamiliarity, in either the PACC or any other region.

Taken together, these results suggest neural overlap between self-specificity and resting-state activity in the anterior midline regions like the VMPFC.

Neuronal Findings IIb: Rest and Self-Specificity—Anterior and Posterior Midline Regions

The strong overlap between resting state and regions recruited during self-specificity was further confirmed in a recent study [31]. They conducted two experiments with different subject groups. Each group underwent a self-reference task, explicit judgment of trait adjectives as self- or non-self-related, and a control task with a valence judgment of the trait adjectives as either positive or negative, and a pure resting state (of about 10 s). This allowed them to compare all three conditions/tasks (self, valence, rest) with each other and to see, in particular, how much self and rest overlap with each other.

As expected, they found in both experiments recruitment of stronger neural activity in anterior and posterior midline regions (VMPFC, DMPFC, PACC, PCC, precuneus) during the self-task when compared to the valence task. Moreover, the rest condition was associated with stronger activity than the valence task in the midline regions, while this was not the case for the self-task when compared to rest. Here the level of neural activity in the midline regions did not differ between the self-task and the resting-state condition.

The authors also conducted analyses that allowed them to directly investigate overlapping and dissociating regions between self and rest. Overlapping regions (i.e., conjunction analysis) between self and rest concerned the PACC, the VMPFC, and the PCC, while dissociating regions included the DMPFC (stronger during self) and the precuneus (stronger during rest). These findings could be confirmed in the second experiment where the relevant regions, as yielded in the first experiment, showed the same pattern of neural activity during the three conditions, self, non-self, and rest.

These findings suggest a close relationship, that is, regional overlap, between the neural activities underlying rest and self in especially the regions of the inner ring, the PACC and the PCC, while in the middle ring (precuneus, DMPFC), neural activities during self and rest seem to dissociate from each other.

The overlap between rest and self is further support by a recent magnetoenceophalographic (MG) study by Lou et al [22]. He investigated judgment of self-related words and focused on three main regions, precuneues, thalamus/pulvinar, and anterior midline regions (including VMPFC, DMPFC, and PACC). Using granger causality analysis (which allows testing of the direction of functional connectivity), he observed that the three regions were bidirectionally connected to each other (i.e., showing high degrees of statistical covariance in their signal changes).

Most interestingly, the increase in functional connectivity occurred already 900 ms before stimulus onset, while it was further enhanced by the stimulus itself for the first 900 ms. Such functional connectivity was strongest in the gamma range between 30 and 45 Hz before and after stimulus onset and strongest in the self-condition after stimulus onset.

Taken together, these studies show strong overlap between high resting-state activity and stimulus-induced activity as elicited by high self-specific stimuli in anterior and posterior midline regions, including both cortical and subcortical (albeit with limited evidence) regions. Hence, the resting-state activity in the anterior regions of the inner ring seems to be closely related to self-specificity in as yet unclear ways. What does this overlap between resting state activity and self-specificity imply for the kind of psychopathological symptoms we observe in schizophrenia and depression? This will be the focus in the subsequent sections.

II. Self in Depression

Psychopathological Symptoms: 'Increased Self-Focus' and 'Decreased Environment-Focus' in Depression

Major depressive disorder (**MDD**) is a psychiatric disorder that is characterized by extremely negative emotions, suicidal thoughts, hopelessness, diffuse bodily symptoms, lack of pleasure, that is, anhedonia, ruminations, and enhanced stress sensitivity (see [32]and [3] as well as [17]for a recent overview and [33]).

How do depressed patient experience their symptoms? This is dealt with in phenomenology that, loosely (and rather broadly) defined (see: Northoff 2004 for details), describes the subjective experience of the depressive symptoms from the First-Person. Such phenomenological approach has to be distinguished from the psychopathological approach that targets the depressive symptoms in the rather objective terms of observation in Third-Person Perspective. Since the three above described psychodynamic key features clearly point to an altered self in depression, the following phenomenological and psychopathological account focuses on the self. For that purpose, I want to start with a quote from a recent paper that nicely describes the alterations of the self in depression.

 "She sat by the window, looking inward rather than looking out. Her thoughts were consumed with her sadness. She viewed here life as a broken one, and yet she could not place here finger on the exact moment it fell apart. 'How did I get to feel this way?' she repeatedly asked herself. By asking, she hoped to transcend her depressed state; through understanding,

she hoped to repair it. Instead, her questions led her deeper and deeper inside herself –further away from the path that would lead to her recovery." [34].

This description of a depressed patient shows three crucial characteristics concerning the self which shall be conceptualized as increased self-focus, association of the self with negative emotions, and increased cognitive processing of the own self (see [35]).

Let me start with the increased self-focus. Similar to our patient, almost all depressed patients look rather inward than outward, they focus very much on themselves while no longer being able to shift their focus on others' (see arrows towards the self in Figure 1). Social-psychological theory speaks of self-focused attention as focus on internal perceptual events, that is, information from those sensory perceptions that react to changes in bodily activity [36]. The Self-focus may also concern enhanced awareness of one's present or past physical behaviour, that is, heightened cognizance of what one is doing or what one is like. In addition to such increased self-focus, the depressed patient's focus is often also on the own body. Depressed patients show heightened awareness of their own body which phenomenologically results in the subjective perception of diffuse bodily symptoms (see [37]). The increased self-focus may consequently be accompanied by what I call 'increased body-focus'.

The increased self- and body-focus imply that the depressed subject's attention is no longer focused on its relation to the environment and environmental events, as in healthy subjects, but rather on itself as the prime focus with the environment shifting into the background. Hence the increased self-focus goes along with what one may call 'decreased environment-focus'. The concept of the decreased environment-focus describes that patient's subjective experience and perception is no longer directed towards the environment and its respective persons and events. Instead, the patient's subjective perception and experience is rather directed the own body and the own cognition thus resulting in what I here describe as 'increased self-focus'. This means that the balance between the environment-focus and the self-focus is unilaterally shifted to the latter in depression at the expense of the former. Therefore, the decreased environment-focus entails the increased self-focus (and vice versa) (see Figure 1).

This also supported by recent empirical data. Empirical research clearly indicates heightened self-focused attention in depression. A variety of studies assessing self-focused attention with diverse measures and methodologies all converge on the finding of an increased and perhaps prolonged level of self-focused attention in depression [36]. What remains unclear though is whether this increased self-focus

Figure 1
Self, body and environment in depression
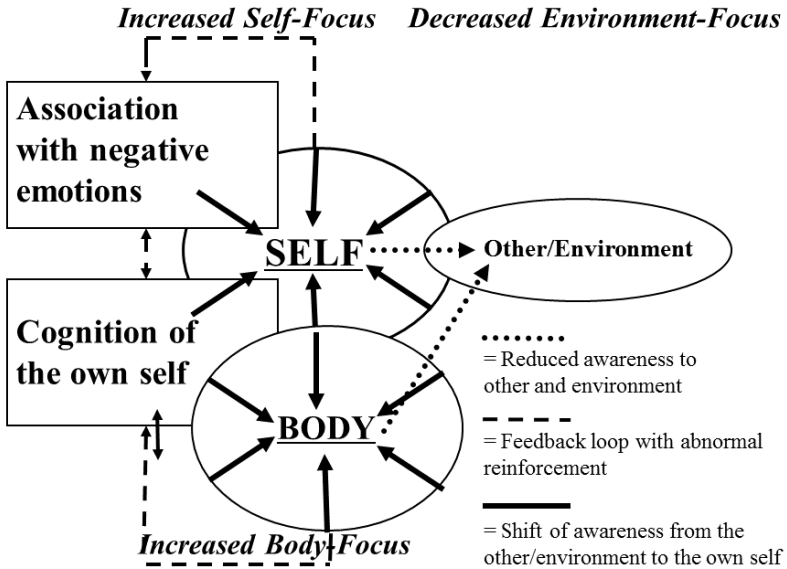
The figure shows the relationship between the different directions of phenomenal consciousness (here denoted as awareness) in depression. Phenomenal consciousness can be directed either externally towards the environment or internally towards either the own self or the body. In depression, there is increased directedness towards the own self and the body ('increased self- and body-focus'), while the directedness towards the environment is decreased ('decreased environment-focus'). The increased self- and body-focus are symbolized by larger circles and inward arrows; the decreased relationship of both self and body to the environment is illustrated by thinned arrows. The consequences of the increased self-focus for subsequent psychological functions are indicated on the very left leading to increased association with negative emotions and increased cognitions of the own self.

is purely explicit and thus conscious or whether it is already present on an implicit and thus unconscious level.

Another characteristic is the attribution of negative emotions to the own self, the association of the own self with negative emotions. The own self is associated with abnormal sadness, guilt, mistakes, inabilities, death, illness, etc. which may ultimately result in paranoid delusions. A recent study investigating symptom clusters in the Beck Depression Inventory (BDI) observed three BDI factors among them the self-blame factor [38]. Interestingly, depressed patients with previous suicide attempts showed significantly higher scores on the BDI self-blame factor than those without suicide attempts. Moreover, the self-blame factor significantly correlated with the total number of suicide attempts and with known risk factors for suicidal behaviour (See [35] for review). Such self-blame possibly results from the association of the self with predominantly negative emotions in depression while at the same time these patients remain apparently unable to attribute any kind of positive emotions to their self.

Finally, there is also increased cognitive processing of the own self. Our patient described above typically suffers from increased cognitive processing, she thinks about herself and her mood and tries desperately to discover the reasons for her depression and thereby gets only deeper and deeper into the depressed mood (see arrows from negative cognitions to the self in Figure 1). This cognitive processing of the own self is described as *rumination* and is often considered a method of coping with negative mood that involves increased self-focused attention and self-reflection [36].

On the basis of a rumination scale, Treynor et al. [34] suggest a two factor model of rumination. They call the first factor reflection, which describes a purposeful turning inward to engage in cognitive problem solving to alleviate one's depressive problems. The second factor is best described as brooding as passive comparison of one's current situation with some unachieved standard oriented on others' and thus on what psychodynamically can be called the ideal-self [39]. Corresponding to the first factor, reflection, Rimes and Watkins [40] suggest an increase in what they call 'analytical self-focus' in depression which they define as thinking analytically 'about' oneself and one's symptoms; such increased 'analytical self-focus' is related to the increased ratings of the self as worthless and incompetent and ratings of depressed mood.

They distinguish the increased 'analytical self-focus' in depression from what they call 'experiential self-focus'. The concept of the 'experiential self-focus' is supposed to describe the focus on the direct experience of one's thoughts, feelings and sensations in the present moment. In contrast to the increased 'analytical self-focus',

the 'experiential self-focus' tends to be rather low and decreased in depression since these patients do no longer experience themselves and their self as such. There is thus an imbalance between 'analytical and experiential self-focus' in depression. Whether this imbalance between 'analytical and experiential self-focus' corresponds to the above postulated dysbalance between an 'increased self-focus' and 'decreased environment-focus' must remain open here requiring future conceptual and empirical investigation.

Neuronal Findings Ia: Resting State and its Anatomical Organisation

Let me first describe the findings of altered resting-state activity in MDD. Because of several excellent reviews about the structural and functional brain changes in MDD ([41-46] we here briefly highlight only the main findings and conclusions from these various reviews and then relate them to functional networks as delineated in normal-healthy brains (see [3, 32] for recent reviews).

Alcaro et al (2010) conducted a meta-analysis of all imaging studies in human MDD that had focused on resting-state activity. This yielded hyperactive regions in the perigenual anterior cingulate cortex (PACC), the ventromedial prefrontal cortex (VMPFC), thalamic regions like the dorsomedial thalamus and the pulvinar, pallidum/putamen and midbrain regions like the ventral tegmental area, substantia nigra, the tectum, and the periaqueductal gray (PAG). In contrast, resting-state activity was hypoactive and thus reduced in the dorsolateral prefrontal cortex (DLPFC), the PCC, and adjacent precuneus/cuneus [41].

These results are well in accord with other meta-analyses (see [46-49]) emphasized the role of the hippocampus, parahippocampus, and the amygdala where resting-state hyperactivity was also evident in MDD. Interestingly, the very same regions and the PACC also show structural abnormalities with reduced gray matter volume in imaging studies and reduced cell count markers of cellular function in postmortem studies (see [46, 48, 49]).

Involvement of these regions in MDD is further corroborated by the investigation of resting-state activity in animal models of MDD. Reviewing evidence for resting-state hyperactivity in various animal models yielded diverse participating brain regions—the anterior cingulate cortex, the central and basolateral nuclei of the amygdala, the bed nucleus of the stria terminalis, the dorsal raphe, the habenula, the hippocampus, the hypothalamus, the nucleus accumbens, the PAG, the DMT, the nucleus of the solitary tract, and the piriform and prelimbic cortex [41]. In contrast, evidence of hypoactive resting-state activity in animal models remains sparse with no clear results [41].

Taken together, these findings indicate abnormally high resting-state activity in extended subcortical and cortical medial regions of the brain. This has led authors like Phillips [50], Mayberg [42-44], and Drevets (see [46, 48, 49]) to assume dysfunction in the limbic system in depression or more specifically in the "limbic-cortico-striato-pallido-thalamic circuit" with reciprocal interactions between medial prefrontal and limbic regions being crucial [46].

Neuronal Findings Ib: Resting State Imbalance between Inner and Outer Ring

How do these findings fit into the delineated anatomical characterization of the healthy brain as characterized by inner, middle, and outer rings? What was conceptualized as inner and middle rings at the cortical level, the paralimbic areas and the cortical midline structures, generally show hyperactivity during resting state in MDD.

Another observation fits well with this anatomical model in the healthy brain. The outer ring covers the lateral regions on the cortical level like the DLPFC and the sensory and motor regions. Especially in the DLPFC and in part also in the motor cortex (see [41]), resting-state hypoactivity has been consistently observed, especially in MDD.

Considering these findings together, resting-state activity in MDD may be characterized by a subcortical-cortical imbalance between inner/middle and lateral rings. More specifically, the inner and middle rings' regions seem to be hyperactive in the resting state. In contrast, subcortical and especially cortical regions of the lateral-cognitive ring, like the lateral prefrontal cortex and the sensory-motor cortices, seem to show hypoactivity in the resting state (see also [17]).

Neuropsychopathological Hypothesis **I: Resting-State Hyperactivity and Increased Self-Focus in Depression**

The aforementioned findings indicate imbalance in the resting-state activity between inner/middle and outer rings. More specifically, they show that resting-state activity in the anterior portions of the inner ring and also to some degree in the middle ring is abnormally elevated, while the outer ring's resting-state activity is decreased. This means that there is a dysbalance in the resting-state activity in depression along the aforementioned inner-to-outer and anterior-to-posterior gradients.

I now assume the balance between the three anatomical rings and their corresponding functional connectivity to be central in constituting the balance between self-specific and pre-intentional organization in the resting state. Why? We recall the neural balance between midline and lateral networks to be central in

designating contents as internal or external as well as in constituting the directedness toward contents. Since the neuronal balance between the three anatomical rings is altered in depression, one would expect a shift toward increased internal contents and decreased external contents. This indeed is the case and surfaces on the phenomenal level, as I describe in the following.

Phenomenally, a core symptom in MDD is the extremely increased focus on the own self. All thoughts and feelings are circulating around the own self, the own person, which we described as increased self-focus (see Northoff [3, 35]; see also Lemogne [51] , who distinguishes such increased self-focus as associated with phasic VMPFC hyperactivity from its phasic DMPFC-mediated cognitive elaboration). Such increased self-focus goes along with detachment from the environment; that is, from the persons, objects, and events with the patients feeling disconnected. We described this as "decreased environment focus" (see [17]). The question is now how the shift in the focus from the environment to the self, that is, decreased environment focus and increased self-focus, are generated. For that, we turn to the resting-state activity in MDD.

One would consequently expect elevated resting-state activity in the midline regions to lead to increased self-specificity and hence to abnormally increased personal concerns in patients with MDD during both resting-state and stimulus-induced activity. While it remains to be demonstrated for the resting-state activity, it holds true indeed for stimulus-induced activity. Grimm ([52, 53]) from our group (and others like Lemogne [51, 54-56] who also distinguish between phasic and tonic activity) observed behaviorally significantly increased scores for self-specificity with regard to especially negative emotional pictures. Neuronally this went along with decreased signal changes during self-specific stimuli in anterior cortical midline regions. This supposedly reflects the abnormally high resting-state activity and its increased assignment of self-specificity to stimuli.

The assumption of increased self-specificity on the phenomenal level is further supported by the observation of a correlation between the increased behavioral scores of self-specificity and the decreased stimulus-induced activity in especially the anterior midline structures. One may consequently hypothesize that the increased self-specificity as observed behaviorally stems from the abnormally increased resting-state activity in the midline regions and their apparently increased self-specific processing [52, 53].

What do these findings imply in neurophenomenal regard? We observed decreased stimulus-induced activity in the anterior midline regions, while at the same time the stimuli were assigned increased degrees of self-specificity. How is it possible that decreased stimulus-induced activity goes along with increased self-specificity?

I assume that this is due to the carryover and transfer of the increased resting-state activity and its abnormal self-specific organization onto subsequent stimulus-induced activity.

Let me be more specific. The increased resting-state activity makes it impossible for the stimulus to induce major activity changes, hence the decreased stimulus-induced activity. And the associated carryover and transfer of the resting state's abnormal self-specific organization is increased and consecutively leads to the assignment of abnormally high degrees of self-specificity to the stimuli, hence the conjunction of neuronal decreases in stimulus-induced activity and behavioral increases in self-specificity.

Neuronal Findings II: Abnormal Exteroceptive Processing in Depression

How about the resting state's preintentional organization in depression? For that, we need to consider how interoceptive and especially exteroceptive stimuli are processed in MDD.

Patients with MDD often suffer from generalized bodily symptoms like heart pounding, increased breathing (with yawning), and multiple-diffuse bodily aches. This seems to go along with abnormally increased awareness of their own bodily processes (body perception), including sensitivity to stress and autonomic-vegetative changes as demonstrated in a recent work [37].

The same study also investigated the neuronal activity during exteroceptive and interoceptive awareness (tone and heartbeat counting) in relation to the brain's resting-state activity. Interoceptive stimuli by themselves (e.g., the heartbeat) induced a "normal" degree of brain signal changes (activation) in the bilateral anterior insula in depressed patients when considered relative to the preceding resting-state activity levels. This suggests that there is no abnormality in interoceptive stimulus processing itself in depression.

In contrast to stimulus-induced activity during interoceptive stimuli, we observed abnormally reduced activity during exteroceptive stimuli. More specifically, we observed that exteroceptive stimuli induced decreased stimulus-induced activity in the insula in depressed patients when compared to healthy subjects. This let us further question whether such reduced activity is related either to the exteroceptive stimulus itself or rather to abnormal resting-state activity levels. The latter was indeed the case, as we observed increased resting-state activity in the insula itself. This is well in line with the resting-state hyperactivity in the inner ring, the core-paralimbic system to which the insula belongs.

To test for independent changes in exteroceptively related stimulus-induced activity, we then calculated the exteroceptively related stimulus-induced activity relative to the preceding resting-state activity level. Interestingly, the initially observed difference between healthy and depressed patients in "absolute," for example, resting-state-independent, signal changes during exteroceptive stimuli when calculating them in such "relative" way, dependent on the preceding resting-state activity level. Hence, when including the resting-state activity level, there was no difference anymore between healthy and depressed subjects in signal changes during exteroceptive processing.

In contrast to the exteroceptive stimuli, no differences between healthy and depressed subjects were evident in interoceptive stimuli in both relative and absolute signal changes. This difference between interoceptive and exteroceptive stimuli with regard to relative and absolute signal changes suggests differential interaction of both kinds of stimuli with resting-state activity. Either rest–stimulus interaction is reduced during exteroceptive stimuli or rest–stimulus interaction is increased during interoceptive stimuli, which cannot be differentiated on the basis of our findings.

What is clear is that there is imbalanced activity between intero- and exteroceptive stimulus processing, including their respective interaction with the resting-state activity level. Because of the paucity of work in this area, additional imaging studies need to investigate changes in interoceptive processing in depression.

The study by Wiebking [37] also investigated psychological measures of body perception, employing the body perception questionnaire (BPQ). They found the BPQ scores to be significantly increased in depressed patients as being indicative of increased bodily awareness. Most interestingly, unlike in healthy subjects, the increased BPQ scores no longer correlated with the signal changes during the resting-state and the exteroceptive condition.

This suggests that depressed patients no longer properly modulate their degree of neuronal activity. They remain apparently unable to properly down-modulate the perception and awareness of their own body and to shift attention from the body to the environment. This may explain the many somatic complaints that characterize MDD. Though tentative, such lack of correlations with abnormally increased neuronal activity has also been seen for other psychological measures in depression like excessive negative affect, self-specificity, and negativistic temporal projections to future possibilities [37, 51, 53, 56].

Neuropsychopathological Hypothesis II: "Self-Perspectival–Intentional Imbalance" in Depression

These findings are indicative of an imbalance in the neural processing between interoceptive and exteroceptive stimuli, with only the latter but not the former inducing decreased neural activity. This may consecutively lead to relatively increased neural processing of interoceptive processing and rest–intero interaction when compared to the apparently absolutely reduced exteroceptive processing and rest–extero interaction. As already noted, this abnormal shift toward interoceptive processing may psychopathologically promote increased bodily awareness and subsequent concerns with undesired bodily symptoms.

Meanwhile, the decreased exteroceptive processing may be accompanied by reduced awareness of and concern with environmental changes, especially positive events that could beneficially impact depression (see Fig. 27-2b). This means that, phenomenally, one may want to speak not only of an increased self-focus but also of an increased bodily focus and a decreased environment focus (see also: Northoff [17]).

Why though are rest–extero interaction and external awareness reduced when compared to rest–intero interaction and internal (i.e., bodily and self-) awareness? Recall that I assumed the increased self-focus and increased self-specificity during stimulus-induced activity to be traced back to the increased resting-state activity and the anterior regions of the inner ring, the midline network. At the same time, however, the resting state data also showed decreased resting-state activity in the lateral regions of the outer ring (see earlier). The question now is how such decreased resting-state activity in the outer ring is manifest pre-phenomenally as well as phenomenally in stimulus-induced activity.

We already know that stimulus-induced activity is reduced during exteroceptive stimuli as supposedly mediated by decreased rest–extero interaction. That, in turn, may phenomenally go along with decreased external awareness, which phenomenally I described as decreased environment focus. Now the question arises how such decreased stimulus-induced activity in lateral regions and the decreased environment focus are related to the abnormally low resting-state activity in the very same regions.

One can assume a preintentional organization of the resting state that we supposed to be associated with the neural balance between midline and lateral networks in the resting state. If now this neural balance is abnormally shifted toward the midline regions, as it seems to be the case in depression, the associated pre-phenomenal balance will also shift. More specifically, the resting state's pre-phenomenal balance will shift toward its self-specific organization and away from its pre-intentional

organization. That means that any subsequent phenomenal consciousness is predisposed and biased toward increased self-specificity while at the same time going along with decreased external directedness of their intentionality.

That is exactly what one observes in depression. The resting state's abnormally strong self-specific organization is manifest in increased self-specificity and increased directedness toward internal contents at the expense of external contents. This is reflected in the description of an increased self-focus and a decreased environment focus (see earlier). That is possible only if the resting state's pre-intentional organization is dysbalanced toward internal contents, which I assume to correspond to the abnormal neural balance between midline and lateral networks; that is, between inner/middle and outer rings.

III. Self in Schizophrenia

Psychopathological Symptoms: Abnormalities of the Self in Schizophrenia

Early psychiatrists like E. Kraepelin and E. Bleuler at the beginning of the 20th century assumed abnormality of the self to be basic in schizophrenia. Unlike in our times, these early psychiatrists had to rely on nothing but clinical observation. Based on that, they assumed an abnormal change of the self to be fundamental in schizophrenia.

More specifically, Kraepelin [57] characterized schizophrenia as "the peculiar destruction of the inner coherence of the personality" with a "disunity of consciousness" ("*orchestra without conductor*"). Bleuler [58] also pointed out that schizophrenia is a "disorder of the personality by splitting, dissociation" where the "I is never completely intact."

A contemporary of Bleuler and Kraepelin, Berze, even referred to schizophrenia as "basic alteration of self-consciousness." Karl Jaspers [59] also noticed "incoherence, dissociation, fragmenting of consciousness, intra-psychic ataxia, weakness of apperception, insufficiency of psychic activity and disturbance of association, etc." to be basic as unifying "central factors" in schizophrenia.

The early descriptions of a disrupted self are complemented by current phenomenological accounts that focus predominantly on the experience of the own self in relation to the world. Josef Parnas [60, 61] describes what he calls "presence" as being altered in schizophrenia. The experience of the world and its objects is not accompanied by a pre-reflective self-awareness anymore.

Let me specify this point. The own self, the self that experiences the experience of the world, is no longer included in that very experience:

"The prominent feature of altered presence in the pre-onset stages of schizophrenia is disturbed ipseity, a disturbance in which the sense of self no longer saturates the experience. For instance, the sense of mineness of experience may become subtly affected: one of our patients reported that this feeling of his experience as his own experience only "appeared a split-second delayed." " [60]

The patients remain unable to refer to themselves in their experience of the world. It is as if the experience of the world is no longer their own experience of their own self. Instead, their experience may belong to and be experienced by someone else, but it is no longer their own self who makes and experiences those experiences. Due to the absence (as opposed to presence) of the own self in their experience of the world, patients with schizophrenia become detached, alienated, and estranged from their own experience. Such detachment of the experiences from their own self makes it impossible for them to experience their experiences as subjective and thus as belonging to their own self.

The experiencing self is consequently no longer affected by its own experiences, which Sass (2003) describes as "disorder of self-affectivity": the own self is no longer experienced as the own self and most important is no longer experienced as the vital centre and source of the own experiences, actions, perceptions, thoughts, and so on. This reflects what Sass [62] calls the "diminished self-affection," meaning that the self is no longer affected by its own experiences.

If, however, the self is not affected anymore by its own experience, the self stands apart from the objects and the events in the world that are experienced. A gulf, a phenomenological distance as Parnas [60] says, opens up between world and self. The objects and events of the world no longer make intuitive sense and are thus not meaningful anymore to the experiencing subject. The own self becomes thus almost objective and mechanical in its experience and perception of the world.

Neuronal Findings Ia: Resting State in Schizophrenia

Various studies investigated recently the default-mode network (DMN) in schizophrenia (see: Kuhn and Gallinat [33] for a recent review). Recent imaging studies in schizophrenia reported abnormal resting-state activity and functional connectivity in the *anterior cortical midline structures* (aCMS). One study [31] demonstrated that the aCMS (and posterior CMS like the posterior cingulated cortex [PCC]/precuneus) show decreased task-induced deactivation (TID) during a working memory task. This was observed in both schizophrenic patients and their

relatives when compared to healthy subjects. That is indicative of decreased task-related suppression and possibly increased resting-state activity.

Furthermore, the very same schizophrenic subjects also showed increased functional connectivity of the aCMS with other posterior regions of the CMS, such as the PCC. Both functional hyperconnectivity and decreased TID correlated negatively with each other. The more decreased task-related suppression, the more increased the degree of functional connectivity. Finally, both decreased TID and increased functional connectivity in aCMS correlated with psychopathology, that is, the predominantly positive symptoms as measured with the PANS scale.

Decreased TID in aCMS was also observed in an earlier study that investigated working memory [63]. Similar to the study described earlier, they let subjects perform a working memory task and observed abnormally decreased TID in aCMS in schizophrenic patients when compared to healthy subjects. And similar to the other study, they also observed abnormal task-related activation in the right dorsolateral prefrontal cortex in schizophrenic patients. Another study [64] also reported abnormal TID in aCMS as well as abnormal functional connectivity from aCMS and posterior CMS to the insula in schizophrenic patients (see also [65-68] .In addition to TID and functional connectivity, another abnormal measure of resting-state activity is the temporal features, more specifically fluctuations or oscillations in certain temporal frequencies. For instance, Hoptman [69] demonstrated that low-frequency fluctuations in the resting state were increased in the aCMS (and the parahippocampal gyrus) in schizophrenic patients, while they were decreased in other regions like the insula. Abnormally increased low-frequency oscillations (<0.06 Hz) in the aCMS (and posterior CMS regions and the auditory network) and their correlation with positive symptom severity were also observed in another study on schizophrenic patients [70].

Neuronal Findings **Ib: Self-Specificity in Schizophrenia**

This concerns alterations in the resting-state activity. How about changes during stimulus-induced activity and their relation to self-specificity?

A recent imaging study by Holt [71] showed that abnormal anterior-to-posterior midline connectivity is related to self-specificity. They investigated schizophrenic patients during a word task where subjects had to judge trait adjectives according to their degree of self-specificity (and also two other tasks: other-reflection, i.e., relation of that word to another person) and perception-reflection (i.e., word printed in uppercase or lowercase letters).

How about their results? Schizophrenic patients showed significantly elevated activity in posterior midline regions like the mid- and posterior cingulate cortex

during self-reflection, while signal changes in the anterior midline regions like the medial prefrontal cortex were significantly reduced when compared to healthy subjects. Finally, functional connectivity was abnormally elevated from the posterior to the anterior midline regions in schizophrenic patients. Analogous results of altered midline activity with a dysbalance between anterior and posterior midline regions are also observed in other studies on self-specificity in schizophrenia (see [72])

Taken together, these results demonstrate abnormal resting-state activity in especially the anterior and posterior midline network in schizophrenia (see [33] for a recent meta-analysis). The very same network also shows alterations in the balance between anterior and posterior midline regions when probing for self-specific stimuli.

Unfortunately, studies testing the linkage between resting-state abnormalities and self-specific stimuli remain to be conducted in schizophrenia. This would be needed to support the argument of the carryover and transfer of the resting state's self-specific organization onto subsequent stimulus-induced activity and the associated phenomenal state of consciousness.

While we currently do not have such neuronal evidence, there is plenty of phenomenal evidence of schizophrenic patients suffering from an abnormality in self-specificity in their experience. This makes it rather likely that abnormal self-specific organization in the resting state may be carried over and transferred onto the subsequent stimulus-induced activity and its associated phenomenal states.

Neuropsychopathological Hypothesis Ia: "Basic Disturbance of the Self" in Schizophrenia

How do these findings relate to the psychopathological and phenomenological descriptions? I assume that what the early psychiatrists described as "the peculiar destruction of the inner coherence of the personality" or "basic alteration of self-consciousness" may correspond to what here refer to the changes in the resting state's self-specific organization. Let me spell this out in further detail.

Following the early descriptions, the basic disturbance in the self is supposed to impact all other subsequent functions and domains of the personality. Analogously, I assume the resting state's self-specific organization to also affect any subsequent stimulus-induced activity and consecutively all functions, including sensory, motor, affective, and cognitive functions (as it seems indeed to be the case in cognitive, affective, sensory, and motor functions). In the same way the basic disturbance of the self is present everywhere, the resting state, metaphorically speaking, "has its hands" in all kinds of neural processing.

How is such overall presence of the "basic disturbance of the self" possible? It must indeed be very basic. As basic as, for instance, the resting state is basic to any kind of subsequent stimulus-induced activity. And it must be carried over and transferred to the subsequent stimulus and its associated contents and functions. I thus assume that what the early psychiatrists described as "basic disturbance of the self" is carried over and transferred to every domain of the subject's mental life. In the very same way the resting state's abnormal self-specific organization is carried over and transferred to the subsequent stimulus-induced activity and its associated functions (see Figure 2).

Figure 2 Basic Disturbance of the Self in Schizophrenia
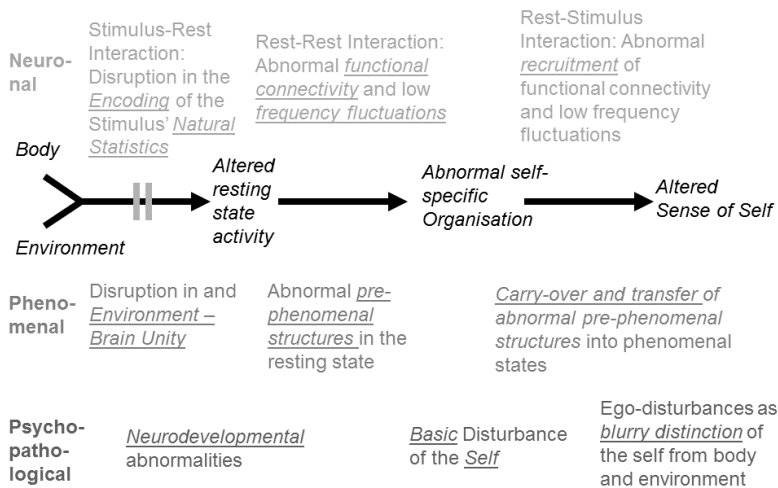
The figure shows different stages in the constitution of an altered sense of self in schizophrenia. I assume the encoding of stimuli and their natural statistics to be abnormal in schizophrenia during stimulus-rest interaction. That leads to abnormal resting state activity with abnormal pre-phenomenal structures. The abnormal resting state will then affect subsequent rest-rest interaction which leads to abnormal self-specific organisation in the resting state. Any encounter with stimuli during rest-stimulus interaction will be affected by that which results in the consecutive experience of an abnormal sense of self.

I indicated three levels, neuronal (upper), phenomenal (middle low), and psychopathological (lower). I assume that the above described neuronal mechanisms correspond to the respective phenomenal and psychopathological features.

Where does this "basic disturbance of the self" and thus the resting state's abnormal self-specific organization come from?

I assume the resting state's abnormal self-specific organization to be ultimately traced back to an abnormal encoding of the environmental stimuli's statistical frequency distribution by the resting state's neural activity. This, in turn, leads to an abnormal "environment brain unity," which in turn affects the resting state's self-specific organization.

Analogously, I assume the resting state and its abnormal self-specific organization to be carried over and transferred to subsequent stimulus-induced activity and thus any kind of neural processing. If so, one would assume the resting state's abnormal self-specific organization to lead to abnormal experiences on the phenomenal level of consciousness. This will be the focus of the next section.

Neuropsychopathological Hypothesis Ib: Abnormal Self-Specific Organization in Schizophrenia

How do these phenomenal descriptions relate to the here postulated neuronal mechanisms?
I assume that they reflect an abnormal self-specific organization in the resting state as it is carried over and transferred to subsequent stimulus-induced activity. This is in accordance with the earlier described neuronal results of an abnormal resting state and abnormal neural activity during self-specific stimuli.

Let me be more specific. Due to the resting-state abnormalities, the stimulus cannot be properly integrated into the resting state's pre-phenomenal self-specific organization. The lack of the stimulus's integration into the resting state's self-specific organization leads then to a lack of self-specificity of the stimulus itself during stimulus-induced activity. This decreased or lacking assignment of self-specificity to the stimulus is then phenomenally manifest in what is described as decreases in both self-affection and sense of mineness and belongingness. This amounts to exactly the way Parnas and others characterize experience in schizophrenia.

How about phenomenal consciousness itself? Unlike patients in a vegetative state, schizophrenic patients still show phenomenal consciousness. I assume the occurrence of phenomenal consciousness in schizophrenia to be possible because of the transfer and carryover of their resting state's self-specific and pre-intentional organization. The problem is that what is carried over and transferred is abnormal by itself. And these abnormalities are then imposed upon the subsequent stimulus-induced activity and become thereby manifest in experience, that is, phenomenal consciousness.

The case of schizophrenia must thus be distinguished from vegetative patients who do not show any phenomenal consciousness. Here I assume the process of carryover and transfer by itself to be affected. Hence, we may need to distinguish between "what" is carried over and transferred and "how" this occurs. The "what" is altered in schizophrenia, while the "how" is disturbed in the vegetative state. As we can see, both lead to radically different consequences for phenomenal consciousness.

This also pertains to the distinction between the form and state of consciousness [1, 2]. The form or organization of consciousness concerns the "what" and thus the resting state and its pre-phenomenal structures. This is abnormal in schizophrenia and depression as neuropsychiatric disorders, while the "how" refers to the state of consciousness, the level of consciousness, that, as I assume, is related to the degree of carryover and transfer. This is deficient in the vegetative state and other disorders of consciousness (see for instance: [73, 74]).

Conclusion

I focused on the abnormalities of resting-state activity across different regions in MDD and schizophrenia. Thereby we put the current findings of resting-state hyper- and hypo-activity into a wider neuroanatomical context, which suggests multiple radial-concentric and vertical integrations between subcortical and cortical regions. The inner-middle ring of subcortical core-paracore and cortical paralimbic-midline regions shows abnormal resting-state hyperactivity in MDD while the outer ring of the lateral subcortical and lateral cortical region is relatively hypoactive. Since the different subcortical-cortical systems, e.g., the different rings, include different functional systems as for affect, bodily perception, reward, cognition, etc., abnormal resting-state activity leads to abnormal rest-stimulus interaction and consecutively to the different kinds of symptoms in MDD. Schizophrenia too showed abnormal neural activity in anterior and posterior cortical midline structures as well as in the resting state, which seem to be related to the abnormal experience of the self.

Both disorders point out the central role of the resting state in these disorders in general and its particular relationship to the abnormalities of the self. The exact neural mechanisms underlying the self in general and its abnormal changes in psychiatric disorders remain unclear, however. Future investigations are therefore necessary to reveal the neural mechanisms which, in turn, may then yield diagnostic and therapeutic markers that allow for the consecutive development of endophenotypes and biomarkers and thus for clinical utility.

References

1 Northoff, G. (2013b).” Gene, brains, and environment-genetic neuroimaging of depression.” *Curr Opin Neurobiol 23*, 133-142

2 Northoff, G. (2013a). “What the brain's intrinsic activity can tell us about consciousness? A tri-dimensional view.” *Neurosci Biobehav Rev 37*, 726-738

3 Northoff, G. (2011). “Self and brain: What is self-related processing?” *Trends Cogn Sci 15*, 186-187

4 de Greck, M.et al (2010). “Decreased neuronal activity in reward circuitry of pathological gamblers during processing of personal relevant stimuli.” *Hum Brain Mapp 31*, 1802-1812

5___, Supady A et al (2009). "Decreased neural activity in reward circuitry during personal reference in abstinent alcoholics--a fmri study." *Hum Brain Mapp 30*, 1691-1704

6 Doering, S. et al (2013). "Assessment of personality functioning: Validity of the operationalized psychodynamic diagnosis axis iv (structure)." *Psychopathology 2013*

7 Gillihan, S., Farah M. (2005). "Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience." *Psychol Bull 131*, 76-97

8 Legrand, D., Ruby, P. (2009). "What is self-specific? Theoretical investigation and critical review of neuroimaging results." *PsycholRev 116*, 252-282

9 Northoff, G. et al (2006). "Self-referential processing in our brain--a meta-analysis of imaging studies on the self." *Neuroimage 31*, 440-457

10 Qin P, Northoff G (2011). "How is our self related to midline regions and the default-mode network?" *Neuroimage 57*, 1221-1233

11 van der Meer, L. et al (2010). "Self-reflection and the brain: A theoretical review and meta-analysis of neuroimaging studies with implications for schizophrenia." *Neurosci Biobehav Rev 34*, 935-946

12 Christoff, K. et al (2011). "Specifying the self for cognitive neuroscience." *Trends Cogn Sci 15*, 104-112

13 Northoff, G., Panksepp, J. (2008). "The trans-species concept of self and the subcortical-cortical midline system." *Trends Cogn Sci 12*, 259-264

14 Northoff, G. et al (2009). "Differential parametric modulation of self-relatedness and emotions in different brain regions." *HumBrain Mapp 30*, 369-382

15 Panksepp, J., Northoff, G. (2009). "The trans-species core self: The emergence of active cultural and neuro-ecological agents through self-related processing within subcortical-cortical midline networks." *ConsciousCogn 18*, 193-215

16 Schneider, F. et al (2008). "The resting brain and our self: Self-relatedness modulates resting state neural activity in cortical midline structures." *Neuroscience 157*, 120-131

17 Northoff, G. et al (2011). "The 'resting-state hypothesis' of major depressive disorder-a translational subcortical-cortical framework for a system disorder." *Neurosci Biobehav Rev 35*, 1929-1945

18 Feinberg, T. (2009). "From axons to identity: Neurological explorations of the nature of the self" (=Norton Series on Interpersonal Neurobiology). WW Norton & Company

19 ___ (2011). "Neuropathologies of the self: Clinical and anatomical features." *Conscious Cogn 20*, 75-81

20 Feinberg, T. et al (2010). "The neuroanatomy of asomatognosia and somatoparaphrenia." *Journal of neurology, neurosurgery, and psychiatry 81*, 276-281

21 Northoff, G. et al (2010). "Rest-stimulus interaction in the brain: A review." *Trends Neurosci 33*, 277-284

22 Lou, H. (2010). "Coherence in consciousness: Paralimbic gamma synchrony of self-reference links conscious experiences." *HumBrain Mapp 31*, 185-192

23 ___, Joensson, M. et al (2011). "Recurrent activity in higher order, modality non-specific brain regions: A granger causality analysis of autobiographic memory retrieval." *PLoSOne 6*, e22286

24 Lou, H., Kjaer, T. et al (1999). "A 15o-h2o pet study of meditation and the resting state of normal consciousness." *Hum Brain Mapp 7*, 98-105

25 ___, Luber, B. et al (2004). "Parietal cortex and representation of the mental self." Proc *Natl Acad Sci USA 101*, 6827-6832

26 Lou, H. et al (2010). "Self-specific processing in the default network: A single-pulse tms study." *Exp Brain Res 207*, 27-38

27 Lou, H. et al (2005). "The mental self." *Prog Brain Res 150*, 197-204

28 Buckner, R. et al (2008). "The brain's default network: Anatomy, function, and relevance to disease." *Ann NY Acad Sci 1124*, 1-38

29 Raichle, M. (2001). "A default mode of brain function." *Proc Natl Acad Sci USA 98*, 676-682

30 D'Argembeau, A. et al (2005). "Self-referential reflective activity and its relationship with rest: A pet study." *Neuroimage 25*, 616-624

31 Whitfield-Gabrieli, S. et al. (2009). "Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia." *Proc Natl Acad Sci USA 106*, 1279-1284

32 Hasler, G., Northoff, G. (2011). "Discovering imaging endophenotypes for major depression." Mol Psychiatry 16, 604-619

33 Kuhn, S., Gallinat, J. (2013). "Resting-state brain activity in schizophrenia and major depression: A quantitative meta-analysis." *Schizophr Bull 39*, 358-365

34 Treynor, W.et al (2003). "Rumination reconsidered: A psychometric analysis." *Cognitive Therapy and Research 27*, 247-259

35 Northoff, G. (2007). "Psychopathology and pathophysiology of the self in depression - neuropsychiatric hypothesis." *J Affect Disord 104*, 1-14

36 Ingram, R. (1990). "Self-focused attention in clinical disorders: Review and a conceptual model." *Psychol Bull 107*, 156-176

37 Wiebking, C. (2010). "Abnormal body perception and neural activity in the insula in depression: An fmri study of the depressed "material me"." *World JBiolPsychiatry 11*, 538-549

38 Grunebaum, M ET AL (2005). "Symptom components of standard depression scales and past suicidal behavior." *J Affect Disord 87*, 73-82

39 Boker, H. et al (2000). [self concept and object relations of patients with affective disorders--individual centered diagnosis with the repertory-grid technique]. *Psychotherapie, Psychosomatik, medizinische Psychologie 50*, 328-334

40 Rimes, K., Watkins, E. (2005). "The effects of self-focused rumination on global negative self-judgements in depression." *BehavRes Ther 43*, 1673-1681

41 Alcaro, A. et al (2010). "Is subcortical-cortical midline activity in depression mediated by glutamate and gaba? A cross-species translational approach." *Neurosci Biobehav Rev 34*, 592-605

42 Mayberg, H. (2002). "Depression, ii: Localization of pathophysiology." *Am J Psychiatry 159*, 1979

43 ___(2003). "Modulating dysfunctional limbic-cortical circuits in depression: Towards development of brain-based algorithms for diagnosis and optimised treatment." *Br Med Bull 65*, 193-207

44 ___ (2009). "Targeted electrode-based modulation of neural circuits for depression*." J Clin Invest 119*, 717-725

45 Phillips, M.et al (2003). "Neurobiology of emotion perception ii: Implications for major psychiatric disorders." *Biological psychiatry 54*, 515-528

46 Price, J., Drevets, W.(2010). "Neurocircuitry of mood disorders." *Neuropsychopharmacology 35*, 192-216

47 Fitzgerald PB, Oxley TJ, Laird AR, Kulkarni J, Egan GF, Daskalakis ZJ: An analysis of functional neuroimaging studies of dorsolateral prefrontal cortical activity in depression. Psychiatry Res 2006;148:33-45.

48 Savitz, J., Drevets, W.(2009a). "Bipolar and major depressive disorder: Neuroimaging the developmental-degenerative divide." *Neurosci Biobehav Rev 33*, 699-771

49 _____ (2009b). "Imaging phenotypes of major depressive disorder: Genetic correlates." *Neuroscience 164*, 300-330

50 Phillips, M. (2003). "Understanding the neurobiology of emotion perception: Implications for psychiatry." *Br J Psychiatry 182*, 190-192

51 Lemogne, C. et al (2012). "Medial prefrontal cortex and the self in major depression." *J Affect Disord 136*, e1-e11

52 Grimm, S.(2011). "Reduced negative bold responses in the default-mode network and increased self-focus in depression." World J Biol Psychiatry 12, 627-637

53 ___, Ernst, J. et al (2009). "Increased self-focus in major depressive disorder is related to neural abnormalities in subcortical-cortical midline structures." Hum Brain Mapp 30, 2617-2627

54 Lemogne, C. et al (2009). "Cognitive avoidance of intrusive memories and autobiographical memory: Specificity, autonoetic consciousness, and self-perspective." *Memory 17*, 1-7

55 ___, le Bastard, G. et al (2009). "In search of the depressive self: Extended medial prefrontal network during self-referential processing in major depression." *Soc Cogn Affect Neurosci 24*, 305-312

56 ___, Mayberg, H.,et al (2010). "Self-referential processing and the prefrontal cortex over the course of depression: A pilot study." *J Affect Disord 124*, 196-201

57 Kraeplin, E. (1913). *Ein Lehrbuch für Studierende und Ärzte*. Leipzig: Barth

58 Bleuler, E. (2010). "[dementia praecox or the group of schizophrenias"]. *Vertex 21*, 394-400

59 Jaspers, K. (1964). *General psychopathology*. Chicago: University of Chicago Press

60 Parnas, J. (2003). "Self and schizophrenia: A phenomenological perspective." In: T.Kirchner, A. David (eds).*The self in neuroscience and psychiatry*. Cambridge Mass: MIT Press 217-241

61 Parnas, J. et al (2001). "Visual binding abilities in the initial and advanced stages of schizophrenia." *Acta PsychiatrScand 103*, 171-180

62 Sass, L., Parnas, J. (2003). "Schizophrenia, consciousness, and the self." *Schizophr Bull 29*, 427-444

63 Pomarol-Clotet, E. et al (2008). "Failure to deactivate in the prefrontal cortex in schizophrenia: Dysfunction of the default mode network?" *Psychol Med 38*, 1185-1193

64 Mannell, M et al (2010). "Resting state and task-induced deactivation: A methodological comparison in patients with schizophrenia and healthy controls." *Hum Brain Mapp 31*, 424-437

65 Calhoun V et al. (2008). "Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder." *Hum Brain Mapp 29*, 1265-1275

66 Jafri, M. et al (2008). "A method for functional network connectivity among spatially independent resting-state components in schizophrenia." *Neuroimage 39*, 1666-1681

67 Park, I. et al (2009). "Medial prefrontal default-mode hypoactivity affecting trait physical anhedonia in schizophrenia." *Psychiatric Research 2009*, 155-165

68 Williamson, P (2007). "Are anticorrelated networks in the brain relevant to schizophrenia?" *SchizophrBull 33*, 994-1003

69 Hoptman, M. et al (2010). "Amplitude of low-frequency oscillations in schizophrenia: A resting state fMRI study. *SchizophrRes 117*, 13-20

70 Rotarska-Jagiela, A.(2010). "Resting-state functional network correlates of psychotic symptoms in schizophrenia." *Schizophr Res 117*, 21-30

71 Holt, D. et al (2011). "An anterior-to-posterior shift in midline cortical activity in schizophrenia during self-reflection." *Biol Psychiatry 69*, 415-423

72 Taylor, S. et al (2007). "Medial frontal hyperactivity in reality distortion." *Biol Psychiatry 61*, 1171-1178

73 Huang, Z. et al (2013). "The self and its resting state in consciousness: An investigation of the vegetative state." *HumBrain Mapp 2013*

74 Qin, P. et al (2010). "Anterior cingulate activity and the self in disorders of consciousness." *Human brain mapping 31*, 1993-2002

...

16

On Autism and Interaction Theories of the Self

.

Inês Hipólito

Instituto de Filosofia da Nova (IFLNOVA), FCSH-UNL, Portugal

.

Abstract: *Persons with Autistic Spectrum Disorder show various difficulties in social skills, cognitive processing and other co-occurring behavioral and physical problems.*
Taking into account that people with autism demonstrate a lack of ability in dealing with others, the study of autism, with its specific constellation of behavioral and cognitive deficiencies may highlight the structure, development and nature of social cognition in general and learning by interaction, in particular. In this article, I will challenge the Theory of Mind view that characterizes social cognition as an individual achievement that happens within a particular person's brain and treats social phenomena as external events that require interpretation, namely, mindreading. Alternatively, I will support Shawn Gallagher's "Interaction Theory" which proposes that the ability to "read" others is primarily a form of body reading rather than mind reading.
I will emphasize that in autism, a first-person perspective based on pre-reflexive, essentially non-conceptual bodily attunement (affective framing) often is replaced with the third-person perspective involving the application of algorithms, general principles and abstract rules of pure logic, and I will claim that learning by interaction is an impairment of Autism not due to an intellectual or cognitive impairment but rather ensue from an embodied cognitive impairment.

1. Clinical Properties of Autistic Spectrum Disorder

The word "Autism" derives from the Greek word "auto" meaning "I", "same", "directed from within". Kanner's early work includes several references to the *extreme self-focus* exhibited by the children he examined. He writes that one child "behaved as if people as such did not matter or even exist", and another gave "the impression of being self-absorbed". Another child is described as following:

"he got happiest when left alone, almost never cried to go to his mother, did not seem to notice his father's homecomings, and was indifferent to visiting relatives (…) he seems to be self-satisfied to get his attention almost requires one to break down a mental barrier between his inner consciousness and the outside world" (Kanner 1943).

Frith and colleagues refer to this self-absorption as naïve egocentrism, and describe how it can be a source of difficulty in social interchange for individuals with autism (Frith & de Vignemont 2005).

As pervasive developmental disorder, autism is characterized by deficits in more than one functional domain. According to Wing and Gould (1979) it has been common practice to base the diagnosis on a triad of impairments: (1) Socialization, (2) Imagination, (3) Communication.

(1) Socialization includes indifference to others, disability in play with other children experiencing an interaction based on need (needs-only basis) and often nonverbal; inability to measure the thoughts and emotions of others (failure of a meeting of minds, Kanner's notion of "Autistic aloneliness"). (2) Imagination: evident disability during games activities. While autistic children do not quickly build some drama of its creation through the toys autistic children simply organize the toys, for example, online (straight lines), build them in towers. Outside of games, problems occur when imagination offer resistance to change and reluctance to routine activities vary. (3) Communication: is a severe disability of autism. The speech (spoken language) is virtually absent in 50% of cases (Yale Child Study Center). In general the development of language is considerably delayed and markedly deviant.[1]

Persons with autism show severely diminished or abnormal social interaction and communication, as well as a restricted repertoire of activities and interests (DSM-IV, 66). These symptoms can be mild, seen in a lack of certain nonverbal behaviors such as eye-to-eye gaze and gestures or any type of social interaction, or a more serious lack of all reciprocal social interaction and other large impairments in language development and language use.

Autists seem to acquire only a portion of self and yet, many children with autism have some concept of self and self-reflection. In fact, their concepts of self and their magnitude of self-directed attitudes is limited due to their inability to few dissociative lines of development that contribute to the typical self-development. Nonetheless, children with autism demonstrate abilities to recognize intentions and actions directed to an objective; however, they have a relative lack of propensity to respond to the attitudes of the other. Indeed, they can copy actions with

intentionality directed (goal-directed actions), but do not tend to identify with the subject of the action.

For Vygotsky, as in "The Genesis of Higher Mental Functions", these "higher mental functions" of cognition such as thinking, planning, memory, are socially founded relations, and their relationships as mental functions reflect social processes.

According to Bakhtin (1981), *reasons*, *self* and *other* are jointly involved in the determination of meaning: "consciousness is never self-sufficient; it always finds itself in the intense relationship with other consciousness".

Recent evidence from neuroscience and more specifically from functional MRI Scanning Mirror Neurons or other insights of self-other correspondence in psychological connectedness and interaction, showed us how something like empathy and identification can operate at the neuronal (Decety and Chaminade 2003, Gallese 2001) and physiological funcionnal level. Behavioral impairments in autism are theorized to result from abnormal neuronal organization in brain development generating 4 systemically related neurofunctional impairments: (a) canalesthesia, wherein abnormal hippocampal system function "canalizes" sensory records, disrupting integration of information; (b) impaired assignment of the affective significance of stimuli, wherein abnormal amygdaloid system function disrupts affect association; (c) asociality, wherein impaired oxytocin system function flattens social bonding and affiliativeness; and (d) extended selective attention, wherein abnormal organization of temporal and parietal polysensory regions yields aberrant overprocessing of primary representations. This model proposes that complex human behaviors may be guided by multiple overlapping neural mechanisms (Waterhouse et al 1996).

The theory of cognitive development argues that the concepts of self are necessary as the child feels and expresses social emotions since these involve an elaborated cognitive process. It is not possible to consider an emotion as "social" without the individual experience of the other person's awareness. At a minimum, this awareness has to be constituted by having an emotion of the kind that sets a personal relatedness, that is: the emotion can establish social cognition.

The typical human development as a development of interpersonal relationships and social understanding depends on the structure of a self-other very early on in life. The self can be seen as an "atom" in the world of relational processes. Thus, there are selves that come together to form relationships, but it is rather a relational process from which the idea of a psychological self can emerge.

2. Learning by Interaction: an impairment in Autism Spectrum Disorder?

One of the most widely discussed approached to explain the cognitive and behavioral aspects of Autism is Theory of Mind (ToM), it characterizes social cognition as an individual achievement that happens within a particular person's brain and body and treats social phenomena as external events that require interpretation (Maiese 2013). Through this perspective, we first remark the other person's behavior and expressions, and then we use mind reading to understand him. The theory theory (TT), a version of ToM, supports that the pervasive way in which we understand other persons depends on our implicit or explicit practice of mentalizing or mind-reading, in which we use a common sense or folk-psychological theory about how mental states (beliefs, desires, intentions) inform the behaviors of others.

Shaun Gallagher has challenged the Cartesian idea that other's mental states are hidden away and inaccessible and rejects the notion that we ordinarily act as spectators of other's behavior. According to him, TT fails to grasp the primary way in which we relate to and interact with others and, as an account of intersubjectivity or social cognition more generally, can be challenged at the level of its basic suppositions:

> "Supposition 1 (the mentalistic supposition): The problem of intersubjectivity is precisely the problem of other minds. That is, the problem is to explain how we can access the minds of others. This is a problem of access because other minds are hidden away, closed in, behind the overt behavior that we can see. This is a mentalistic and clearly Cartesian supposition about the very nature of what we call the mind. The mind is conceived as an inner realm, in contrast to behavior, which is external and observable, and which borrows its intentionality from the mental states that control it. Both theory theory and simulation theory set the problem as one of gaining access to other minds, and their explanations of social cognition are framed in precisely these terms.
>
> Supposition 2 (the supposition of universality): Our reliance on theory (or our reliance on simulation or some combination of theory and simulation) is close to universal. That is, this folk-psychological way of understanding and interacting with others is pervasive in our everyday life". (Gallagher, 2004)

Supposition 1 is to be found, explicitly or implicitly, in almost every description of intersubjective interaction that proponents of theory of mind propose. A typical example is Happé's characterization:

"to have a theory of mind is to be able to attribute independent mental states to self and others in order to explain and predict behavior." Autism, then, is "an impairment of the fundamental human ability to 'mind-read'"(Happé 1995, 38).

Supposition 2 is easily documented in the many strong claims made for universality by proponents of theory of mind. According with Gallagher (2004), "our normal procedure is to treat them [others] as bearers of mental states hidden behind their embodied, behavioral manifestations, and that our primary form of interacting with them is predicated on our attempts to explain or to predict their next move through a process of mentalizing".

Proponents of theory of mind do appeal to developmental psychology, and especially to false-belief tests, for support. In false-belief task, a subject is asked about the thoughts and actions of another person (or sometimes a puppet or character in a story) who lacks certain information that the subject has. Importantly, these are the same tests that many autistic subjects fail to pass, and it is primarily on this basis that the theory of mind account of autism is developed.

In this context, we shall raise the question: does the false belief paradigm capture everything there is to say about children's abilities to understand others, or does it even capture the most important things? According to Gallagher (2004), false-belief tasks are thus extremely limited in terms of trying to capture the nature of intersubjective understanding. One reason for this is that subjects are asked to predict the behavior of others with whom they are not interacting. The subject is installed in the role of third-person observer, and in this role the child is asked to predict what the other person will do. The fact that what is tested is ability for third-person observation suggests that the results of these experiments may not be applicable to second-person (I–you) interaction, which is arguably the primary and ordinary way of encountering the other person.

Alternatively, Gallagher proposes "Interaction Theory", which emphasizes how our capability to comprehend other persons ultimately rests in a form of embodied practice that is emotional, sensory-motor, perceptual and non-conceptual. Whatever theory of mind we have, depend on these embodied practices that we carry out during interpersonal interaction, and these are the basic ways of understanding each other.

Interaction Theory comprehends Primary Intersubjectivity and Secondary Intersubjectivity. Primary Intersubjectivity is the innate or early developing capacity to interact with others manifested at the level of perceptual experience— we see or more generally perceive in the other person's bodily movements, facial gestures, eye direction, and so on, what they intend and what they feel. On this view, in second-person interactions, the mind of the other is not entirely hidden or private, but is given and manifest in the other person's embodied behavior. Interaction theory contends that these embodied practices constitute our primary

access for understanding others, and continue to do so even after we attain theory of mind abilities.

> "What we might reflectively or abstractly call their belief or desire or mental state is expressed directly in their behavior" In brief, the developmental evidence for primary intersubjectivity suggests that pretheoretical (nonconceptual) sensory-motor capabilities for understanding others already exist in very young children. Infants already have a sense from their own proprioception and movement of what it means to be an experiencing subject-agent. They can sense that certain kinds of entities (but not others) in the environment are indeed subject-agents like themselves; and that in some way these entities are similar to and in other ways different from themselves". (Gallagher 2004)

As contemporary research on mirror neurons (Rizzolatti & Craigheiro 2004; Gallese & Sinigaglia 2010; 2011) endorses, our perception of the other person induces a sensory motor process that reverberates kinetically and kinesthetically with their intentions[2] (see: Gallagher 1986; Petit 1999).

The perception of others' intentional actions engages our motor and affective process in a way that TT fails to acknowledge. In neonate imitation, the newborn demonstrates proprioceptive awareness of her own body, the acknowledgment of a distinction between self and no self, and also the recognition that the other is, in fact the same kind of entity as herself.

"This interaction depends not only on a distinction between self and non-self, and a proprioceptive sense of one's own body, but on the recognition that the other is in fact of the same sort as oneself" (Bermúdez 1996, cit. in: Gallagher 2004 see as well: Gallagher and Meltzoff 1996).

The newborn infant not only can pick out a human face from the crowd of objects in its surroundings, but also can imitate the gesture it sees on that face and use facial gesture to provoke response from others (Maiese 2013). This is an imitation that cannot be mediated by theorizing or even strictly cognitive simulation, rather occurs in a direct, unmediated and fully embodied manner. In fact, the infant is capable of perceiving other persons as intentional agents. This perceptual ability is, as Scholl and Tremoulet suggest, "fast, automatic, irresistible and highly stimulus-driven" (2000, 299). By the innate body schema appeal, infants can map other's facial expressions and bodily movements into their proprioceptive bodily experiences.

Infants as young as 6 months perceive grasping as goal directed; at 9 months they follow the other person's eyes and start to perceive various body movements as

goal-directed movements; and at 10 to 11 months they are able to construe some kinds of continuous action according to intentional boundaries. In other words, infants are capable of non-mentalistic, perceptually based understanding of the dispositions of other persons and they look to other's bodies and expressive movements to make sense of their behavior.

According to Gallagher, the ability to "read" others is primarily a form of body reading rather than mind reading. Indeed, the basic bodily capacities that make humans naturally attuned to the expressions of others are largely pre-reflective, emotional, sensory-motor, perceptual and not intellectual governed. All face-to-face interactions are based in embodied intersubjectivity. Social understanding should be, thus, understood as a product of the embodied social interactions of primary intersubjectivity, which are both enactive and essentially emotive.

There are good reasons to believe that our appreciation of others as a person involves a disposition to affect them, and to be affected by them. It is during this process of mutual affection that joint meanings are generated and individuals' distinct perspectives are intersubjectively merged and modified. Interpersonal engagement is enactive in the sense that individuals "do not passively receive information from their environments, which they are translate into internal representations whose significant value is to be added later", but instead actively participate in the generation of meaning.

Even in the early stages of interpersonal engagement, the infant's body serves as sense-giving orientation through which all experience is structured, and surroundings take on meaning and significance on the basis of the infant's basic needs and preferences. This is not only because the infant must depend on others to satisfy her basic needs, but also because the infant spontaneously cares about other persons.

People use expressions, gestures and other body movements to invite the infant into some sort of communication. In early interactions between child and caregiver that require joint attention, the body plays a key role in allowing the infant to gain familiarity with her caregiver's intentions and attune her behavior accordingly. This is a result of infant's directly perception of other's desiderative feelings in their actions, gestures and expressions, while, simultaneously, becomes directly aware of the mutual influence that exists between her and her caregiver. Very young children have some basic awareness of their own ability to modulate and impact the behavior of others. Moreover, the infant's experiences inform her that others can elicit gestures, actions and expresions from her. She implicit sense that where adults point their fingers or direct eyes often shapes her own patters of attentions. As a

social agent, involved in a social interaction, the infant is "at once prodder and prodded" (Maiese 2013).

Some of the bodily processes that help to give rise to this sense of relatedness and cooperation include the key elements of primary intersubjectivity identified by Gallagher (see: Maiese 2013):

a) A perpetual "intentionality detector"

b) An "eye direction detector", and

c) A "shared attention mechanism"

This suggests that there is a link between the perception of others and the goal-oriented potentialities of one's own body, and that the direct perception of others as persons involves bodily responsiveness. In this regard, too, primary intersubjectivity is based on something that is cognitively closer to the perception of embodied comportment, rather than to a theoretical inference about emotional states. It is, of course, impossible to ask the infant to provide phenomenological descriptions of primary intersubjectivity. Yet phenomenologists have described something that seems to correspond to what is indicated by the behavioral measures of developmental psychology.

Contemporary theorists have called attention to this sort of bodily responsiveness or "motor resonance". Daniel Hutto observes that imitation and motor mimicry can be characterized as "instinctual responses to situations or other people, for which our innate systems naturally are calibrated"(Hutto 2004). This information is bound up with intentional directedness informed by our biological needs; perception of one's social environment often is quick and reliable. To engage in acts of imitation and mimicy, infants or adults need not rely on psychological principles or make inferences about people's beliefs and desires, since in many cases subjects simply read another person's basic desires straight from her reactions and expressions.

According to Hobson (2002)

"(t)he defining feature of secondary intersubjectivity is that an object or event can become a focus between people. Objects and events can be communicated about (…) the infant's interactions with another person begin to have reference to the things that surround them. The child, seeing an adult who tries to manipulate a toy in the right way and who appears frustrated about being unable to do so, quite readily picks up the toy and shows the adult how to do it" (see: Meltzoff 1995; Meltzoff and Brooks 2001).

Quite obviously this understanding depends on shared attention and the pragmatic context, just as we understand our own actions on the highest pragmatic level

possible. So again, the evidence indicates that well before the development of a theory of mind mechanism, the child looks to the body and the expressive movement of the other to discern the intention of the person or to find the meaning of some object. In this kind of second-person interaction two-year-olds are even capable of recognizing pretend behavior, for example, the mother pretending the banana is a telephone (Leslie 1994). Our understanding of the others' expressive movements depends on their involvement in meaningful instrumental/pragmatic contexts. Therefore, according to Gurwitsch, the meaningful encounter with others is not only embodied, but contextually embedded and pragmatic. "Prior to all specific cognition, and independent of it, we are concerned with other people in our 'natural living' of daily life" (1978, 35)— in the pragmatic contexts of life.

From the "Interaction Theory" perspective, social interaction, intersubjectivity, and the perception of others are not explainable in terms of socially contextualized behavior is irreducible to either pragmatically contextualized or abstractly noncontextualized behavior.

The interaction theory of intersubjectivity, then, includes two important and related, but nonetheless distinct, elements:

- – Primary intersubjectivity: embodied, sensory-motor (emotion-informed) capabilities that enable us to perceive the intentions of others (from birth onward);
- – Secondary intersubjectivity: embodied, perceptual, and action capabilities that enable us to understand others in the pragmatically contextualized situations of everyday life (from twelve to eighteen months of age onward).

According to Gallagher (2004), "on this view, the theory of mind is, at best, a set of specialized cognitive abilities that allow us to *mentalize* on rare Cartesian occasions (from four years onward)".

There are good reasons to believe that, in addition to their impaired ability to understand other people's mental states, autistic subjects exhibit wide ranging sensory abnormalities that make bodily engagement with physical world dramatically different that of ordinary. Because autistics lack a clear sense of the meaning and significance of their own bodily feelings, they will find it difficult to engage in social referencing behaviors or shared attention. The sensory disturbances commonly associated with autism, which are symptomatic of a disruption is embodied emotion consciousness, lead directly to an inability to become attuned to other people.

For this reasons, the distinction that Baron-Cohen (1999) sorts between understanding the physical world and understanding the mental world is

misguided because It fails to acknowledge the sense in which understanding other minds is largely a matter of becoming attuned to other's living. As Baron-Cohen himself highlights, pointing gestures, gaze monitoring, showing gestures, and others aspects of joint attention often are absent in children with autism.

Autistics are incapable to monitor or direct another person's focus of attention, use gestures to engage in social communication, and coordinate their own behavior and attention to others. It seems unlikely that autism results essentially from a disturbance in meta-cognitive or higher-level mechanisms of mentalizing or theorizing. Social cognition rests largely on the capacity for body attunement rather than theorizing and it is disrupted or etiolated patterns of essentially embodied experience and bodily attunement that make it difficult for autistic subjects to connect with and understand others.

Because autistic subjects cannot perceive the intentions or emotions of others in their bodily comportment or participate fully in reciprocal bodily modulation, high-functioning autistics may employ theorizing strategies or mentalizing as a way to compensate. Nevertheless, it is indeed because of the lack of ordinary way of understanding others, and so must try to make algorithmic and explicit what for most of us is second in nature (Maiese 2013). In autism, a first-person perspective based on pre-reflexive, essentially nonconceptual bodily attunement (affective framing) often is replaced with the third-person perspective involving the application of algorithms, general principles and abstract rules of pure logic..

Uta Frith and Frederique de Vignemont (2005) depict social cognition among autistic subjects as relying primarily on "abstract allocentrism". Autistic subjects take up the allocentric stance in direct social interaction. They rely heavily on rules and predict other people's behavior on the basis of regularities between inputs and outputs. Therefore, while autistics are able to understand social structures and interpersonal relationships in a detached way, their sense of how the social world should work often is very logical, formal and far removed from the reality of most everyday social encounters. They are capable only of an abstract allocentric stance that is disconnected from embodied, face-to-face interactions, so they tend to rely on normative rules rather than the desiderative bodily feelings of affective framing when empathizing with others. One might say that people with autism operate in the way one might expect all humans to operate if TT were true, by appealing to a set of generalizations and approaching social cognition from a detached, third person perspective. However, it is important to keep in mind that this sort of rule following is not the way the ordinary subjects understand other people, but instead a way that high-functioning autistics attempt to compensate for their social cognition deficits. The application of rules helps to make sense of what autistic subjects appear to be sensitive to moral considerations in certain respects, and yet

lack moral competence in other. The data listed above, lead us to infirm that there is not an impairment on the intellectual o cognitive level, but rather, an impairment on the level of perception, this is, on embodied cognition which precludes learning by interaction, in other terms, it precludes we-experiences.

Conclusions

If the account of bodily attunement that I have presented is correct, then Autism should not be understood as "mindblindness" or impairment in the ability to theorize about what is going on in others' minds, infer their mental stantes, and predict their behavior. The available evidence (Frith & Vignemont 2005) does not indicate that children with autism are impaired with respect to emotion recognition, rather they suffer from emotion processing abnormalities and find it more difficult to regulate and reflect on their emotions – this may indicate that although they have an autonomic affective response, they are unable to decipher fully the meaning of what they are feeling, and, as a result, their emotions do not come to the interaction scenario during information processing. In Autism the cognitive and the affective elements come apart, so that the desiderative bodily feelings or affective framing no longer effectively focus attention or guide the subject in making sense of others' behavior (Maiese 2013). It doesn't seem reliable that they are insensitive to others or that they simply do not care, rather than their caring is more detached from cognitive processing than it is from ordinary subjects. Maiese argues that this apparent detachment can be explained as a deficiency in bodily attunement and an impairment capacity for affective framing. Although autistics do exhibit a weakened sense of emotional connectedness, they still are able to participate in moral life to some extent and are capable of some mode of empathy. Athwart the development of moral rules and principles of conduct, they can fashion long-term stable values and concerns for themselves and care about the interests of others. While many of us are empathically engaged spontaneously and become immediately attuned to the concerns of others, autistics rely more on rules of conduct and general principles in order to navigate the social landscape. This might be described as a "cold" methodology that engages the interest to a great extent in order to bring about the sort of affective, bodily attunement that for ordinary subjects is already there and occurs spontaneously and pre-reflectively via affective framing.

Since autistics' mode of empathy is more rule-driven, their participation in social and moral scene differs from ordinary subjects. As autistics are unable to grasp human emotions intuitively and pre-reflectively through bodily attunement, they, instead, rely on pure logic.

Endnotes

[1] Autism diagnose is guided by checklists such as is found in DSM. Items are used to assess socialization and imagination, that include "pervasive lack of responsiveness to other people" and bizarre responses to various aspects of the environment. Regarding communication, there are items such as "gross deficits in language development", "if preach is present, peculiar speech patterns, such as immediate and delayed echolalia and pronoun reversal" and "abnormalities of speech, such as question-like melody or monotone robotic-like voice".

[2] This insight was much developed by Merleau-Ponty who wrote that "between this phenomenal body of mine and that of another as I see it from the outside, there exists an internal relation which causes the other to appear as the completion of the system" (1962, 315). Gallese has been developing the implications of the mirror neuron research for these kinds of questions, although he takes the simulationist position.

References

Andersen, J., Cushing, S. (2013). *Philosophy of Autism*. London: Rowman Littlefield

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders. [DSM -IV]* 4th ed., text rev.). Washington, DC

Bakhtin, (1981). *The Dialogic Imagination. Four Essays by M.M. Bakhtin.* Austin: University of Texas Press

Baron-Cohen, S., (1995). *Mindblindness.* Cambridge, Mass: The MIT Press

___ (2003). *The Essential Difference*. New York: Basic Books

___ et al (1985). "Does the autistic child have a "theory of mind"?" *Cognition 21*, 37-46

Barkow, J. et al (1992). *The Adapted Mind*. New York. Oxford University Press

Barresi, J. & Moore, C. (1996). "Intentional relations and social understanding". *Behavioral and Brain Sciences 19(1)*, 107-122

Braten, S. (1998). "Infant Learning by Altercentric Participation: The Reverse of Egocentric Observation in Autism". In: S. Braten (ed.). *Intersubjective Communication and Emotion in Early Ontogeny.* Cambridge: Cambridge University Press, 105-26

Bechtel, W., Richardson, R. (1992). Discovering Complexity. Princeton, NJ: Princeton University Press

Blake, R. et al (2003). "'Visual Recognition of Biological Motion' is Impaired in children with Autism." *Psychological Science 14*, 151-158

Castelloe, P., Dawson, G. (1993). "Subclassification of Children with Autism and Pervasive Developmental Disorders. A Questionnaire bases on the Wing and Gould Subgrouping Scheme." *Journal of Autism and Developmental Disorders 33*, 229-241

Ceponiene, R. et al (2003). "Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend." *Proceedings of the National Academy of Sciences 100*, 5567-5572

Cundall, M. (2006). "Autism's Role in Understanding Social Cognition." *Journal of Humanities & Social Sciences1*, 1

http://www.scientificjournals.org/journals20 07/articles/1002.htm

Currie, G., Sterelny, K. (2000). "How to Think about the Modularity of Mindreading." *The Philosophical Quarterly 50*, 145-162

Cummings, L. (2005). *Pragmatics: A Multidisciplinary Perspective.* Edinburgh: Edinburgh University Press

Dawson, G., et al (1995). "Subgroups of Autistic Children Based on Social Behavior Display Distinct Patterns of Brain Activity." *Journal of Abnormal Child Psychology 23*, 569-583

Decety, J., Chaminade, T. (2003). "When the Self represents the Other: a new cognitive neuroscience view on psychological identification". *Consciousness and Cognition 12(4)*, 577-596

Fodor, J., (1980). "Special Sciences, or the Disunity of Science as a Working Hypothesis." In: N. Block (ed). *Readings in the Philosophy of Psychology*. Vol. I. Cambridge, MA: Harvard Publishers

Fodor, J., (2000). *The Mind Doesn't Work That Way*. Cambridge, Mass: The MIT Press

Frith, C., Frith, U. (1999). "Interacting minds – a biological basis." *Science 286*, 1692–1695

Frith, U., & de Vignemont, F. (2005). "Egocentrism, allocentrism, and Asperger syndrome". *Consciousness and Cognition 14(4)*, 719–738

Gallagher, S. (2001). "The Practice Mind". *Journal of Consciousness Studies 8 (5-7)*, 83-108

___ (2004). "Understanding interpersonal problems in austism: interaction theory." *PPP 11 (3)*, 199-217

___ (2009) "Inference or Interaction: Social Cognition". *Philosophical Explorations: An International Journal for the Philosophy of Mind and Action 11 (3)*, 163-174

___ (2000). "Philosophical conceptions of the self: Implications for cognitive science". *Trends in Cognitive Sciences 4(1)*, 14–21

___(1986). "Lived Body and Environment". *Research in Phenomenology 16 (1)*, 139-170

Gallagher, S., Meltzoff, A. (1996). "The Earliest Sense of Self and Others: Merleau-Ponty and Recent Development Studies". *Philosophical Psychology 9*, 213-236

Gallese, V., Sinigaglia, C. (2010). "The Bodily Self as a power for action." *Neuropsychologia 48 (3)*, 746-55

_____ (2011). "What is so special about embodied simulation?" *Trends of Cognitive Sciences 15(11)*, 515

Goldman, A., (2006). *Simulating Minds*. New York. Oxford University Press

Gopnik, A., Meltzoff, A. (1998). *Words, Thoughts and Theories*. Cambridge, Mass: The MIT Press

Hamlyn, D. (1974) "Person Perception and our understanding of Others". In: T. Michel (ed.) *Understanding Other Persons*. Oxford: Blackwell, 1-36

Happé, F. (1994). "An advanced test of theory of mind." *Journal of Autism and Developmental Disorders 24 (2)*, 129-154

Hobson, P. (2011). "Autism and the Self". In: S. Gallagher (Ed.). *The Oxford Handbook of The Self*. New York: Oxford University Press, 571-591

___ (2008). "Interpersonally Situated Cognition". In: *International Journal of Philosophical Studies 6*, 377-397

Horn, L. (1978). *Experience and Growth of Understanding*. London: Routledge & Kegan Paul

Hutto, D. (2004). "The Limits of Spectatorial Folk Psychology." *Mind & Language 19(5)*, 548-573

Leslie, A. (1994). "Pretending and Believing: Issues in the Theory of ToMM." *Cognition 50*, 211-238

Lord, C., Paul, R. (1997). Language and communication in autism. In: D. Cohen & F.

Kanner (1943). "Autistic disturbances of affective contact." *Nervous Child 2*, 217–250

___ (1997). "An Investigation of Language Impairment in Autism. Implications for Generic Subgroups". In: *Lang Cogn process 16 (2-3)*, 287-308

Maiese, M. (2013). *Autism, Empathy and Affective Framing in Philosophy of Autism*. United Kingdom: Rowman& Littlefield Publishers

Merleau-Ponty, M. (2005[1964]). *Phenomenology of Perception*. Transl. from the French by Colin Smith. Routledge & Kaegan Paul

Malle, B., et al. (2001). *Intentions and Intentionality: Foundations of Social Cognition*. Cambridge, MA: The MIT Press

Mead, G. (1934). *Mind, Self and Society*. Chicago: University of Chicago Press

Malcolm, N. (1971). *Problems of Mind: Descartes to Wittgenstein*. Crows Nest: Allen & Unwin

Rizzolatti, G., Craigheiro, L. (2004). "The mirror neuron system." *Annual Review of Neuroscience 27*, 169-192

Schmid, H. (2012). *Cognitive Pragmatics*. Mouton De Gruyter

Strawson, P. (1962[1958]). "Persons". In: V. Chappell (ed). *The Philosophy of Mind. Englewood Cliffs*. NJ: Prentice-Hall, 127-46

Uddin, L. (2011). "The self in autism: An emerging view from neuroimaging." *Neurocase, 17(3)*, 201-208

Volkmar, F. et al (Eds.), *Handbook of autism and pervasive development disorders*. 2nd edition. New York: John Wiley

17

Autism: a Relational Self Impairment?

.

Anna Ciaunica
University of Fribourg, Switzerland

.

Abstract: *There is controversy over the basis for young children's experience of themselves and others as separate yet related individuals, each with a mental perspective on the world – and over the nature of corresponding deficits in autism[1]. On the dominant view, it has been argued that the human aptitude at inferring mental states is one of the crucial preconditions for the evolution of cooperative social structure in human societies. Consequently, limitations in psychological perspective-taking among autistic children have been regarded as signs of a "Mindblindness" deficit in having metarepresentational mental states (Baron-Cohen et al 1985; Leslie 1994). Lombardo and Baron-Cohen (2010) recently acknowledged that Mindreading researches have left a gap in terms of devices that may be responsible not only for atypical self-referential processes in autistic children but also in their integration into the bigger picture of how individuals navigate and interact with the social world. This paper defends the idea that the intersubjective relatedness of the self is foundational for the development of higher order mentalizing skills and not the other way around. Hence, autistic individuals might suffer from a specific Relational-Self Impairment which triggers metarepresentational deficits.*

Keywords: autism, enactivism, intersubjectivity, relational self

Introduction

Central to the study of the development of self/other connectedness are questions that concern our most basic understanding of other people. Primates, and particularly human beings, are social animals whose cognitive capacities develop upon interaction with other conspecifics. But how can we *relate* to what is going on in other people's minds? Within the dominant cognitivist[2] research paradigm, there has been a two-party debate between Theory Theory (TT) and Simulation Theory (ST). TT theorists (Baron-Cohen 1985; Leslie 2000) hold that we explain and predict others' behavior by relying on (i) an *innate* modular or (ii) *acquired*

mechanism/theory of how people generally behave. ST theorists (Currie and Ravenscroft 2002; Davies and Stone 1995b; Goldman 2006) argue that we explain and predict others' behavior by using an inner model designed to simulate the other person's mental states. The basic idea is that we understand others by having some grasp of their mental states and by using this understanding to explain and predict their behavior on the basis of these mental states (Spaulding 2010). At the core of our intersubjective skills there are high-level cognitive mechanisms.

Over the past two decades progress has been limited to articulating hybrid TT/ST accounts. However, a new party to the debate has recently emerged, stemming from (1) *embodied/enactive cognition* (Lakoff and Johnson 1999; Bermúdez 2003; Hutto 2008; Gallagher 2005; Menary 2007; Chemero 2009; Hutto and Myin 2013); (2) *phenomenology* (Thompson 2001; Ratcliffe 2007; Overgaard 2007; Zahavi 2007); and (3) *Wittgenstein*'s late philosophical work (Hobson 2009). These theorists reject the idea that at the core of our intersubjective skills there are high-level cognitive mechanisms. The Embodied/Enactive Social Cognition approach is roughly the view that our normal everyday interactions consist in non-mentalistic embodied engagements. This relatively new research paradigm in cognitive science challenges the conceptual foundations of cognitivism, namely the idea that mentality is essentially content-involving and hence an exclusively heady affair (Chemero 2009; Hutto and Myin 2013) and claims instead that we have evolved from creatures whose neural resources were devoted primarily to perceptual and motoric processing, and whose cognitive activity consisted largely of immediate online interaction with the environment. Consequently, it has been suggested that perhaps "it is time we set the horse before the cart: personal relations are what provide the basis for cognitively elaborated concepts of mind in toddlerhood and beyond" (Hobson 2009, 246). If this line of reasoning is correct, then it has profound implications for our understanding of minds not as inferential or simulationist devices situated in individuals' heads, but rather as *constitutively* dependent upon worldly and social interactions. The intersubjective relatedness of the self might be foundational for and not merely subsidiary to the development of mentalizing skills. In other words, the lack of mindreading abilities in autism may be the *effect* and *not* the cause of interpersonal relatedness impairments.

This paper is structured as follows: in section 1 I briefly present and discuss current understandings of autism within the cognitivist Mindreading paradigm. Then I will interpret recent empirical findings as corroborating the idea that Mindreading theories overestimate the role of conceptual thinking and thereby fail to encapsulate the relational aspects of the self. Part 2 argues that the so-called "minimal self" (Zahavi 2005) is primarily a "relational self", i.e. constitutively dependent upon social interactions. More specifically, I discuss empirical findings suggesting that

social cognition is not reducible to intracranial mechanisms but instead emerges from within the dynamics of the interactive process itself (Krueger 2010). Section 3 defends the idea that autistic individuals might suffer from a specific Relational-Self Impairment (RSI henceforth), and reviews behavioral and neural evidence supporting this hypothesis. I conclude that the enactive research paradigm, by stressing the role of social interactions in shaping the relatedness of the self, might provide us with better tools in understanding autistic disorders.

1. A Lacking Device? The Mindblindness Hypothesis

The discipline of developmental psychopathology represents a uniquely powerful tool in characterizing the underpinning of a person's understanding of self vis-à-vis other people. The social disabilities found in autism have been emphasized as a, if not the, central defining feature of the disorder. Indeed, autistic children (AC henceforth) show a fundamental impairment in socialization from the first years, perhaps even the first days or weeks of life. Unlike typically developing infants, they display a lack of reciprocity in social contact, inappropriate gaze behavior, and paucity of joint play and communicative skills (Klin et al 1992). Later in life, symptoms of social impairment persist even in the case of the highest functioning individuals who have a pervasive difficulty in responding empathically to others (Cohen 1980).

The social impairments in autistic children have been conceptualized by three explanatory theories. (1) The "*Theory of Mind*" (ToM) (Leslie 2000; Baron-Cohen 1985) proponents explain these deficits in terms of very specific and primarily cognitive incapacity to attribute psychological states to others and to self. Underlying it is the assumption that what other people feel and think is internal and occluded from us, and the only clue we have is their observable behavior. AC are thought to have more trouble to "read" other people's occluded minds and to conceive of other people's subjectivity, hence the impairment in empathy display and social cognition in general. (2) The *Weak Central Coherence* (WCC) theory (Frith 1989; Happé 1994) argues that AC focus on piecemeal information and have difficulty integrating what they perceive as well as perceiving things in context. (3) Finally, the *Executive Function* (EF) theory (Ozonoff et al 1991; Russell 1998) holds that AC lack control over their actions and attention, associated with activity in the frontal lobes. A full review of this literature is beyond the scope of this paper. What seems important at this stage of our discussion is the fact that all these theories are shaped within a cognitivist-representationalist framework. Indeed, according to a longstanding philosophical tradition, to know is to have a mind that possesses internal detailed and complete representations of the outside world. The concepts of "representation"[3] and "metarepresentation"[4] have played a very important role in research in autism. Indeed, one of the main premises behind the *Mindblindness*

theory is the idea that while typically developing children (TD) have an intact mechanism for *representing* or attributing mental states to both self and other, this mechanism (Carruthers 2009) or set of mechanisms (Baron-Cohen 1985; Leslie 1994) is profoundly impaired in *autistic children* (AC henceforth).

Before setting the discussion about the Relational Self Impairment" hypothesis, we need to have a quick look at the evidence on which Mindreading statements are grounded. A popular distinction in the developmental literature is the distinction between *automatized* and *controlled* processes. Automatized processes refer to prewired or over-learned structures of behavioral organization in which *information is represented implicitly*, embedded in procedures, unavailable to other representational systems of the mind. In contrast, deliberative processes refer to voluntary and conscious operations that are flexible and modifiable and can be governed by higher-order cognitive goals. Now, the important question is: how an infant does develop awareness of and come to represent the sets of internal states cues? Are infants initially more sensitive to internal or to external stimuli? On the one hand, it has been hypothesized that infants make use of innate cross-modal capacity to map the adult's visual movements onto the proprioceptive feelings of his or her own movements that the adult is imitating. For example, Meltzoff and Gopnik (1993) proposed that there are innate mechanisms that allow the infant to attribute emotions to other minds starting from birth. Experiments on neonatal imitation (Meltzoff and Moore 1977) and on the innate basis for primary emotions (Ekman et al 1972; Izard 1977) suggest that by imitating the caregiver's facial emotion expression, the infant activates through prewired connections the corresponding physiological emotion state in himself or herself. The imitation-generated internal emotion state is then introspectively accessed, and the felt affect is attributed to the other's mind. The basic assumption behind this view is that the infant's initial state is characterized by direct introspective access to internal emotion states, and that they have conscious access to their internal basic emotion states from the beginning of life.

Opposing these views, Gergely and Watson (1999) argued that at the beginning of life "the perceptual system is set with a bias to attend to and explore the external world and builds representations primarily on the basis of exteroceptive stimuli." (1999, 110) In other words, infants are born with an innate Contingency Detection Module (CDM). More specifically, the CDM is supposed to identify the contingent relatedness between internal proprioceptive stimuli and external affect-mirroring display.

These debates have been abundantly discussed in the literature and the details will not be pursued here. For the moment suffice it to say that current research attempt to appraise the infants' preconceptual senses of the self and others and aims to

explain how these contribute to their increasingly sophisticated and linguistically elaborated understanding of persons with minds. The core assumption of the tradition representationalist view is that there are subpersonal contents and subpersonal operations that are truly cognitive in the sense that these operations can be properly explained only in terms of these contents.

In autistic children, for reasons that are not yet understood, there are failures of both the basic set of metarepresentational skills and at the intersubjective connectedness level. Indeed, both clinical descriptions (Kanner 1943) and experimental studies (Hobson et al. 2006,) illustrate how AC appear to lack a grasp of self and other in social interactions. Numerous studies have revealed that AC show not only limited responsiveness to others in settings that elicit social referencing (Charman et al 1997; Sigman et al 1992) but also impairment in the assimilations of the stance of the other (Hobson 1993) and reduced one-to-one intersubjective engagement and responsiveness along with impairment in joint attention and other forms of 'secondary intersubjectivity' (Wimpory et al 2000).

In short, AC have but partial and fragile capacities (a) to perceive expressions of emotion as expressive of subjective mental life, (b) to relate to other persons as persons with subjective states of their own and (c) to identify with other people's attitudes, and co-ordinate experiences of self and other (Hobson 2009).

Unsurprisingly, Mindblindness defenders explain these results by arguing that AC display a fundamentally cognitive impairment. But consider the opposite hypothesis, namely that the ability to socially interact and read others' minds might have evolved *first*, with the turning inward of the metarepresentational spotlight upon our own inner states developing only later (Happé 2003). Indeed, as Hobson insightfully argued:

> *"to suppose that children with autism lack cognitive equipment for mentalistic reasoning that is normally 'triggered' by perceptual input, or to conceptualize their difficulties in apprehending and responding to people's feelings in terms of their limited ability to interpret socio-emotional 'cues', is to overintellectualize and over-individualize the socially embedded nature of human psychology"* (Hobson 2009, 255).

In the next section, I argue that our concepts of psychological states are dependent upon our interaction and attitudes towards what people express through their bodies. Then I review experimental studies suggesting that the complexity of interpersonal coordination[5] is already present in early infancy. More specifically, I argue that an impairment of the biologically-based tendencies for interpersonal engagement triggers cognitive dysfunctions and not the other way around.

2. Building the Self through Others: the Interaction Component

Clearly, the story of the development of self-other relations begins in the earliest months of life, when infants adjust to the interactive style of their caregiver (Bigelow 1998), apprehend and copy certain of the facial and manual actions of other people (Field et al 1982; Meltzoff and Moore 1977) and register if someone fails to maintain responsiveness towards them in face-to-face interactions (Cohn and Tronick 1983; Murray and Trevarthen 1985). The centrality of what has come to be termed as "primary intersubjectivity" characterizes the mental experience of infants during infant–caregiver interactions from the earliest phases of life (Braten 1988; Hobson 1993; Stern, 1985; Trevarthen 1979; Trevarthen and Aitken 2001). Given this atunement between a baby's subjective experience of her own as well as of the caregiver's experience during the organized patterns of mother–infant interactions from birth (Rochat and Striano 1999), it has been hypothesized that there is a basic human-specific drive to *share* psychological states with others (Tomasello et al 2005). Trevarthen (2005), for example, showed that the newborn baby is already interested in the attention of a nearby person. Indeed, the infant appears to be endowed with the tendency to make emotionally expressive movements with voice, face and hands which are ideally adapted for interpersonal emotional expression.

Moreover, attention has been paid to the intricate organization of the early bi-directional affective and imitative interaction sequences and their characteristic contingent 'protoconversational' turn-taking structure (Beebe et al 1997; Jaffe et al 2001). Several studies showed innate attentiveness to and preference for the pattern of the human face, a prewired interest in eye contact and an innate propensity to follow gaze shift when this is subsequent to direct eye contact (Farroni et al 2002), the early sensitivity and motivation to explore and analyze the causal contingency structure of interactions (Gergely and Watson 1999) or the implicit understanding of others' pretend actions at 15 months (Onishi and Baillargeon 2005). Later, towards the end of the first year, infants relate to others' actions and attitudes with reference to a shared world, for example by showing things to others and making requests, imitating others' actions on objects, and engaging in social referencing (Bretherton 1992). A further stage that occurs between the middle and the end of the second year is when children show conceptual understanding of self and other, for example in adjusting their actions to the needs and feelings of others, referring to themselves as 'I' and addressing others as 'you', and talking about their own and others' mental states (Bretherton and Beeghly 1982). This latter stage has been typically taken to reflect a child's 'theory of mind', but also the child's increasingly sophisticated concept of 'selves', i.e., of individual persons who have distinctive psychological relations with the social and physical world.

Recently, evidence has been mounting to suggest that infants' understanding of others is more robust within *interactive* contexts (Moll et al 2008; Kiraly 2009). In other words, the more engaged the interactions infants/agents are the more robust the infants' understanding of others becomes. Consistent with a burgeoning literature suggesting a common basis for both the production and perception of action, it has been hypothesized that infants' own actions inform their perceptions of goal-directed behavior (Falck-Yter et al 2006; Kanakogi and Itakura 2010) and thereby contribute to their ability to view others as goal-directed agents[6]. Furthermore, goal-relatedness is differently perceived by infants in social versus physical event configurations (Woodward et al 2001). They regard the direction of the others' eye-gaze as being referentially significant and hence intentional. Studies using looking-time measures have revealed a difference between understanding passive visual (looker-object) versus active, manual (actor-object) relations in infants (Woodward 1998). Similarly, fMRI evidence is accumulating that infants' brain organization in typically developing children may well adapted to be an "intersubjective system" (Aitken and Trevarthen 1997; Tzourio-Mazoyer et al 2002). Moreover, several developmental studies seem to suggest that higher level cognitive processes are strongly affected by the history of social interactions (Jaffe et al 2001; Threvarthen 2005, 2010).

The idea behind these studies is that the environment is never for a living system a neutral world waiting to be internally represented and evaluated in order to become meaningful. In so many ways, the ongoing online processes are already meaningful, without the need of adding an explicit conceptual representation of it. These findings suggest that infants' capacity to understand other minds is rooted in their capacity to actively *engage* in intersubjective scenarios. In other words, there is an irreducible embodied aspect to such interactions (De Jaegher and Di Paolo 2007). The "interaction component" thesis stipulates that interaction *constitutes* (not just causally contributes to) intersubjective understanding.

Indeed, by the end of first year of life, infants have a non-mentalizing perception-based, embodied and pragmatic understanding of the emotions and intentions of other persons (Gallagher 2011). Moreover, typically developing newborns show rudimentary forms of social mimicry (Meltzoff and Moore 1977), which expands (develops) considerably across the first two years of life (Masur and Ritz 1984; Piaget 1962). The importance of imitation skills to early peer engagement has been established for typically developing toddlers (Nadel and Peze 1993). The imitative process begins with imitation of facial movements, available at birth, and develops to include affectively related body movements and mirroring of facial expressions, body postures, gestures and tone of voice. Importantly, imitation difficulties seem to be a core part of the primary symptom set associated with early autism. Several

studies demonstrated that social impairments in autism involve dyadic as well as triadic engagement (Dawson et al 1998; Hobson 1993). Crucially, unlike motor stereotypes and repetitive behaviour, the lack of social and emotional mirroring is unique to the disorder (Charman et al 1997; Stone et al. 1997).

In the next section I introduce the Relational-Self Impairment (RSI) hypothesis and review a mix of behavioral and neural studies supporting the idea that the lack of mentalizing abilities may be the effect and not the cause of intersubjective connectedness impairments.

3 The Relational-Self Impairment Hypothesis

Recently, Lombardo and Baron-Cohen (2010) acknowledged that Mindreading researches have left a *gap* in terms of devices that may be responsible for atypical self-referential processes in autistic children and their integration into the bigger picture of how individuals navigate and interact with the social world. Consequently, they argue that what is needed is a more *relational* approach in order to reveal the deeper complexities involved in interpersonal relations. While I agree with Lombardo and Baron-Cohen on this final note, I disagree with their proposal to replace the search for *individual mechanisms* by a search of *relations between mechanisms*. In this section I argue that a good way to bridge the above mentioned gap is to examine whether the aspects of social-cognitive impairments that extend beyond the autistic children's limitations to form representations are foundational in nature.

More specifically, building upon Neisser's (1991) distinction between the *interpersonal* self[7] and the *ecological* self[8] and starting with premise that the early manifestation of a sense of self in the physical and social domain is not a given but rather "develops via the active process of intermodal perception and exploration" (Rochat and Striano 2000, 515), the idea is to argue that failures in emotional engagement and perspective-taking may be *primary* in development and *cause* children's representational limitations. Indeed, one could briefly define an impairment of the Relational-Self as a dissociation between autistic children's self-consciousness in being observed and their ability to be *affected by* and *engaged with* the attitudes of a particular embodied person. Against theories focusing upon cognitive aspects of autism, Peter Hobson (1989, 1993) has long championed a view, which insists on the idea of a profound disruption of patterned intersubjective engagement between the child and other is basic in autism. Furthermore, several theorists proposed that emotion deficits are primary to autism, producing other social and cognitive dysfunctions (Mundy and Sigman 1989; Rogers and Pennington 1991). In the remainder of this section, I present and discuss several findings that are in line with the RSI hypothesis.

At the *behavioral* level, Klin and colleagues (1992) found deficits in early social skills such as reaching in expectation of being picked up: these impairments in primitive social abilities are not easily explained as the result of impaired cognitive abilities. Furthermore, it seems that a deficit at the level of interpersonal connectedness could prevent AC from registering and assimilating the bodily-anchored psychological stance of another person as suggested by a recent study by Hobson and colleagues (2009) in which AC witnessed one adult tearing another (nonresponsive) adult's drawing. In contrast to children without autism, who expressed dismay, questioned the perpetrator, and showed concern towards the victim, most children with autism showed very little indication of *feeling for* the person whose drawing was destroyed. The victim had shown no overt expression, yet the children without autism immediately orientated towards, and showed concern for, this person. Similar findings reveal that AC display limited propensity to register and assimilate the bodily-expressed attitudes of others in such a way as to apprehend and relate to the world through other people's minds and emotions. Also they exhibit limitations in perceiving and/or adjusting to the perspectives of different figures within a story, and in shifting from one-person anchored perspective to another (Hobson and Meyer 2005).

Moreover, several studies illustrated that facial imitation[9] is tied to interpersonal social engagement, including communication and mirroring of emotional expressions and imitation of sounds. Evidence is accumulating that people mimic others' emotional facial expressions (Bush et al 1989; Dimberg 1982). Importantly, numerous studies reveal that autistic children are significantly more impaired in overall imitation abilities than typically developing children. Stone et al (1997) found dissociation between imitation of actions and objects and imitation of body/facial actions in young children with autism. De Myer et al 1972 and Stone et al 1999 reported that imitations of functional actions on objects were less impaired than imitation of body-movements without objects. Also Hobson and Lee (1999) found significant impairments in imitating the emotional quality of the experimenter's movement.

Crucially, not all areas of emotional functioning are impaired. Indeed, autistic individuals can express a wide range of emotions and show attachment behaviors (Sigman et al 1992). But the interesting point is that autistic individuals reveal failures in *automatic* affective processes, such as those involved in creating emotional reciprocity (Hobson 1993, Kasari et al 1993, Rogers and Pennington 1991). Consequently, McIntosh and colleagues (2006) investigated automatic facial mimicry – a process involved in rapid emotional communication by using the psychophysiological method of facial electromyography (EMG)[10] (Cacioppo and

Berntson 1992). The main hypothesis to be tested was that imitation in autism involves two different processes:

(a) An "affective mechanism" available at birth and modulating social exchanges;

(b) A more "executively constructed, cognitively mediated, intentional imitation system that allows one to learn from instrumental means-end relations from others" (Rogers et al 2003, 777)

Hence, they compared automatic[11] and voluntary mimicry of emotional facial expressions in autistic individuals and a matched comparison sample of typical individuals and found that autistic participants did not automatically mimic facial expressions, but they were equally successful as typical participants on voluntary mimicry. This suggests that their absence of automatic mimicry was not due to deficits in perception, praxis, motivation or task understanding.

Consequently, it has been argued that imitation performance on complex goal-oriented tasks tells us little about automatic processes that might contribute to rapid sharing of affective, more basic states. Indeed, given that autistic individuals perform generally better in imitating actions on objects, it has been suggested that perhaps autistic individuals make use of the second mechanism (b), but without the benefit of the first (a). Hence, their imitation is more effortful and unaccompanied by the social / emotional markers of pleasure and intimacy that typically developing children demonstrate as they imitate another person's movements. This model predicts a gradual development of imitation abilities in autism as the second imitative system, more cognitively effortful matures.

In another study designed to test the children's abilities to recognize emotions, Hobson and colleagues (1988) presented an array of four emotionally expressive photographs of different individuals' faces. Then they gave the children a pile of full-face photographs to sort, with instructions either to sort by the people's identities, or by their emotional expressions. The experimenters selected photographs so that the task was either to match emotions across different individuals, or to match photographs of the same individual when that person's emotion differed. When they did this with the mouths and brows of the faces blanked-out, so that only the eyes were available as a basis for sorting, children without autism were much more able than those with autism to sort emotional expressions. When the experimenter used full faces- but with the array and the faces to be sorted upside-down- the task was beyond the participants without autism. Yet the performance of the children with autism substantially improved over sorting part-faces, and they scored significantly higher than the comparison children in sorting upside-down emotions.

At the neural level, Hall and colleagues (2003) investigated whether attenuation in neural activation to facial stimuli is present in individuals with autism (adolescents and adults) when the salience of emotional cues is increased and additional prosodic information is provided. They found that although participants with autism performed as well as the comparison group on the gender-recognition task; also they responded as quickly, but made significantly more errors than the comparison subjects during the emotion-recognition task. They found that high-functioning adults with autism, when recognizing emotion, place less processing emphasis on the extraction of facial information. They also found that in emotion processing, adults with autism showed greater thalamic activation which is consistent with the suggestion that autistic individuals process facial stimuli trough a selective analysis of features rather that holistically. They used cross-modal emotional stimuli and suggested that autistic individuals may process these cross-modal stimuli as competing rather than complementary sensory experiences. They draw the conclusion that autistic individuals show difficulties in recognizing and understanding emotions because they rely on prototypical representations of emotions and use categorical knowledge to solve novel problems of emotional experiences. In short, autistic individuals draw on categorical knowledge to interpret emotional signals.

To summarize: individuals with autism are notably impaired in both (1) the recognition of emotional prosody and (2) the perception of facial emotion. In order to compensate, they use effortful cognitive strategies based on learned associations and prototypical references to label emotional expressions. Indeed, for many emotion tasks, autistic individuals, given sufficient time, might be able to use non-affective compensatory strategies to accomplish the task. But these strategies are costly. Moreover, this could explain why AC manage eventually to acquire language skills while failing to show any propensity whatsoever in engaging in an emotional and psychological attunement with other people. Hence, lack of *communication* and *participatory sense-making* (De Jaegher and Di Paolo 2007) rather than lack of language or mindreading abilities seem to characterize autism.

Conclusion
I argued here that impairments of the relational self in autistic individuals might induce important mentalizing abilities deficits. In other words, modes of social relatedness may be primary over modes of cognitive abilities. I reviewed several empirical studies emphasizing both the constitutive role of interpersonal relatedness in self-awareness and the basic human-specific drive to share psychological states with others. Then I presented recent empirical work suggesting that the fundamental social process of automatic mimicry may be impaired in disorders such as autism. These findings, I claimed, support the Relational Self

Impairment hypothesis according to which the lack of metarepresentational abilities may be the *effect* and *not* the cause of socio-emotional and communication deficits. If this line of reasoning is correct, then the case of autism illuminates both the perceptual-relational underpinnings of human beings' grasp of psychological states concepts, and the importance of a shared form of life for successful mentalizing functioning. Indeed, adopting this view has profound implications for our understanding of minds not as mechanisms situated in individuals' heads, but rather as *constitutively* dependent upon worldly and social interactions.

Endnotes

[1] Autism is a developmental disorder with a genetic basis and a prevalence of 0.1 to 0.6%. It is diagnosed on the basis of early emerging qualitative abnormalities in social interaction, communication and imagination. Given the limited space available here, I shall use the term "autism" as an umbrella term for "Autistic Spectrum Disorders".

[2] Cognitivism is roughly the view that that our cognitive capacities should be understood in terms of computational procedures operating on symbolic, internal mental states, and thus, that cognitive science should be focused on studying these internal states and processes.

[3] I will use the term "representation" to mean *mental representation* as defined in Von Eckardt's *MIT Encyclopedia of the Cognitive Science*s entry. According to Von Eckardt, a (mental) representation has four important aspects: "(1) it is realized by a representation bearer; (2) it has content or represents one or more objects; (3) its representation relations are somehow 'grounded'; (4) it can be interpreted by (will serve as a representation for) some interpreter" (1999, 527).

[4] A metarepresentation is a "representation of a representation" or as it is sometimes labeled, a "higher-order representation".

[5] Note that the notion of "coordination" and other dynamical systems conceptual tools are not mysterious. Not only the interaction dynamics can be mathematically expressed and empirically measured (Kelso 2009a) but also more importantly, it can be related to neural activity. Indeed, the field of second-person neuroscience is growing (Schilbach et al 2013).

[6] For example, Sommerville and colleagues (2005, 2008) demonstrated that 3-month olds focus on the relation between an agent and her goal if they reached for (and not merely looked at) a toy before observing another actor grasping it. Also 10-month olds who received active training in pulling a cane to retrieve a toy subsequently registered another person's cane-

pulling action as goal-directed behavior, while infants who underwent mere observational training were unable to do this.

[7] The interpersonal self grows out of the infant's transactions with others, in particular the developing sense of shared experience and reciprocity.

[8] The ecological self is the sense infants develop of their own physical body as a differentiated and situated agent in the environment.

[9] Niedenthal et al (2001) illustrated that this mimicry plays a causal role in the processing of emotional expression. Participants watched one facial expression morph into another and had to detect when the expression changed. Some participants were free to mimic, whereas others were prevented from mimicking by holding a pencil laterally between their lips and teeth. Consistent with the embodiment hypothesis, participants free to mimic the expressions detected the change in emotional expression earlier (more efficiently) for any facial expression than did participants who were prevented from mimicking the expressions.

[10] EMG monitors electrical charges in muscle activity over the cheek and brow region.

[11] Automatic Mimicry occurs without external prompting and involves mere replication of a model's actions, without any insight into why those actions are affective (Want and Harris 2002).

References

Aitken, K., Trevarthen, C. (1997). "Self–other organization in human psychological development." *Development and Psychopathology 9*, 651–675

Baron-Cohen, S. et al (1985). "Does the autistic child have a ''theory of mind''?" *Cognition 21*, 37–46

Beebe, B. et al (1997). "Mother– infant interaction structures and presymbolic self and object representations." *Psychoanalytic Dialogues 7*, 113–182

Bermúdez, J.(2003). *Thinking Without Words*. New York. Oxford University Press

Bigelow, A. (1998). "Infants' sensitivity to familiar imperfect social contingencies in social interaction." *Infant Behavior and Development 21*, 149–161

Braten, S. (1988). "Dialogic mind: The infant and the adult in protoconversation." In: M. Carvallo (Ed.). *Nature, cognition and system*. Vol. I. Dordrecht: Kluwer Academic Publishers, 187–205

Bretherton, I. (1992). "Social referencing, intentional communication, and the interfacing of minds in infancy." In: S. Feinman (ed.). *Social referencing and the social construction of reality in infancy*. New York: Plenum, 57–77

Bretherton, I. Beeghly, M. (1982). "Talking about internal states: the acquisition of an explicit theory of mind." *Developmental Psychology 18*, 906–921

Bush, L., et al (1989). "The effects of sex and prior attitude on emotional reactions to expressive displays of political leaders." *Psychophysiology 23*, 427

Cacioppo, J., Berntson, G. (1992). "Social psychological contributions to the decade of the brain: doctrine of multilevel analysis." *American Psychologist 47*, 1019–1028

Carruthers, P. (2009). "How we know our own minds: The relationship between mindreading and metacognition". *Behavioral and Brain Sciences 32*, 121–138

Charman, T. (1997). "Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation." *Developmental Psychology 33*, 781–789

Chemero, A. (2009). *Radical Embodied Cognition*. Cambridge, MA: MIT Press

Cohn, J., and Tronick, E. (1983). "Three-month-old infants' reaction to simulated maternal depression." *Child Development 54*, 185–193

Currie, G., Ravenscroft, I. (2002). *Recreative minds: imagination in philosophy and psychology*. Oxford: Oxford University Press

Davies, M., Stone, T. (1995b). *Mental simulation: evaluations and applications*. Oxford: Blackwell.

Dawson, G., et al (1998). "Children with autism fail to orient to naturally occurring social stimuli." *Journal of Autism and Developmental Disorders 28*, 479–485

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

DeMyer, M., et al (1972). "Imitation in autistic, early schizophrenic, and nonpsychotic subnormal children." *Journal of Autism and Childhood Schizophrenia 2*, 264–287

Dimberg, U. (1982). "Facial reactions to facial expressions." *Psychophysiology 19*, 643–647

Ekman, P. et al (1972). *Emotion in the human face*. New York: Pergamon Press

Falck-Ytter, T. et al (2006). "Infants predict other peoples' action goals." *Nature Neuroscience 9*, 878-879

Farroni, T., et al (2002). "Eye contact detection in humans from birth." *Proceedings of the National Academy of Sciences of the USA 99*, 9602–9605

Field, T. et al (1982). "Discrimination and imitation of facial expressions by neonates." *Science 218*, 179–181"

Frith, U. (1989). "Autism and "Theory of Mind"." In: C. Gillberg (Ed.). *Diagnosis and Treatment of Autism*. New York: Plenum Press, 33-52

Gallagher, S. (2005). *How the body shapes the mind*. USA: Oxford University Press

Gallagher, S. (2011) "In Defense of Phenomenological Approaches to Social Cognition: Interacting with the Critics", *Review of Philosophy and Psychology*, Volume 3, 187-212

Gergely, G., and Watson, J. (1999). "Early social-emotional development: Contingency perception and the social biofeedback model." In: P. Rochat (Ed.), *Early social cognition: Understanding others in the first months of life*. Hillsdale, NJ: Erlbaum, 101–137

Goldman, A. (2006). *Simulating minds: the philosophy, psychology, and neuroscience of mindreading (philosophy of mind)*. USA: Oxford University Press

Hall, G., et al (2003). "Enhanced salience and emotion recognition in autism: a PET study." *American Journal of Psychiatry 160*, 1439–1441

Happé, F. (1994). *Autism: An Introduction To Psychological Theory.* Hove: Psychology Press

Happé, F. (2003). "Cognition in autism: one deficit or many?" *Novartis Foundation Symposium 251*, 198-207

Hobson, R. (1993). "The emotional origins of social understanding." *Philosophical Psychology 6*, 227–249

___ (2009) "Wittgenstein and the developmental psychopathology of autism". *New Ideas in Psychology* 27 (2), 243 - 257

___ (2012) "Emotion as personal relatedness." *Emotion Review 4 (2)*, 169-175

___, Lee, A. (1999). "Imitation and identification in autism". *Journal of Child Psychology and Psychiatry 40*, 649–660

___, Meyer, J. (2005). "Foundations for self and other: A study in autism". *Developmental Science 8*, 481–491

___, et al (2006). "Foundations for self-awareness: An exploration through autism." *Monographs of the Society for Research in Child Development 284*, 1–165

Hutto, D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*, Cambridge, Mass.: MIT Press

Hutto, D. and Myin, E. (2013), *Radicalizing Enactivism*. Cambridge, MA: MIT Press

Izard, C. (1977). *Human Emotions.* New York: Plenum Press

Jaffe J, et al (2001). "Rhythms of dialogue in infancy: coordinated timing in development." *Monographs of the Society for Research in Child Development 66 (2)*, i-viii; 1-132

Kanakogi, Y., and Itakura, S. (2010). "The link between perception and action in early infancy: From the viewpoint of the direct-matching hypothesis." *Japanese Psychological Research 52(2)*, 121–131

Kanner, L. (1943). "Autistic disturbances of affective contact." *Nervous Child 2*, 217–250

Kasari, C., et al (1993). "Affective development and communication in children with autism." In A. Kaiser and D. Gray (Eds.). *Enhancing children's communication: Research foundation for early language interventions*. New York: Paul H. Brooks, 201–222

Kelso, J. (2009b). "Coordination dynamics." In: R.A. Meyers (ed). *Encyclopedia of Complexity and System Science.* Heidelberg: Springer, 1537–1564

Király, I. (2009). "Memories for events in infants: Goal-relevant action coding." In: T. Striano and V. Reid (eds.). *Social Cognition: Development, Neuroscience and Autism.* Oxford: Wiley-Blackwell

Klin, A., et al (1992). "Autistic social dysfunction: Some limitations of the theory of mind hypothesis". *Journal of Child Psychology and Psychiatry 33*, 861–876

Krueger, J. (2010). "Comment: Radical Enactivism and Inter-Corporeal Affectivity"

In: T. Fuchs, et al (eds). *The Embodied Self.* Stuttgart: Schattauer

Lakoff, G., Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books

Leslie, A. (1994). "ToMM, ToBy, and Agency: Core architecture and domain specificity". In: L. Hirschfeld S. Gelman (Eds.). *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge University Press, 119–148

Leslie, A. (2000) "'Theory of mind' as a mechanism of selective attention." In: M. Gazzaniga (ed.). *The New Cognitive Neurosciences*. 2nd Edition. Cambridge, MA: MIT Press. 1235–1247

Lombardo, M., Baron-Cohen, S. (2010). "The role of the self in Mindblindness in autism". *Consciousness and Cognition* (2010); doi:10.1016/j.concog.2010.09.006

Masur, E., Ritz, E. (1984). "Patterns of gestural, vocal, and verbal imitation performance in infancy." *Merrill-Palmer Quarterly* 30, 369–392

McIntosh, D. et al (2006). "When the social mirror breaks: Deficits in automatic, but not voluntary mimicry of emotional facial expressions in autism." *Developmental Science* 9, 295–302

Meltzoff, A., Moore, M. (1977). "Imitation of facial and manual gestures by human neonates." *Science* 198, 75–78

Meltzoff, A., Gopnik, A. (1993). "The role of imitation in understanding persons and developing a theory of mind". In: S. Baron-Cohen, et al (eds.). *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press, 335-365

Menary R. (2007). *Cognitive Integration: Mind and Cognition Unbounded*. Palgrave

Moll, H. et al (2008). "Fourteen-month-olds know what 'we' have shared in a special way." *Infancy 13*(1), 90-101

Murray, L., and Trevarthen, C. (1985). "Emotional regulation of interactions between two-month-olds and their mothers."

In: T. Field and N. Fox (Eds.). *Social perception in infants*. Norwood, NJ: Ablex, 177–197

Mundy, P., Sigman, M. (1989). "The theoretical implications of joint attention deficits in autism." *Development and Psychopathology 1*, 173–183

Nadel, J., Peze, A. (1993). "What makes immediate imitation communicative in toddlers and autistic children?" In: J. Nadel and L. Camaioni (Eds.). *New perspectives in early communication development*. London: Routledge, 139–156

Neisser, U. (1988). "Five kinds of self-knowledge," *Philosophical Psychology*, 1: 35–59

Niedenthal, P. et al (2001). "When did her smile drop? Facial mimicry and the influences of emotional state on the detection of change in emotional expression." *Cognition and Emotion* 15, 853–864

Onishi, K., Baillargeon, R. (2005). "Do 15-Month-Old Infants Understand False Beliefs?" *Science, 308(8),* 255–258

Overgaard, S. (2007). *Wittgenstein and Other Minds: Rethinking Subjectivity and Intersubjectivity with Wittgenstein, Levinas, and Husserl*. London: Routledge

Ozonoff, S., et al (1991). "Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind". *Journal of Child Psychology and Psychiatry and Allied Disciplines 32*, 1081-1105

Piaget, J. (1962). *Play, dreams, and imitation in childhood.* New York: Norton

Ratcliffe, M. (2007). *Rethinking Commonsense Psychology*. London: Palgrave MacMillan.

Rochat, P., Striano, T. (1999). "Social-cognitive development in the first year." In: P. Rochat (Ed.). *Early social cognition.* Mahwah, NJ: Lawrence Erlbaum, 3–34

_____ (2000). "Perceived Self in Infancy", *Infant Behavior and Development 23 (3-4)*, 513–530

Rogers, S., Pennington, B. (1991). "A theoretical approach to the deficits in infantile autism." *Development and Psychopathology 3*, 137–162

Rogers, S., et al (2003). "Imitation performance in toddlers with autism and those with other developmental disorders." *Journal of Child Psychology and Psychiatry 44, 763–781*

Russell, J. (ed.) (1998). *Autism as an Executive Disorder*. Oxford: Oxford University Press

Schilbach, L. et al (2013).Towards a second-person neuroscience. *Behavioral Brain Science.36 (4)*, 393-414

Sigman, M. et al (1992). "Responses to the negative emotions of others by autistic, mentally retarded, and normal children". *Child Development 63*, 796–807

Sommerville, J. et al (2005). "Action experience alters 3-month-old infants' perception of others' actions." *Cognition 96*, B1–B11

Stern, D. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. New York: Basic Books

Stone, W. et al. (1997). "Motor imitation in young children with autism: What's the object?" *Journal of Abnormal Child Psychology* 25, 475–485

Stone, W. et al (1999). "Can autism be diagnosed accurately in children under three years?" *Journal of Child Psychology and Psychiatry 40*, 219–226

Tomasello, M., et al (2005). "Understanding and sharing intentions: The origins of cultural cognition." *Behavioral and Brain Sciences* 28, 675–691; discussion 691– 735

Thompson, E. (2001). "Empathy and Consciousness." *Journal of Consciousness Studies 8(5-7),* 1-32

Trevarthen, C. (1979). "Communication and cooperation in early infancy: a description of primary intersubjectivity." In: M. Bullowa (ed). *Before speech*. Cambridge: Cambridge University Press, 321–347

___ (2005). "Stepping away from the mirror: Pride and shame in adventures of companionship – reflections on the nature and emotional needs of infant intersubjectivity." In C. Carter et al (eds.). *Attachment and bonding: A new synthesis (*Dahlem Workshop Report 92). Cambridge, MA: MIT Press, 55–84

___ (2010). "What is it *like to be a person* who knows nothing? Defining the active intersubjective mind of a newborn human being." *Infant and Child Development 12/2010; 20(1),* 119 – 135

___, & Aitken, K.J. (2001). "Infant intersubjectivity: Research, theory, and clinical applications." *Journal of Child Psychology and Psychiatry 42*, 3–48

Tzourio-Mazoyer, N. et al (2002). "Neural correlates of woman face processing by 2-month-old infants." *NeuroImage* 15, 454–461

Von Eckardt, B. (1999). "Mental representation". In: R. Wilson and F. Keil. *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, Massachusetts: MIT Press, 527-529

Want, S., and Harris, P. (2002). "How do children ape? Applying concepts from the study of non-human primates to the developmental study of 'imitation' in children." *Developmental Science* 5, 1–13

Wimpory, D. et al (2000). "Are infants with autism socially engaged? A study of recent retrospective parental reports." *Journal of Autism and Developmental Disorders 30*, 525–536

Woodward, A. et al (2001). "How infants make sense of intentional action." In: B. Malle, et al (Eds.), *Intentions and intentionality: Foundations of social cognition*. Cambridge, MA: MIT Press

Woodward, A. (1998). "Infants Selectively Encode the Goal Object of an Actor's Reach." *Cognition 69*, 1–34

Zahavi, D. (2005). *Subjectivity and Selfhood: Investigating the first-person perspective*. Cambridge, MA: The MIT Press

___ (2007). "Expression and empathy." In D. Hutto and M. Ratcliffe (Eds.). *Folk Psychology Re-Assessed.* Dordrecht: Springer, 25-40

The editors:

Alexander Gerner is a theatre director, playright and researcher in philosophy based in Lisbon, Portugal and a full member of the CFCUL at the Universidade de Lisboa. Currently he is a Post-Doc Reasearcher in Philosophy at the CFCUL working on "Philosophy of Cognitive Enhancement" http://cognitiveenhancement.weebly.com/ inside Philosophy of Human Technology at the Centre of Philosophy of Science at the Science Faculty of the University of Lisbon. In 2012 Alexander Gerner received his PhD degree in History and Philosophy of Science at the University of Lisbon on "Philosophical Investigations of Attention". (fc. Springer =SAPERE). With Olga Pombo he co-edited books on Diagrammatology and Diagram Praxis (2010) and on Peirce's concept of Abduction and the Process of Scientific Discovery (2007). ASASE is the first of two books on the cognitive and social foundation of the self Contact: amgerner@fc.ul.pt

Jorge de Almeida Gonçalves was born in Lisbon, Portugal, on the 25th of April 1960. He graduated in Psychology (1988) and Philosophy (1997), both in Lisbon. He has a Masters degree in Philosophy (2002) and a PhD (2007) also in Philosophy. Between 1988 and 1999 he worked in Clinical Psychology. He currently works at Institute of Philosophy of Language (New University of Lisbon) and his current research interests are Consciousness and Self Studies, Philosophy of Psychology, Philosophy of Psychiatry, and Philosophy of Film.
Contact: jorgalvesenator@gmail.com