



**DET PSYKOLOGISKE FAKULTET**

***Implicit Mentalization as Therapeutic Skill.***  
***Using Micro-phenomenology to Access Implicit Knowledge***

HOVEDOPPGAVE  
*profesjonsstudiet i psykologi*

Ester Camilla Blikstad

Vår 2022

### Abstract

In mentalization theory and Mentalization Based Treatment (MBT), implicit mentalization is characterized as an intuitive, pre-conscious, procedural and non-verbal understanding. It entails fast and automatic reflexive processing, requiring little or no attention, intention, awareness or effort. This way of making sense of the world and of others is contrasted with explicit mentalization, which is a slower and more reflective process, where we selectively attend to and deliberately try to figure out, rationally and linguistically, what things mean. The quality of the therapist's implicit mentalizing is claimed to be essential for the outcome of psychotherapy. Therapists implicit mentalization can thus be seen as a skill that potentially can be improved. This thesis examines how the framework of mentalization theory conceptualizes implicit mentalization and explores contemporary critiques of the theoretical and philosophical underpinnings of the construct. It is argued that the current understanding of the phenomenon is constrained by cognitivistic and representationalist assumptions about the mind that may make access to the implicit aspects of our understanding difficult. The framework of phenomenology may be better suited to access implicit processes, and the thesis proposes that the method of microphenomenology can be used to address the question of how to improve the implicit mentalizing skills of therapists.

*Key words: implicit mentalization, Theory of Mind (ToM), interaction theory, phenomenology, microphenomenological interview*

### Sammendrag

Innenfor mentaliseringsteori og mentaliseringsbasert terapi (MBT) blir implisitt mentalisering karakterisert som en intuitiv, før-bevisst, prosedural og ikke-verbal forståelse, som en rask og automatisk refleksiv prosessering som krever lite eller ingen oppmerksomhet, intensjon, bevissthet eller anstrengelse. Denne måten å forstå verden og andre på kontrasteres med eksplisitt mentalisering som er en langsommere og mer reflektiv prosess hvor vi bevisst retter vår oppmerksomhet mot å finne ut hva ting betyr, på en språklig og rasjonell måte. Kvaliteten på terapeutens implisitte mentalisering hevdes å være vesentlig for utbyttet av psykoterapi. Terapeutens implisitte mentalisering kan i lys av dette forstås som en ferdighet som kan trenes og forbedres. Denne teksten ser på hvordan implisitt mentalisering er konseptualisert innenfor mentaliseringsteorien og undersøker sider av kritikken som er rettet mot det teoretiske og filosofiske grunnlaget for dette begrepet. Det hevdes at forståelsen av dette begrepet er begrenset av kognitivistiske og representasjonalistiske antakelser om sinnet som gjør det vanskelig å adressere de implisitte aspektene ved vår forståelse. Fenomenologien synes å være et bedre egnet rammeverk for å forstå implisitte prosesser og teksten undersøker hvorvidt mikrofenomenologisk intervju kan benyttes som metode for å adressere spørsmålet om hvordan kvaliteten på terapeuters implisitte mentalisering kan bedres.

*Nøkkelord: implisitt mentalisering, Theory of Mind (ToM), interaksjonsteori, fenomenologi, mikrofenomenologisk intervju*

**Contents**

Introduction ..... 4

Development of the Concept of Mentalization ..... 10

Implicit and Explicit Mentalization..... 13

Theory of Mind (ToM)..... 16

    Theory Theory ..... 16

    Simulation Theory ..... 17

    ToM and Mentalization Theory..... 19

Criticism of ToM..... 21

Interaction Theory ..... 25

Intentionality ..... 28

On Becoming Aware ..... 31

    The Basic Cycle..... 33

        Epoche ..... 33

        Intuitive evidence..... 38

    A Session ..... 42

    The Context ..... 42

Microphenomenology ..... 42

How Can We Improve the Quality of the Therapist's Implicit Mentalizing Skills? ..... 48

    Short-comings Of the ToM-Framework and Some Implications for Mentalization Theory ..... 49

        Theoretical Constraints of the ToM-framework..... 49

    A Potential Role for Phenomenology..... 52

    How Can Microphenomenology Aid in Improving the Therapist's Implicit Mentalizing Skills? ..... 54

References ..... 58

## Introduction

In social encounters the ability to understand each other is of crucial importance. Sometimes we immediately think we know what the other person intends, thinks, or feels. At other times we wonder and try to figure out more deliberately why someone behaved like they did or what someone meant by a certain look. The concept of mentalization is meant to capture the core aspects of social cognition. It is commonly described as the process of understanding oneself and others based on inner mental states (Luyten, Campbell, Allison, & Fonagy, 2020, p. 19) such as personal desires, needs, feelings, beliefs and reasons. Over the last 30 years, the concept has worked its way into the core vocabulary of developmental and clinical psychology. At the same time, it frequently figures in philosophical debates, especially in the area of philosophy of mind, and it is a central concept in cognitive science and neuroscience. In spite of its centrality in various fields, the concept does not have a clear-cut definition everyone will agree upon. It is an ambiguous concept and its connotations have changed considerably during its development. The concept of mentalization in psychologically oriented mentalization theory is today basically construed as a multidimensional concept spanning four polarities: cognitive/affective, external/internal, self/other and explicit/implicit (Fonagy & Luyten, 2009). Cognitive mentalization, understanding cognitive content such as beliefs, is different from affective mentalization, the understanding of desires and emotions. Focusing on internal aspects, directly considering thoughts, feelings, and experiences is different from focusing mentalization on external factors like physical and visible features, behaviors or actions. Mentalization can be directed towards identifying one's own thoughts and feelings, or the mental states of others. And finally, mentalization can be conducted in an explicit way where we deliberately reflect on mental states or implicit in a non-conscious and automatic way (Fonagy & Luyten, 2009). Effective mentalizing is all about balancing these different aspects, taking information based

on all these dimensions into account when making a judgement. Deficits in the ability to mentalize, or imbalances in these dimensions, serve an explanatory role in various different mental health conditions. Different types of psychopathology can be distinguished based on different combinations of impairments along these four dimensions (Luyten et al., 2020). The development of mentalizing is thought to depend on the quality of the social learning environment. Different aspects of this environment have been considered important during the development of the concept of mentalization, but the early attachment relation has all the way been understood as particularly relevant. Mentalization is fundamentally understood as a capacity that develops and grows in relations.

Mentalization Based Treatment (MBT) was originally developed as a method for treating borderline personality disorder, characterized by a mentalization profile with imbalances primarily regarding the implicit, affective and external poles (Bateman & Fonagy, 2010). Important goals in MBT are learning the patient to flexibly shift perspectives between oneself and others, not to jump to conclusions based on what is immediately and implicitly taken to be the case, but rather explicitly consider several aspects of a situation and learning to be able to think and feel at the same time, in other words, balancing the different polarities on the four core dimensions seen as constituting mentalizing. These skills are assumed to lead the patients to be better able to cope with the intense emotions and interpersonal problems people referred to MBT often experience. What enables this shift to happen, and these new skills in the patient to develop, is to be found in the *stance* through which the therapist relates to the patient. Key elements in this stance are the therapist's consistent focus on the mind of the patient and continuous monitoring of and adaptation to the mentalizing capacity he or she displays at a given time, a curiosity and open-mindedness based on an attitude of not-knowing, which requires tolerance for ambiguity and uncertainty (Allen, Fonagy, & Bateman, 2008). The therapist's characteristic stance of simultaneous openness and finely

tuned responses is the way the so-called implicit mentalizing capacities of the therapist comes to expression. This aspect of mentalization is of crucial importance in mentalization theory and MBT. Implicit mentalization is characterized as an intuitive, pre-conscious, procedural and non-verbal understanding (Allen & Fonagy, 2006), a fast and automatic reflexive processing, requiring little or no attention, intention, awareness or effort (Bateman & Fonagy, 2012). This way of making sense of the world and of others is contrasted with explicit mentalization, which is a slower and more reflective process, where we selectively attend to and deliberately try to figure out, rationally and linguistically, what things mean. This explicit-implicit dimension is said to be the most fundamental of the dimensions of mentalization (Bateman & Fonagy, 2012; Liljenfors & Lundh, 2015). We all fluctuate along this dimension due to both individual differences and across situations where various demands are placed on us, but ideally, implicit mentalization enables us to navigate social interactions in a fast and smooth way. The tendency, however, to rely too much on poorly developed implicit mentalization is assumed to be a crucial component of psychopathology. Take patients with borderline personality disorders as an example; "they are often overwhelmed by their emotions (affective pole), make too quick assumptions (implicit pole) and focus on external cues displayed by others perceived as indicating abandonment or rejection" (Volkert, Hauschild, & Taubner, 2019, p. 25). They further tend to cling to such biased implicit judgements in a rigid and inflexible manner. In such instances, their implicit mentalizing can be considered inaccurate and maladaptive. In other words, implicit mentalizing can be both of an adaptive and maladaptive kind. Rapid unarticulated processing is not good or bad in itself, what determines its functional value is the way it aligns with the current context and facilitates communication.

The quality of implicit mentalization thus seem to be a key factor in regard to both patients and therapists. With a therapist that has good implicit mentalization skills the patients

are assumed to be contained in a way that helps them balance the tendency to jump to conclusions based on limited or biased perceptions or reasoning. Instead, they learn to take into consideration other aspects of the situation as well, both through the way the therapist him- or herself is embodying this ability and by encouraging the patient to explore all of the mentalizing dimensions. In turn, this will impact both the explicit and implicit mentalization skills of the patient.

There are numerous citations in the literature on mentalization that clearly points to the therapist's skills in implicit mentalization as an important factor in psychotherapeutic effectiveness.

The therapist's mentalizing *in a way* [my italics] that fosters the patient's mentalizing is seen as a critical facet of the therapeutic relationship and the essence of the mechanism of change (Fonagy & Bateman, 2006, p. 415).

As clinicians, our helpfulness to our patients may have more to do with the quality of implicit mentalization we offer through a general attitude rather than explicit elaboration of specific mental contents as has been frequently suggested in the past (Fonagy, 2003, p. 271).

Thus much psychotherapeutic effectiveness consists of fostering a safe and a secure climate—a largely implicit mentalizing skill (Allen & Fonagy, 2006, p. 19).

Implicit mentalizing is perforce the foundation of any therapeutic work (Allen et al., 2008, p. 167).

Based on statements like these, I take it that the quality of the therapist's implicit mentalizing is understood to be a cornerstone of the effectiveness of psychotherapy. It would thus be of relevance to take a closer look at the implicit aspect of mentalization on the part of the therapist. More specifically, how can the quality of this aspect be developed and refined in order to enhance the quality of therapy?



In beginning to research this topic however, a sort of aporia or confusion became apparent that made me wonder whether the framework of mentalization theory has sufficient resources to address this question in a satisfying way.’’ Concerns have been raised by several authors to the apparent lack of sufficient theoretical and philosophical underpinnings of the theory to deal with the topic of implicit mentalization (Davidsen & Fosgerau, 2015; Froese & Gallagher, 2012; Fuchs & De Jaegher, 2009; K oster, 2017; Liljenfors & Lundh, 2015). The implicit pole of mentalizing is given a kind of primary importance. At the same time the main architects behind mentalization theory clearly express that implicit mentalization is a problematic concept, for instance when they describe the explicit-implicit distinction as having "the most vexing form of heterogeneity in the concept" (Allen & Fonagy, 2006, p. 7) and further characterizing implicit mentalization as "elusive" (p. 10). Placing the most elusive aspect of an already ambiguous concept at the core of mentalization theory and MBT may easily be deemed problematic. The criticism is not, however, directed at the elusive core. Rather, the critics claim that mentalization theorists have a theoretical bias when they rely so heavily on the Theory of Mind (ToM)-paradigm to explain the mechanisms involved in mentalization. ToM is a theoretical position well suited for explaining many of the explicit aspects of mentalizing, but may be less able to account for the process of implicit mentalization (Davidsen & Fosgerau, 2015). To give a preliminary hint of this critique, the ToM-framework can be said to be based, to a large degree, on cognitivistic and representationalist assumptions about the mind embedded in cartesian dualism (to be clarified later), poorly suited to capture the intersubjective and processual aspects of minds and interactions. Aspects considered crucial in the development and performance of implicit mentalization by both mentalization theorists and the critics referred to. There seems to be a theoretical constraint on mentalization theory that in certain ways contradicts their own clinical insights.

In order to get to the implicit aspect of mentalizing, in an attempt to try to make the elusiveness a bit more tangible, the ToM-framework needs to be challenged. The critics referred to above, in their different ways, all point in the direction of alternative conceptualizations of social interaction and intersubjectivity than what has usually been referred to in mentalization theory in order to better understand what is involved in implicit mentalization. Several of them also point to phenomenology as an alternative theoretical framework.

In this thesis I will pursue this criticism of a lack of sufficient theoretical and philosophical underpinnings of mentalization theory to deal properly with the concept of implicit mentalization. I will do this from within the phenomenological tradition. It is my hope that this critical examination can open the space for a fuller dialogue between mentalization theory and phenomenology, specifically with regard to the explication of implicit aspects of our experiences through the so-called micro-phenomenological interview. In doing this I hope to make visible a potential tool or method by which therapists can work on developing and improving the quality of their own implicit mentalizing skills. To do this, we first need to take a closer look at the theoretical and philosophical basis of mentalization theory through an examination of the development of the concept of mentalization and its theoretical underpinnings, especially ToM and the criticism that can be raised against it. By doing this, the alternative phenomenologically inspired approach may be seen as offering important insights. In this attempt at offering insights from an alternative theoretical and philosophical tradition, I will primarily draw on work in classical phenomenology, interaction theory, enactivism and neurophenomenology as well as phenomenologically inspired work of detailing the aspects of becoming aware.

### **Development of the Concept of Mentalization**

From its inception, mentalization theory has been a constantly developing conceptual framework. During this development the concept of mentalization has gradually changed and different theories have been referred to and elaborated on accordingly, in order to attain a theory that can account for both how the capacity for mentalization is developmentally achieved and how it is actually performed. The concept of mentalization can be traced back to French psychoanalytical thinking in the 1960's, and the mentalization model is firmly rooted in the object relations tradition and in attachment theory (Freeman, 2016). From the start, Bowlby's attachment theory has been the main inspiration for the developmental perspective and Theory of Mind (ToM) has served as an overarching theoretical model for the workings of the mind. The current use of the concept in mentalization theory was established by Fonagy in the late 80's and early 90's and it started out with the definition of mentalization as "the capacity to conceive of conscious and unconscious mental states in oneself and others" (Fonagy, 1991, p. 641). Right from the start the concept aimed at capturing the interplay between conscious and unconscious mental states at play in social cognition. In order, however, to deal effectively with the concept, to develop a research paradigm and therapeutic effective interventions based on this concept, there was a tendency in the early development of mentalization theory to focus primarily on the explicit and conscious aspects of mentalization. During the 90's the concept was operationalized as 'reflective function', the capacity to interpret the experience of attachment figures with reference to mental states. Considerable research was conducted on the relationship between a child's attachment style, adverse experiences and parents' reflective function as evidenced through the Adult Attachment Interview (Fonagy et al., 1995). The concepts of mentalization and reflective function were during this period often used interchangeably, and they were conceptualized as unitary processes and as composed of a single mechanism (Duschinsky &

Foster, 2021). The focus during this period clearly implied that the mentalizing process which had been described so far was primarily an explicit, reflective and conscious process.

Moreover, as it was treated as a unitary process, different relevant dimensions in the capacity assessed by the Reflective Functioning Scale was operationalized in a limited and one-dimensional way (Choi-Kain & Gunderson, 2008). Concerns about problems in the construct of mentalization as well as growing awareness of different developmental trajectories to symptoms of mental illness, led Fonagy and Luyten (2009) to describe the four polarities or dimensions of mentalization; implicit/explicit, self/other, internal/external and cognitive/affective now seen as central for the concept. With the dimensional construct of mentalization the scope of the concept was considerably enlarged. The different dimensions were directly linked to neuroscientific findings and mapped onto the developmental trajectories of the human species (Luyten & Fonagy, 2015). Theoretically, mentalization theory was still based on the importance of the primary attachment relation for the development of mentalization and on ToM to explain the mechanisms involved in mentalization. The theoretical foundation was broadened and substantially updated according to recent advancements in neurobiology, genetics, evolutionary science and developmental psychology (Fonagy & Luyten, 2009; Luyten & Fonagy, 2015). In this new conceptualization, the implicit aspect of understanding self and others was given more attention and even considered as the most fundamental pole in the construct (Liljenfors & Lundh, 2015). Even though the importance of the implicit pole of mentalization was now highlighted, theoretical ambiguities on how to conceptualize this pole still existed without being sufficiently addressed. In parallel with the refinement of the dimensional understanding, there has gradually been a tendency as well towards understanding mentalization more as an activity and not a static property of mind (indicated by favoring the wording 'mentalizing' over 'mentalization' (Allen & Fonagy, 2006), and towards the social

and intersubjective aspects of mentalizing (Bateman & Fonagy, 2010). The social and intersubjective aspects are now at the core of the most recent reorientation in mentalization theory. Over the last few years the focus in mentalization theory has shifted from the importance of the primary attachment relation for the development of mentalization to a more socially and contextually oriented conceptualization of mentalization based on epistemic trust and salutogenesis (Luyten et al., 2020). In this new approach the phenomenology of the patient's experience of having his or her personal narrative mirrored in a relevant way opens the "evolutionary highway" to the development or restoration of epistemic trust, an "openness to the reception of social knowledge that is regarded as personally relevant and of generalizable significance" (Fonagy, Luyten, Allison, & Campbell, 2019, p. 95). Epistemic trust, the evolutionary prewired capacity to trust others as sources of social information (Luyten et al., 2020) fosters salutogenesis, the capacity to benefit from positive influences in one's environment, and hence enables social learning. This mechanism of socioecological learning is now regarded as crucial for the development of mentalization; whereas earlier formulations focused on the unique role of dyadic attachment, the current view has evolved to a more comprehensive set of considerations concerning the role of family, peers and broader sociocultural factors in the development of mentalizing, through different communicative processes (Luyten et al., 2020).

Through this historical account of the overall features in the development of the concept of mentalization, I want to draw attention towards two examples I think illustrate a theoretical bias in mentalization theory. According to Duschinsky and Foster (2021), Fonagy and colleagues in the early 90's circumscribed the object of concern in certain respects when they introduced mentalization to describe social cognition with reference to mental states. Their decision to cut the understanding of social and cultural conventions from the term was important in order to limit the scope of causal explanations and theoretical concern. This

omission has however haunted the paradigm over decades, as Fonagy and colleagues recently have acknowledged in their reorientation towards a more sociocultural learning theory centered around relational and contextual conditions for epistemic trust (Duschinsky & Foster, 2021). A theory of social cognition that omits socio-cultural processes from its fundamental level, is problematic. A similar case could be made regarding the conceptualization of implicit mentalization in mentalization theory; by introducing this term without sufficient theoretical elaboration, implicit mentalization remains ambiguous and elusive. Focusing initially on explicit mentalization, the theoretical scaffold built on this ground ends up with a theoretical bias that makes subsequent access to implicit mentalization almost impossible without getting caught up in paradoxical statements. If implicit mentalizing is considered the fundamental pole in the explicit/implicit dimension, a theory primarily elaborated around the explicit pole could be expected to come short in grasping the process of implicit mentalization.

### **Implicit and Explicit Mentalization**

Perhaps the most important of the four distinctions drawn by Fonagy and Luyten (2009) in their dimensional model is the one between automatic/implicit and controlled/explicit mentalizing. This distinction was taken from existing discussions in cognitive science of social cognition. According to Morton and Frith (1995) mentalizing was said to be primarily unconscious or implicit, a property of our cognitive apparatus that comes into action when triggered by particular stimuli, and which “makes sense” of other people’s and our own behaviour fully automatically. This influential equation of implicit, unconscious and automatic mentalizing influenced the way Fonagy and Luyten (2009) introduced automatic mentalizing as a quick, nonconscious, affect-led and unintended process of identifying and making use of knowledge of mental states, running in the background of lived experience without awareness or effort.

The distinction between automatic and controlled mentalizing found in cognitive science was said to match Fonagy and Luyten's mutual commitment to psychoanalysis, and the importance of non-conscious processes, including clinicians' own implicit understandings (Duschinsky & Foster, 2021). However, the way they introduced these terms into mentalization theory has got them into subsequent trouble. Our post-Freudian belief in the unconscious entails the presumption that it can, at least partly, be brought to consciousness through self-conscious reflection or by disciplined procedures like psychoanalysis. In the cognitivist tradition that has dominated cognitive science since its inception in the 50's, the unconscious is understood as mental processes that cannot be brought to consciousness at all. According to Nisbett and Wilson (1977), apparently the most cited paper in cognitive science, we have no introspective access what so ever to mental processes. This strong dichotomy between what is and what is not available to conscious awareness runs in the background of Fonagy and Luyten's conceptualizations. Further, Duschinsky and Foster (2021) in their extensive biography of the development of the concept of mentalization, claim that Fonagy and Luyten's primary characterization of automatic and controlled mentalization in their 2009 paper presumed alignment with non-mentalizing and mentalizing, respectively. If so, this would imply that implicit mentalization is the same as automatic, non-mentalizing processing. That would be paradoxical.

The distinction between implicit and explicit mentalization has been used interchangeably with the distinction between automatic and controlled mentalization. This has led to confusions about how to understand implicit mentalization: sometimes it is treated as unconscious and automatic mentalization, at other times as non- mentalizing. And in their equation of automatic with implicit, and controlled with explicit mentalizing, they also presumed alignment of automatic with fast, and controlled with slow, mentalizing. This conceptualization does not seem to provide room for the implicit, slow processes of 'feeling

things out' that precisely, as Fonagy has acknowledged elsewhere, comprise a large bulk of work in psychotherapy (Duschinsky & Foster, 2021; Fonagy, 2003). Implicit mentalizing can be a potential way to interpersonal understanding, an embodied form of experience that is not just considered a source of error, as is often the case for BPD patients, but rather a source of information of a pre-verbal level of relational interaction. I take it that this is what mentalization theory directs our attention to through the accentuation of the quality of the therapist's implicit mentalization as crucial for therapeutic outcome. This level opens a possible understanding of the therapeutic relation as a relation between subjects that are intertwined through implicit processes, of which it is the aim of therapy to provide meaning to. In my reading of the mentalization literature I take this to be the gist of what they say. If this is right, the relation between the theoretical elaborations and the clinical intuitions concerning implicit mentalization, becomes problematic. According to (Køster, 2017) "implicit mentalization suffers the somewhat unfortunate fate within mentalization theory of being attributed a gross significance in the literature while never receiving a systematic theoretical articulation" (p. 464). This presumed lack of systematic theoretical articulation makes mentalization theory susceptible to certain theoretical constraints. Arguments in line with this has been put forward in several papers (e.g. Daly, 2014; Davidsen & Fosgerau, 2015; De Jaegher, 2018; De Jaegher & Di Paolo, 2007; Duschinsky & Foster, 2021; Froese & Gallagher, 2012; Fuchs & De Jaegher, 2009; Gallagher, 2007; Køster, 2017; Liljenfors & Lundh, 2015). Fonagy and colleagues have recently addressed this problem in a paper looking at a way therapists can learn to better mentalize, and the constraint on MBT by its representational conceptualization is acknowledged (Sharp et al., 2020). This representational constraint can be traced in the reliance on ToM as a conceptual framework for the understanding of social cognition – a feature I now turn to.



### **Theory of Mind (ToM)**

For 30 years, 'Theory of Mind' has been the prevailing paradigm for understanding social cognition in cognitive science. Our capacity to understand each other is explained by reference to theoretical inferences or mental simulation. There has been a longstanding debate in the literature between 'theory theory' accounts and 'simulation theory' accounts, the recent trends now favoring more hybrid accounts. I nevertheless present the original accounts in order to more clearly address the different implications they entail. The same implications apply to the hybrid accounts as well.

#### **Theory Theory**

The 'theory theory' (TT) accounts, though diverse, commonly claim that we understand others by making inferences based on theories about another person. These theories are often referred to as folk psychology or belief-desire psychology. We represent and attribute mental state attitudes to others, we represent and attribute the contents of these mental state attitudes, and we have an understanding of how these attitudes structurally interrelate (Hutto & Ravenscroft, 2021). These theories and the subsequent inferences based on them, often tacit, may involve an innate mechanism for reading other's mind (Baron-Cohen, 1995), often understood as different sub-personal domain-specific 'Theory of Mind Mechanisms' (ToMM) instantiated in a neural module (Leslie, Friedman, & German, 2004). Such mechanisms may be hardwired (Fodor, 1983), or acquired through the interaction of in-built abilities with the social environment during development (Karmiloff-Smith, Klima, Bellugi, Grant, & Baron-Cohen, 1995). However, theories and inferences can also be explained without references to neural modules giving rise to conceptually articulated, propositional representations as in these versions of *modular or nativist TT*. Rather, so-called *scientific TT* sees theories and inferences more as a flexible dynamic modelling of wordly structures that are learned from the social environment by collecting, evaluating and

responding to evidence in a trial-and-error fashion during development, much in the same way a scientist constructs a scientific theory (Gopnik & Meltzoff, 1997). In so-called *model TT*, the brain is understood as instantiating sub-personal predictive processing more or less analogous to what scientists do when making inferences and testing hypotheses. Outside of our awareness, our brains develop generative models that enable them to advance hypotheses—and those hypotheses are further developed, refined, and improved by being tested in our dealings with the world (Hutto & Ravenscroft, 2021).

Whether our theories about other's minds are understood as neural modules, as learned from the social environment or based on subpersonal predictive processing, all TT accounts share the assumption that we take a theoretical, scientific-like stance towards others in our social encounters with them, that involves postulating the existence of mental states in others and using these mental states to explain and predict the other's behavior (Davidsen & Fosgerau, 2015). It involves a spectator's stance towards the workings of the minds of others; a third-person observational setting where the other's mental states are hidden because they are perceptually opaque.

The understanding characterizing all the different TT-accounts can be said to be modelled on the way explicit mentalization operates, with representations, inferences and predictions, although these processes are often attributed to sub-personal mechanisms beyond our awareness. In other words, processes which in mentalization theory would be seen as more implicit aspects of social cognition; automatic, subconscious and fast processing of others mental states, is modelled on the way explicit mentalization functions.

### **Simulation Theory**

The 'simulation theory' accounts (ST) do not refer to the creation of theories about the other's mental states, they rather appeal to one's own mental experience as an inner model for understanding others. ST accounts emphasize a more practical than theoretical understanding

(Gangopadhyay, 2017). The other's mental state, beliefs, desires, needs and emotions are assumed given an inner representational simulation. It is a way of putting oneself in the other's shoes. When thinking about why others behave in certain ways or how they probably will behave, you simulate the other's relevant beliefs and goals, feed these simulated mental states into your decision-making mechanism and let the mechanism produce a simulated decision. This decision is then projected on or attributed to the other (Barlassina & Gordon, 2017). Simulation is often referred to as empathizing, of imaginatively adopting someone else's perspective. It's about re-creating someone else's mental states.

It has become common among ST-theorists to argue for the existence of two types of processes: high-level and low-level simulation processes. High-level simulation processes are cognitive processes that are typically conscious, under voluntary control, and stimulus-independent. They are implemented by the *reuse* of a certain cognitive mechanism and their output states *resemble* the output states generated by the use of the same cognitive mechanism. Further, these processes are under the control of a single cognitive mechanism: imagination. (Barlassina & Gordon, 2017). This kind of processing seem to be more or less what mentalization theory refers to as explicit mentalizing. Low-level simulation processes, are contrary to this, characterized in a way that resembles implicit mentalizing: cognitive processes that are typically unconscious, automatic and stimulus-driven. In the same way as high-level processes, they are implemented by the reuse of certain cognitive mechanisms and their output states resemble states generated by the ordinary use of these mechanisms (Barlassina & Gordon, 2017). Mirroring processes, as described on the basis of the discovery of mirror neurons by neurophysiologist Rizzolatti and colleagues in the 90's, is the prototype of low-level simulation.

In high-level simulation "the attributor creates in herself pretend states intended to match those of the target. In other words, the attributor attempts to put herself in the target's

"mental shoes"" (Goldman, 2005, p. 80). But if we are using our imagination to put ourselves in another's shoe, we apparently already have some idea of what's going on with that person, as Gallagher (2007) points out. The question then is where that knowledge comes from. It is as if the high-level simulation is one step behind. Referring to low-level simulation instead, or seeing them as primary, is a possible way out. How one is to conceive of implicit mechanisms *as simulation* has however been extensively debated and is beyond the scope of this thesis to address. But what is common to all versions of simulation, at an overall level, is that simulation entails using a model to explain how people act, no matter whether this model is used explicitly or implicitly.

### **ToM and Mentalization Theory**

There are numerous examples throughout the literature of how mentalization theory is heavily influenced by the ToM-framework, both the TT- and ST-accounts, right from the start and up until the most recent publications. References to representation, inference, sub-personal ToM-modules or predictive processes, simulation and mirror neurons flourish. It is explicitly stated that "we construe theory of mind as our folk-psychological *conceptual framework* for explaining behavior in terms of mental states; hence the activity of mentalizing *employs* our theory-of-mind framework as well as contributing to its development and refinement (Allen et al., 2008, pp. 48-49). It is said that "[m]entalizing, the representation of our mental states, is the spine of our sense of self and identity" (Bateman & Fonagy, 2010, p. 5), and that our full capacity for mentalization is developmentally achieved, based on early dyadic interactions, and assumed to depend on the development of "a symbolic representational system for mental states" (Fonagy, Gergely, & Target, 2007, p. 289). When it comes to inference, Allen et al. (2008) says that "Accurate mentalizing of others entails two distinct processes: first, inhibiting one's own perspective; second, inferring another person's perspective" (p. 47). The influence of TT shines forth e.g. through Fonagy's

equation of "mentalization proper" with "thinking explicitly about mental states" (Allen & Fonagy, 2006, p. 54) and in the way cognitive vs affective mentalization is distinguished on the basis of neuroscientific mechanisms "the theory of mind mechanism (TOMM) that mediates agent–attitude–proposition (or M-representations) (...) and (b) the empathizing system (TESS) that uses self-affective state-proposition (E-representations)" (Fonagy & Luyten, 2009, p. 1360). The repeated talk of mentalizing as an "imaginative capacity" resonates clearly with the basic tenets of simulation theory (Køster, 2017) as well as the numerous references to mirror-neurons.

The dependence on a ToM-framework is clear in statements like these. With the recent reorientation towards a more socio-cultural learning theory with explicit references to phenomenology (Fonagy et al., 2019) and social interaction (Luyten et al., 2020), this dependence could be supposed to have been considerably weakened. Together with fewer direct references to ToM in recent publications, mentalization theory now seem to take a direction where they try to distance themselves from the theoretical bias they have been criticized for. When for instance Luyten et al. (2020) says that "balanced mentalizing also includes embodied affective features that ground mentalization in an affectively felt reality" (p. 302) this can easily be taken to express exactly what the phenomenological oriented critics referred to earlier point to as central. I see this as a telling example. For, taking the context of this statement into account, it becomes apparent that the theoretical underpinnings of mentalization theory still is the same as earlier. They qualify this statement by citing research that link embodied affectivity and empathy when they say that "the capacity for empathy is underpinned by a more basic "emotional contagion" system" (p. 303). This research (Shamay-Tsoory, Aharon-Peretz, & Perry, 2009), however, relies fundamentally on the ToM-framework, both in its conceptualizations and its methodology. There are several examples of this modus operandi throughout the most recent publications within

mentalization theory. I take this to indicate that the ToM-thorn still can be said to be in the side of mentalization theory. Although the focus and the language has changed considerably, the conceptualizations can still be said to be marked, to a large extent, by the same theoretical and philosophical underpinnings as earlier.

### **Criticism of ToM**

Mainstream theories of social cognition are mainly based on a representationalist view. In general, representationalist approaches assume that the social world or the social other is something pregiven, i.e. they are fully determined in the way they appear to us. Social cognition, then, consists in the internal mapping or modelling of the other's characteristics and actions and, from there, interpreting and giving explanations for their behaviour (Fuchs & De Jaegher, 2009).

Concepts such as theory of mind, simulation or mentalization all have in common that they conceive of social understanding as putting into operation a 'theory' or a 'model' of how people act. Moreover, research into the so-called 'social brain' and the mirror neuron system, also favors a third-person paradigm of social cognition as a passive observation of others' behaviour, based upon an inner modelling process in the individual brain. One could say that the person who perceives another does not actually interact with him or her but deals with internal models or simulations of her actions (Fuchs & De Jaegher, 2009). This is what is meant by representation. Although this may sound like a caricature, it reveals something of the problem. Clearly, no-one would deny that we interact and that interacting is of importance to our social capacities, but in our readiness to accept this 'obvious' fact, there is a danger that the interaction itself is ignored. This is due to the fact that traditional approaches see interacting as that which we do on the basis of inferential or simulative models. Representational approaches have a strong commitment to a Cartesian understanding of the mind as closed off and in need of mediation through inner mental representations (Køster,

2017). It is this Cartesian premise that makes reference to representations, inferences and simulations necessary in order to establish contact between two essentially closed off subjects.

However different the ToM-accounts are, they all embrace certain assumptions. Fuchs and De Jaegher (2009) list four:

1. 'Inner world' hypothesis: Both TT and ST conceive of the mental as an inner realm separated from others by an epistemic gulf that can only be crossed by inference or projection. We are hidden from each other in principle; therefore, we must infer or simulate the other's inner states in order to understand him. (Gallagher, 2010) terms this the mentalizing supposition.

2. Missing interaction: Both TT and ST assume that we primarily observe others from a third-person stance (Fuchs & De Jaegher, 2009), although ST-accounts depend on a first-person process of using the resources of our own minds to simulate theirs (Gallagher, 2010). The research paradigms of ToM focus on one-way, removed social situations and are biased towards localising social cognition in one participant or in his brain.

3. Missing embodiment: Social cognitive science largely assumes a disembodied sender–receiver relation between two Cartesian minds; the body usually functions only as a transmission device. Even though simulation theories increasingly include the body in the modelling of others, they still do not take into account the reciprocity of embodied agents.

4. Missing development: Traditional approaches to social cognition have been criticised for being overly concerned with which capacity follows which in time, without attention for how the different capacities follow from each other (and therefore also remain connected to each other throughout the lifespan). Moreover, the explanation of social cognition by brain modules or mirror systems remains static and unidirectional in that they assume that the brain mechanisms guide development. But there is increasing evidence that

these neuronal systems develop and are continuously modified only through social interaction, particularly in early childhood

These points are also included by Froese and Gallagher (2012) when they say that the different versions of ToM all share the assumptions of methodological individualism and neuro-reductionism. ToM is claimed to assume a specific version of methodological individualism where a person's social cognition is essentially independent of the process of social interaction and therefore exclusively explainable in terms of that individual's capacities alone (e.g. belief-desire inference or pretense). They further assume a specific version of neuro-reductionism where the individual's social cognition is essentially independent of their first-person experience (including embodiment) and is therefore exclusively explainable in terms of subpersonal mechanisms alone (e.g. a TOMM and/or mirror neurons).

Gallagher has extensively criticized the way ST uses the discovery of mirror neurons to support the idea of automatic, non-conscious and sub-personal processing as crucial for social cognition. Gallagher (2007) points out that simulation is a personal-level concept that cannot be legitimately applied to subpersonal processes. In doing this one commits a category mistake, a serious logical fallacy that forces one to reconsider the premises of the theory or the inferences made on their basis.

As a result of the tendency in recent social cognitive science to rely on brain mechanisms such as mirror neurons or other special modules to explain social cognition, intersubjectivity is taken as an inferential or projective process encapsulated in the brain (Fuchs & De Jaegher, 2009). However, such explanations single out one section only of the whole circle of organism–environment interaction. They fail to address social interaction as a *structured* and *structuring process* which in turn influences brain functions. This is not to say that the link between action and perception found in mirror neuron research does not play an important role for social understanding. It does, however, need to be embedded in a context



of embodied and meaningful interactions. If intersubjectivity is regarded as a circular process in which the cogniser constantly influences the other by his actions and vice versa, then cognising and acting are interdependent, and there is no pre-given other. On this condition, the concept of inner mapping or representation is too static to be an adequate description of the process (Fuchs & De Jaegher, 2009). This understanding permeates so-called 4E-approaches to cognitive science, arguing that cognition not solely occur inside the head, but also is embodied, embedded, enacted and/or extended (Newen, De Bruin, & Gallagher, 2018) as well as it permeates the field of non-linear dynamic systems theory.

If this line of criticism of ToM is considered valid, it will have implications for mentalization theory. I suspect the default reaction among mentalization theorists will be to consider this criticism as inadequate when it comes to mentalization theory. And in certain respect I will agree that what is said here not necessarily seem to pinpoint a representation of mentalization theory they will recognize or support. For instance, to claim that mentalization theory is missing interaction or development is absurd and contrary to what they accentuate. In MBT, as in all therapeutic approaches, interaction is key. And the accentuation of implicit mentalization in mentalization theory and MBT is meant, I suppose, to capture key aspects of this interaction. Interaction being secondary would also be hard to swallow given the role of the primary attachment relation in earlier versions of mentalization theory and of sociocultural factors in the latest version of mentalization theory. The point is, however, in relying on the ToM-framework, these implications are part of the picture. If they cannot support them, they should reconsider their reliance on this framework. There is a theoretical tension in mentalization theory, where the implications of the different frameworks they rely on, ends in opposition to each other.

The inconsistency that is here brought to light, may be traced to earlier historical omissions as alluded to earlier, to priorities made early in the development of mentalization

theory. Available theoretical and scientific models will inevitably be steppingstones on the path ahead to a mature theory. There is a clear reorientation going on in mentalization theory today, towards processes several branches of science now to a larger degree address and try to implement, processes such as embodiment, interaction and dynamic couplings between systems. The on-going reorientation will still, however, need to deal more radically with its own premises if it is to escape paradoxical implications.

### **Interaction Theory**

An alternative account of social cognition set forth on the basis of the criticism raised against ToM, is interaction theory. On this view, our primary and pervasive way of encountering others is not characterized by detached observation, but by interaction, characterized in terms of sensory-motor processes (primary intersubjectivity) and context (secondary intersubjectivity). Our primary way of understanding others is worked out in real second-person interactions within pragmatic and social contexts (Gallagher, 2010). A third layer of intersubjectivity gives explicit meaning to our understanding through narrative practices, where the abilities for intersubjective perception and pragmatic interaction frame our everyday interpretations of others (Gallagher, 2007).

Interaction theory draws inspiration from phenomenology which from its inception has been interested in the problem of intersubjectivity. This approach to social cognition insists on the *primacy* of a person's embodied, interactive, and directly perceptual (i.e. not theoretical or explicitly conceptualized) grasp of another's mind (Zahavi, 2011). This is in direct contrast to the original 'problem of other minds', also characterizing mentalization theory, that presupposes the existence of a principally *hidden* mental, interior and private space represented secondarily through different linguistic, corporeal and gestural manifestations. Rather than our minds being opaque and invisible to each other, phenomenology assumes that social cognition has to do with perception and that a human

being in principle is 'visible' to other human beings. (Koubová, 2014). Here, mentalization theorists may object and say that an assumed direct grasp of the other is a substantial part of the pathology involved in borderline personality disorder. The *primacy* of visibility is however understood to potentially be able to integrate invisibility, disruption and misunderstanding (Koubová, 2014). It does not mean that the other is a pre-given entity we have direct access to. Primary visibility involves mutual modulation and coordination between participants (De Jaegher & Di Paolo, 2007), and in interaction, only those aspects that make sense to us as participants, become visible. The other person appears due to our own participation in the emergence and breakdown of joint relational sense-making (Koubová, 2014). Optimally, mutual coordination and modulation allows misunderstandings to be detected and dialectically resolved. When this does not happen, distorted perceptions result as a consequence of the way the emergent sense structure our perceptual field. The emotional dysregulation characterizing borderline personality disorder, is within this framework considered a genuinely interpersonal phenomenon and involves emotions operating upon a disordered world (Ratcliffe & Bortolan, 2020). From the perspective of interaction theory, social understanding first and foremost takes place during our immediate engagement with other people in social interaction, and can in some cases even be constituted by the social interaction process itself (De Jaegher, Di Paolo, & Gallagher, 2010). The underlying processes involved in the intersubjective structuring of our world are the same whether our engagement with others are considered healthy or pathological. This may leave us with an understanding of borderline pathology as having to do with the intersubjective structuring of our world. In every perception something remains hidden or unavailable. There will always be an invisible side that "functions as a subtle source of ungraspable meaning (...) whose function is performative" and which in part constitutes the way we participate in

social interactions (Koubová, 2014). These hidden and performative structures are a main concern in phenomenology, as will be detailed later.

Interaction theory takes interaction as the source of intersubjectivity, not as the end stage of the social cognitive machinery, as in ToM. The interaction process includes several components such as bodily resonance, affect attunement, coordination of gestures, facial and vocal expression (Fuchs & De Jaegher, 2009), components also at play in the moment-to-moment assessment of the patient's mentalizing capacities that is part of the mentalizing stance (Allen et al., 2008). Social cognition is not a solitary task of deciphering or simulating the actions of others, but emerges from the dynamical process of skilfully interacting with them (Fuchs & De Jaegher, 2009). In addition to its roots in phenomenology, interaction theory further draws inspiration from dynamical systems theory - the rigorous scientific approach of observing and describing the coordination processes between intentional and embodied agents or systems. These two sources of inspiration link two sides of the same process – the interaction. By combining systems theory with the experiential phenomenological approach to the same interaction, interaction theory may allow a better grasp of the second-person constitution of intersubjectivity and ‘participatory sense-making’ (Fuchs & De Jaegher, 2009).

By combining phenomenology and dynamic systems theory we have the possibility of an alternative explanatory framework, which avoids the category mistake inherent in ST, by employing the neutral mathematics of dynamical systems theory to account for the experiential and interactive aspects of social cognition. This amounts to a new explanatory tool that can address unanswered problems in social cognition research by putting social cognition back where it belongs: between individuals and not only in their heads (Froese & Gallagher, 2012).

This understanding of social interaction as enabling social cognition via dynamic couplings between autonomous agents and their environment can be traced to the seminal work *The embodied mind: Cognitive science and human experience* by Varela, Thompson, and Rosch (1991). This work has been the initial inspiration for much of the criticism raised against the ToM-paradigm as outlined above. The work of Varela in the areas of neurophenomenology and enactivism has also inspired another strand of research, primarily conducted in French philosophy and psychology, detailing the process of becoming aware and explicating the implicit aspects of our lived experiences (Depraz, Varela, & Vermersch, 2000; Petitmengin, 2006, 2017; Petitmengin, van Beek, Bitbol, Nissou, & Roepstorff, 2019; Vermersch, 2018). This work is firmly rooted in the phenomenological tradition and will be the focus of the remaining part of this thesis and serve as the basis for considerations about improving the therapist's implicit mentalizing skills.

Before we turn to the process of becoming aware, an introduction to intentionality, a core concept in the phenomenological tradition, is necessary.

### **Intentionality**

According to the phenomenological tradition from Husserl, intentionality is the defining aspect of our consciousness (Husserl, 1977). Intentionality here means 'directedness', as our consciousness is always directed towards something or open to the world in a certain way. We imagine, perceive, judge, remember, think or feel, and this activity is always about something. Our intentionality has a double structure; it has to be understood as the unity of what we intend and the way this is intended. The intentional content and the intentional act are correlated. But not in such a way that we can think them apart from each other. We rather have to understand them as two poles of every conscious or intentional experience. In Husserl's terminology these poles are called *noema* and *noesis*. The exploration of this correlational structure of our intentionality, Husserl called static phenomenology. He also

outlined a genetic phenomenology, where the development or genesis of these very structures is thematized. Here, his notion of the life-world comes into play. We always already find ourselves in a world where our life unfolds and where we intend different objects in ways given by our earlier experiences. In the genetic phenomenology a distinction is made between an active and a passive genesis. In the active genesis we play an active role in the constitution of objects, and the results are tools, works of art, scientific theories etc. This active genesis presupposes however a passivity whereby we already are affected. The ways we relate to the things in the world presupposes an openness towards the world. Such openness puts a mark on us at a sensory-motor and affective level: "In passive genesis, the lived body constitutes itself and its surrounding environment through the involuntary formation of habits, motor patterns, associations, dispositions, motivations, emotions and memories" (Thompson, 2007, p. 30). There is a passive genesis in our experiences where intentionality primarily must be understood as an openness to the world through our lived bodies, and where there no longer is meaningful to talk about a subject-object relation in the classical Cartesian sense. The aspects of this passive genesis is the basis of the noetic structures of our experiences, and thus in important ways shape the way things appear or shine forth for us as a result of our lived experiences. In our intentional directedness to the world, a fundamental constitution of meaning takes place, every sensory input is imbued with an affective valor with a significant meaning. All noematic content in our consciousness is shaped by these noetic structures. But in our directedness towards the objects, this constitutive element of our experiences is necessarily lost. Our attention is captured by the noematic content, and the noetic structures disappears from our attention. We are engulfed by the things in our world at the same time as our own transcendental constitution of these very things become invisible to us. The noetic structures more or less disappear, or become transparent, when our focus is caught. But they don't have to. A growing field of research is devoted to exploring the extent to which we can

cultivate our attention to notice these processes. This appears to be highly relevant for an understanding of the relation between implicit and explicit aspects of our understanding of others. We can cultivate our awareness in a way that makes the implicit aspects experientially available in a more tangible way, making implicit mentalization less elusive.

Merleau-Ponty conceptualized intentionality at different levels. He separated an operative intentionality from a more explicit object-directed intentionality. This operative intentionality is an elaboration of the processes involved in Husserl's passive synthesis that partly constitute his genetic phenomenology (Reuter, 1999). The operative or bodily intentionality is conceived as a pre-reflective and procedural experience emerging from a first-person perspective. The subsequent explicit intentionality gives opportunity to make one's own mental processes thematically available, that is to reflect on one's own experiences and their content.

Such an understanding of intentionality may serve as a starting point when we will try to make sense of the claim that implicit mentalization is a skill that can be trained. According to this understanding it is possible to pay attention to, or be aware of, the implicit pre-reflective aspects of experience even though it is outside the focus of attention. Colombetti (2011) makes the point that by directing attention to our bodily feelings, they can appear in different pre-reflective modes. Legrand (2007) also makes the point that it is possible to overcome the 'phenomenal self-forgetting' and bring the subjective character of experience to the focus of attention *without* making it an explicit intentional object. Increased attention to these pre-reflective structures makes the subsequent explicit intentionality directed at these structures, more likely to provide nuanced descriptions of the phenomenal character of experience.

These two forms of intentionality may serve as the basis on which to conceptualize implicit and explicit mentalization in a way that better accounts for the important role assigned to the quality of the therapist's implicit mentalization for therapeutic outcome.

An example of how they can be combined to help people becoming attentive to different kinds of pre-reflective material can be found in Petitmengins elaboration of the micro-phenomenological interview. This approach heavily relies on the descriptions of the process of becoming aware as outlined by Depraz et al. (2000).

### **On Becoming Aware**

In their book *On Becoming Aware*, Depraz et al. (2000) seek to lay the ground for a practical approach to exploring human experience. They seek to describe experience as an activity and investigates this so far as consciousness perceives itself unfolding in an operative and immanent mode, at once habitual and pre-reflective. This endeavor may be relevant in the attempt to grasp the process of implicit mentalization and develop a richer understanding of how therapists may cultivate and improve the quality of their own implicit mentalizing skills. The authors' point of departure is that among all acts of consciousness, there lives, unperceived, a form of pre-reflexivity on the basis of which consciousness is able to perceive its very self at work. This immanent ability, they claim, is habitually ignored or at best practiced unsystematically or blindly. Exploring human experience amounts to developing and cultivating this basic ability. In their book, they explore how we gain access to this pre-reflective and pre-given aspect of our subjectivity in making it conscious as well as exploring other levels of pre-reflective experience that become available when this process is rigorously explored (p. 2).

The phenomenological tradition is an important source and provides an explicit conceptual framework for this approach. At the center is the concept of *reduction* as the method of phenomenology. This is not reductionism in any kind but must be understood in



accordance with what is previously said about intentionality as the description of the passage from focusing on the object to the act through which the object is perceived.

The process of becoming aware is in this perspective understood and examined as a three-fold nested process, unfolding across three different temporal horizons. These three levels of how the process unfold can be linked to different aspects involved in improving the implicit mentalizing skills in therapists. The first layer is here called the basic cycle and can be seen as an outline of the process of implicit mentalizing in the sense that it brings to awareness the pre-reflective aspects of our experiences. The basic cycle is a way of making visible the processes that are usually performed in an automatic way without us noticing how they influence our subsequent judgements about affairs in the world. These processes tend to be relatively quick, at least they are often experienced as a sudden insight. The second layer, here called a session, gives us a procedural description of the unfolding of the process of becoming aware in the sense that what becomes visible through the basic cycle at this stage is expressed and validated. It is an explication of the implicit aspects of understanding. This can be done in various contexts, and in the last part of this thesis, I will focus on the way this can be done in the micro-phenomenological interview. This is a procedure that often lasts about an hour in order to explicate a few seconds of lived experience. The third layer, here called the context, examine the preliminaries and the after-effects of this process of becoming aware and is thus relevant for apprenticeship and for the way a therapist can work on improving skills on a long-term basis.

In what follows I will outline Depraz et als (2000) descriptions of the first phase of the process of becoming aware, the basic cycle, and point to what I see as relevant aspects in mentalization theory. The next two phases of this process, the session and the context, will be exemplified through a presentation of the micro-phenomenological interview.

## The Basic Cycle

The basic cycle consists of two movements, epoche and intuitive evidence, each of them again consisting of several interlocking movements. Epoche is often characterized as the 'bracketing' of our default attitudes about the world, a necessary process in order to attain a 'pure' description of our experiences. The criterion of truth internal to each act in this cycle is intuition or intuitive evidence. These processes can, and in most instances will, unfold without being subjected to explication.

**Epoche.** Epoche involves a *suspension*, a bracketing of our so-called natural attitude. In our everyday dealings with the world, we naturally take lots of assumptions for granted in order to navigate our environment smoothly and effectively. In order to gain access to the way an experience shows up in itself, we have to break with our "realist prejudices" about the relation between appearance and reality. In parallel with this bracketing, we do at the same time have to accomplish two other phases: a *redirection* of attention and an *acceptance* of what shows up. We must redirect our attention from the exterior to the interior. From the default absorption in the world, we must turn the attention inward to our experience. When we do this, we substitute an apperceptive act for perception. Our cognitive apparatus usually makes automatic judgements all the time. The challenge here is to put this process at pause, not doing what we usually do, and perceive without judgement. This is not easy. This process of directing the attention away from the objects of the world towards the way these objects appear in us is caught up with both psychological and pragmatic difficulties. According to Piaget, as cited by Depraz et al. (2000), we are psychologically more attuned to our goals than to our instruments; what he calls positive information or the content of our perception, immediately captures our attention at the expense of negative information or the tools that allows us access to the content. In Jamesian terms, this can be conceptualized as the distinction between center and fringes in our attentional field. Our attention is usually caught

by what happens in the center or in the nucleus, the fringes, the structural relations providing relevance to the attentional focus, remain at the periphery of awareness (Mangan, 2007).

Phenomenologists talk about this using various terms addressing the ways we are captured by the world. Pragmatical difficulties may show up as the risk involved in having to direct attention inwards to aspects one rather wants to avoid or as a diminished sense of control.

The process of directing attention inwards requires a sense of safety and trust in yourself.

Having managed to change the direction of our attention, we then have to change the quality of attention as well. From voluntary turning attention inwards we now have to adopt a stance where we accept and listen. We change our attention from actively looking for something to letting something come to us, to be revealed, i.e., to a more passive and receiving attitude.

This is the phase of letting-go or accepting. This whole process is a reversal of two of our default cognitive processes: a redirection of attention from the external to the internal and a changing of the quality of attention from actively seeking to a passively and acceptingly letting-come. This change from active to passive, means we have to be attentive, but at the same time waiting. What is unfolding or given in this mode is pre-reflective, and the challenge is to balance a sustained attention without immediate fulfillment. We have to provide a focused and open attitude at the same time. It is often experienced as a relative emptiness both in relation to content and to time. In this mode we can't adjust the details of the content that is gradually taking shape, we can just tune and refine our attention.

Paying attention to what might show up (the most strange, i.e., unusual, state) is such a waiting time, which is at once empty and subjectively long, seems to be the primary hindrance to discovering and spontaneously putting the act of becoming aware into practice (Depraz et al., 2000, p. 38).

This description of epoche, the double change of attention and the difficulties inherent in putting this into practice, very much resembles what Allen et al. (2008) describes when they discuss the mentalizing stance of the therapist. In their book *Mentalizing in clinical*

*practice* they put their fingers on a seemingly paradoxical situation; almost all clinicians intuitively know and already use their implicit mentalizing skills in work with their clients, so the authors presumed they would need to devote only a small part of the MBT-training program to develop and cultivate the mentalizing stance. To their initial surprise, they say, "this has become the most important and most difficult skill to get right. And without a mentalizing stance, all other interventions are unlikely to be helpful" (p. 182). The most difficult skill to teach therapists is to adopt and maintain a mentalizing stance. It is through this mentalizing stance the therapist's implicit mentalizing skills come to expression.

I suggest one of the reasons why this appeared as a paradox and a surprise for the authors may have to do with the way mentalization theory initially has conceptualized implicit mentalization. By this I do not at all mean to imply that the authors don't have a deep understanding or practical skills in this area. On the contrary, the authors *do* point to the centrality of implicit mentalization, details various aspects and demonstrate excellent skills themselves in clinical encounters. What I find unfortunate, is that their insights are tied to a theoretical and philosophical perspective that in certain respects constrain them. From a ToM-framework, there seems to be no room for the very space where the epoche unfolds. ToM entails a clear divide between implicit, automatic, unconscious processes and the explicated content we have conscious access to.

The phenomenological tradition with its elaborations on how to conduct the epoche may serve as an alternative to the ToM-framework. Phenomenology may provide a ground from which the mental processes or acts themselves can be thematized rather than the resulting mental content. This focus together with detailed knowledge on how to gain and cultivate access to the enabling processes for our subsequent judgement, may be useful in further elaborating the process of implicit mentalizing in mentalization theory.

There are several challenges one will face in performing the epoche, like fear or uncertainty, boredom and the challenge of accepting silence. It can further be quite frustrating to find out how little control we actually have over our own mental processes. And in turning our attention inwards we also run the risk of suddenly becoming aware of traits or aspects of ourselves that has earlier gone unnoticed. In the redirecting of our attention, we have to inhibit both our outer actions and these kind of internal effects of the re-directing, in order to leave open a place for the subsequent apperception. By turning inwards, we can cultivate and maintain this openness. In letting-go or acceptance, what you have to inhibit is an immediately fulfilling intuition that you might be tempted to bring about by projecting your categories, presumptions, and identifications onto what begins to appear (Depraz et al., 2000). This letting-go is in line with the not-knowing position of MBT (Allen et al., 2008). Accomplishing this is a real challenge. And it is important to note that the skill of letting-go in contemplative traditions throughout the world is considered an advanced skill. Although therapists to a certain extent can be said to already be in possession of the relevant skill, as Allen et al. (2008) notes, it may nevertheless be a bit naïve to expect the refinement of this skill to be relatively easy to accomplish. The challenges involved in this attentional redirection is acknowledged in various traditions that focus on awareness of subjective experience and a substantial amount of practice is devoted to cultivating this skill or attitude. As I see it, there is no disagreement between the mentalization theorists and the authors of *On Becoming Aware* when it comes to the importance of this attitude in the practice of psychotherapy. As Depraz et al. (2000) notes, therapists give their clients all their attention at the same time as they are careful not to interfere with the open and patient acceptance of what the client brings to light by making interior commentaries or counter-transferential movements. The good therapist is able to simultaneously register the verbal, the non-verbal (modification of posture, gestures, breathing rhythm, depth, placement, mimicries, micro-

movements), the epi-verbal (what is said by the manner in which it is said), and the para-verbal (the variations of intonation) by an open listening and observation, without looking to seize hold of anything (Depraz et al., 2000). This is what point to as well, when they list inquisitiveness, curiosity, open-mindedness, uncertainty, not-knowing, interest in understanding better and consistent focus on the mind of the patient among the key elements of the mentalizing stance (Allen et al., 2008). This attitude, stance or skill is considered crucial for all psychotherapy and understood as a common factor for all effective therapy (Wampold, 2013). In therapy with BPD patients there is the additional challenge of maintaining this stance in the relational field with clients who easily experience intense non-mentalizing states. The higher the emotional temperature gets, the more difficult it is to maintain this stance.

Where there may be differences between mentalization theory and the perspective developed by (Depraz et al., 2000) is in how this attitude is developed, cultivated and maintained. For psychoanalysts, to develop this skill you have to go through analysis yourself, it is developed through insight. MBT's break with psychoanalysis partly consist in its accentuation of distortions of mental *processes* rather than mental representations as the core problem, especially for clients with problems in the personality spectrum. In this way, working with insight-oriented interventions with BPD patients, is secondary (Allen et al., 2008). This focus on distorted mental processes rather than content in MBT may contribute to a kind of ambiguity when it comes to how to approach the implicit mentalizing skills in therapists. Therapists are not considered to have distorted mental processes. And there is a legitimate question whether improving this skill in therapists should be addressed in other ways than how it is done with patients. In line with what Allen et al. (2008) say about treatment of choice for patients on the neurotic spectrum rather being insight-oriented interventions because they are assumed to have distorted mental representation as their core

problem, insight-oriented approaches may be beneficial for therapists, as well. Working on insight may improve implicit mentalizing skills in therapists by e.g., identifying triggers and patterns of countertransference. This approach is however largely outside the focus of interest in mentalization theory as a result of their break with the psychoanalytic tradition. This leaves the focus on therapist's implicit mentalizing skill in a kind of vacuum: an assumed unwillingness to go down the psychoanalytic alley on the one hand, on the other, an understanding of implicit mentalizing inscribed in a theory of psychopathology that does not necessarily refer to the challenges of implicit mentalizing therapists face.

For mentalization theory, the phenomenological focus on the processes rather than the content of our consciousness as presented through the epoche, without inscribing this into a theory of psychopathology, may serve as an alternative approach to work on improving implicit mentalizing skills without having to revert to the psychoanalytic tradition they in certain respects have sought to distance themselves from. Through this approach, it seems possible to maintain a focus on processes rather than content in non-pathological cases as well.

**Intuitive evidence.** The process of intuitive evidence or intuition follows immediately after the epoche and is a necessary stage before one can start to express the results of the epoche. It is an intermediate step between the attention that allows something to appear through the suspension of our natural attitude and the process of expressing this in a way that makes it available for others as well. Intuitive evidence is thus a kind of knowledge situated in the tension between the subjective and the intersubjective. The notion is derived from the latin *intueri* which means to look in an active form. As a form of seeing or looking it has nothing to do with symbolic knowledge. It is understood as a direct givenness (Depraz et al., 2000). In the philosophical tradition, intuitive knowledge is extensively discussed, here I will only point to one major class of intuitive acts often highlighted: perception. It is important to

note that the intuitive evidence in the form of perceptual givenness, presupposes the previous reduction, accomplished through the epoche. Without this, the evidence of what gives itself is just what we always and already take for granted in the way we relate to the world. Exactly the natural attitude we are to suspend, to bracket, in order to get to the 'thing itself'. This is a highly relevant point in relation to the quality of our implicit mentalization. In interaction theory or other phenomenologically and enactively oriented approaches to social cognition, there is an insistence on the direct perceptual grasp, the intuitive givenness, of another's mind. From our elaboration of the epoche we now have a framework through which we can address some of the implicit and pre-reflective aspect of this grasping. Direct perceptual givenness does not amount to full transparency, that we have immediate direct perceptual access to the explicit content of the other's mind. According to Koubová (2014), transparency and opacity exist side by side, and shape our social interactions. In our everyday dealings, what normally remains hidden is the way our perceptions are structured, not the 'what', but the 'how'. This point directly to Husserl's passive synthesis or Merleau-Ponty's operative intentionality. We do however have perceptual access to this domain, but this access is via intuitive evidence based on the epoche. The direct perception hypothesis means that our experience of others is not mediated through conscious inferences or simulations, as ToM would have it (Køster, 2017). Our experience of others is directly perceived in the way of intuitive evidence. It is implicit knowledge. This can later be expressed and intersubjectively validated through appropriate methods.

Two important aspects characterize the intuitive act; it is *non-mediated* and *passive* (Depraz et al., 2000). Intuition is non-mediated by any conceptuality. This can be a tricky point, for according to the phenomenological tradition we are always at the same time caught in a hermeneutic circle (Gadamer & Jordheim, 2003), our intuitions are always inscribed in a cultural horizon, always mediated by our means of expression and the context in which the



intuitive act unfolds. In hindsight we can often point to causal relations related to the results of our intuitions. But as a direct lived experience the intuitive act is novel and discontinuous. It is lived as a genuine emergence. It is a creative process to be distinguished from a reflexive activity where every step follows the preceding according to certain laws of e.g., coherence or inference. Intuitive givenness *based on epoche* ruptures the very familiarity of the fabric of the world and lets the content of our consciousness shine forth in a genuine, surprising and non-mediated way (Depraz et al., 2000). Intuition is also passive in a certain sense. It is involuntary in the sense that you cannot bring forth the result in a controlled manner. It is a passive act in that you will have to wait and accept what comes. But what you *can*, and indeed have to do, is to create the conditions for it in the practicing of the *epoche*. Although you cannot anticipate the content of the result, you can anticipate the form or the quality (Depraz et al., 2000). You can make sure nothing is hindering the unfolding fulfillment, create an opening and making sure this opening is not altered. This process of creating and holding space must be learned and cultivated.

The experience of intuitive evidence often strikes us as a sudden clarity, as a eureka or a lightning-bolt. It can be described as an instantaneous breakthrough of a threshold of intensity that fills up, satisfies or saturates us (Depraz et al., 2000). We are however hit by this lightning-bolt only because we have been out in the thunderstorm flying a kite, with a reference to Benjamin Franklin. The sudden insight is long prepared through the initial movements of *epoche*. Note the relevance of this for the understanding of the temporality of implicit mentalization and the contradiction in mentalization theory between the theoretical conceptualization of mentalization as a fast processing and the clinical insight of a slow feeling things out. The first intuitive giving often tend to be instable, so a reiteration of the process is usually necessary. Reiterating the process drives the intuition towards fulfillment and encourages internal, sensory, emotional gestures that allow you to evaluate the quality of

fulfillment, its insufficiency or its completion. This process tending towards fulfillment has three dimensions: Cognitive, emotional and intersubjective. When the criteria of these are simultaneously satisfied, we have an intuitive and convincing experience (Depraz et al., 2000). From a cognitive point of view, this evaluation is captured in the tension between confusion, absence, emptiness, unreadiness and that which still moves within on the one hand, and clarity, distinctness, evidence, completion, stability, coherence, being finished and at peace on the other. From the emotional point of view, we feel frustration or disappointment when the criterion is not fully satisfied opposed to the feeling of adequation, justice, joy, satisfaction or congruence when it is. From the intersubjective point of view insufficient fulfillment can be experienced as an inner conflict whereas sufficient fulfillments is more like "being carried along by a convincing argument which make you approve of it as you agree to recognize its truth" (Depraz et al., 2000, p. 63).

This lengthy description of the basic cycle, consisting of epoche and intuitive evidence, is a phenomenological version of detailing the processes involved in implicit mentalization; our automatic, pre-reflective, intuitive, non-verbal, and procedural way of understanding that runs unconsciously in the background of our lived experiences without the need on our part to pay attention to or be aware of it. But even though it goes unnoticed, it will nevertheless shape our experiences in important ways. Being mindful of these processes can help in maintaining a mentalizing stance when the emotional temperature in the therapy room is high. Cultivating this awareness is not easy, and the inherent challenges in doing this may serve as an elaboration of why Allen et al. (2008) found this aspect of the training more difficult than they initially assumed. One has to be able to stabilize and re-direct attention in a controlled manner, be able to switch between attention positions and avoid bias. One has to know what dimension of our experience we should direct attention to, know what scale of precision is appropriate and possible and we need a way to facilitate retrospective analysis as

this process is impossible to access in real-time due to the rapidity and complexity of the processes (Petitmengin, 2006).

### **A Session**

Following the basic cycle, the next phases of becoming aware is *expression* and *validation*. The basic cycle *can* stand on its own, the initial intuitive completion can very well lead to something that remains unsaid (Depraz et al., 2000). But if we want to *explore the possibility* of developing the quality of our own implicit mentalizing through the microphenomenological interview, as is the intention of this thesis, expression and validation are necessary. This is the focus of a work session, as will be exemplified below through microphenomenology.

### **The Context**

Reflecting on the process of becoming aware, explicating implicit aspects of our knowledge, both has its preliminaries and its aftereffects. The context in which this reflection is carried out may be tailored to improve our implicit mentalizing skills in a favorable way. This will be discussed briefly towards the end of this thesis.

### **Microphenomenology**

The microphenomenological interview was developed by the French philosopher Claire Petitmengin as an elaboration of Vermersch' explicitation interview (2018), originally developed in the late 90's in order to explicate implicit aspects of professional know-how. It is an interview that combines the two forms of intentionality referred to above. The basic cycle, just described, is a detailed description of the operative intentionality. Microphenomenology is an example of how the explicit intentionality can be directed at what is intuitively given through the epoche.

The interview is unstructured, in that the interviewer follows the descriptions the interviewee provides and there are no pre-determined questions. There is however a clear

strictness to it, and extensive training is required in order to conduct such an interview. In a microphenomenological interview one is taken through the steps or phases of the epoche in a systematic fashion. This is done through non-inductive and content-free questions directed at the structure of the experience under investigation and through the iterative process of the interview, enabling successively more and more fine-grained descriptions (Bitbol & Petitmengin, 2017).

There are several difficulties involved in performing the epoche. And the interventions in the microphenomenological interview is designed to overcome these obstacles (Petitmengin, 2006). The first phase of the interview is to evoke a memory or reenact an experience. It has to be done within a frame of suspension, where the main focus is to avoid comments, judgements, references to knowledge about a subject and so on. All these aspects of our "natural attitude" must be bracketed. This means the interviewer needs to have constant focus on this, re-directing attention from utterances related to this attitude, towards the inner space where our experience plays out.

The memory evoked has to be a specific singular lived experience (Vermersch, 2018), what Petitmengin (2006) calls an affective memory in contrast to an intellectual memory based on conceptual knowledge. The memory has to be reenacted in its embodied depth. It is hard to awaken such memories discursively, they tend to appear involuntarily through sensory triggers. But by directing attention towards visual, auditive, tactile, kinesthetic or olfactory aspects of the experience, they can indirectly be awoken (Bitbol & Petitmengin, 2017). The client needs to be stabilized in what Vermersch calls an 'evocation state', indicating contact with the past experience, before the attention can be directed to the inner processes with the aim of describing them. Indications of this state is the slowing down of speech, an unfocused gaze, dropping of the eyes and the use of present tense (Vermersch, 2018). As the attention is often unstable, the interviewee often needs to be brought back through questions or

reformulations aimed at the sensory aspects of the experience or one's own descriptions of the experience. One can also repeat the interviewee's gestures and use of generic terms (Gendlin, 1962; Petitmengin, 2006). When there is contact with the past experience, the attention must be directed away from the content, the 'what' of the experience that usually absorbs our attention towards the 'how', the structure of the experience. The aim is to draw focus towards the rapid phases which precede the emerging content and notice "the subtle inner microgestures that are performed to elicit, stabilize, recognize, evaluate, rule out or enrich" the content (Bitbol & Petitmengin, 2017, p. 734).

This change of attention can be done in relation to all kinds of activities, from the most common (like imagining, remembering, memorizing, problem solving) to more specialized activities in different professions (Petitmengin, 2006). As such, using this method to address implicit mentalizing skills in psychotherapists, is clearly possible. A difficulty one often faces is the tendency to give general descriptions, based on previous knowledge, when one shall attempt to describe the acts of experiencing. The skilled interviewer manages however to guide the attention towards the specific experience and to get more and more nuanced descriptions of both the diachronic and synchronic structure of the experience as it is re-enacted (Petitmengin, 2006).

To get descriptions of the diachronic structure, how the experience temporally unfolds, the interviewer asks content-free questions directed at different moments of the unfolding process, like: 'How do you start?', 'What happened then?' or questions aimed at the action verbs used, like e.g. 'When you did this, what did you do exactly?' or 'At the moment you felt this, what did you feel?' (Bitbol & Petitmengin, 2017). These kind of questions do not suggest any content or instill any presuppositions, they are just aimed at the different phases or stages of the experience. By going through the sequence of stages several times, more detail, nuances and sub-phases may be noticed. The same kind of content-free

questioning is also aimed at structural or generic aspects of different modalities of the experience, enabling description of a successively finer mesh of synchronic structures, including more dimensions to the experience. The elaboration of the descriptions is encouraged through frequent reformulations and by the iterative structure of the interview (Petitmengin, 2006).

Through these techniques, through directive but content-free and non-inducing questions and frequent reformulations aimed at stabilizing attention, the interviewee is guided through the process of *epoche*, kept in an attentive state where one can notice the gradual appearance of different aspects of our experiences. What is here implicitly and intuitively experienced, can be brought to attention and expressed through the guidance of the interviewer.

It seems impossible to describe the experience at the same time it is experienced. Expression then consists in a continuous alternation between getting in contact with the lived experience and describing the trace the experience leaves (Petitmengin, 2006). It is important to refresh this contact frequently as the trace may rapidly diffuse and the descriptions will then become more and more vague and empty. The reiterative structure, repeatedly focusing on the sensory aspects of the experience, helps maintaining this contact. There is however a serious concern whether expression and description distort the experience, and one can legitimately ask the question of whether or to what extent this happens. Petitmengin (2006) acknowledges this problem when she says there would be an interference if verbalization and experience were concomitant, but if verbalization is carried out *a posteriori*, there will be no interference of the experience. Hence, the importance of alternating between re-enacting the experience in silence and describing the internal trace. The question still remains how reliable such descriptions are. And Petitmengin's main focus of research has been on the aspect of validation of such expressions. She underscores the point of not having the epistemological

naivety of believing that a description, wherever disciplined it is produced, can be 'true' in the sense that it exactly reflects the initially lived experience (2006). Here we encounter the comprehensive domain concerning theories of truth. Petitmengin distances her approach from the dominant tradition of correspondence theories where truth is regarded as exact representations or correspondence between statement and facts. Rather, her approach is more in line with phenomenological notions of truth directed at the underlying implicit processes that makes phenomena apparent, let them shine forth, and subsequently allows for statements of correspondence or coherence (Heidegger, 2002). She acknowledges that every moment of explication introduces a transformation, and that this obvious fact is not to be avoided, but that we rather have to observe and describe these very transformations, i.e. describe

the inner operations or 'gestures' that enable me to enter into contact with my own experience, or cut myself off from it, that enable me to evoke a past experience, to divert my attention from 'what' to 'how', to direct my attention to the various dimensions of my experience, to alternate putting into words and 'refreshment' of the past experience (Petitmengin, 2006, p. 258).

By giving rigorous descriptions of these operations, the phenomenological theory of 'truth' may be strengthened through descriptions and refinement of its criteria of validation.

The current criteria for the reliability of the descriptions provided cover different aspects of the whole process of validation. Methodologically, one has to ensure that the interviewee's attention on the described experience is stabilized, that the attention is converted from the 'what' to the 'how', that the descriptions provided are not representations and general beliefs, but rather descriptions of a singular lived experience, that attention is directed towards the different dimensions of the experience and that the descriptions become iteratively more precise (Petitmengin, 2006). This is accomplished through the use of trained interviewers that act as skilled guides in their capacity of having sufficient knowledge about the processes and the inner structural landscape of experiences. Conducting the interview in the right way contribute to the satisfaction of methodological criteria. The intersubjective

criterion of reproducibility is a strong indicator of validity. When there are convergence of descriptions by different subjects, especially in cases where there are a sparse vocabulary or a lack of pre-established descriptive categories, convergence is a very convincing criterion for authenticity (Petitmengin, 2006). Research on the pre-ictal experiences of epilepsy conducted in a neurophenomenological methodological framework, may serve as an illustrating example (Petitmengin, Navarro, & Le Van Quyen, 2007). By using the microphenomenological interview to explore what is experienced prior to an epileptic seizure, similarities of the detailed descriptions by several patients enabled the subsequent detection of subtle patterns of activation of EEG related to epileptic activity that was earlier considered background noise. The main criterion of validity, however, is what Vermersch (2018) calls 'the speech position of the subject'. He distinguishes two types of utterances based on whether the one talking is in contact with his experience or not, with Petitmengin (2006) acknowledging the probability of a whole range of intermediate positions here. The distinction is drawn between an embodied and a disembodied speech position depending on whether or not the subject is in contact with the experience that is described. As this contact is loosened, the expressions are increasingly based on a vague memory of the expression, the memory of an account of the experience or on representations, beliefs or judgements about the experience (Petitmengin, 2006). Identification of the speech positions can probably be based on both subjective and objective indicators, both for the person speaking and for the person listening. Objective indicators are verbal (as the use of "I", present tense, specific context indicators of time and place and concrete character of the vocabulary used), non-verbal (as the direction of the eyes into "empty space" and use of gestures) and para-verbal (as the slowing of the flow of speech and increased silence between words). According to Petitmengin (2006), such clues make clearly perceptible the moment when the interviewee abandons his or her representations and comes into contact with the experience. The subjective, internal criteria of an embodied speech



position, that makes the subject able to distinguish between genuine contact with the experience and a position from which one refers to general knowledge, have not yet been sufficiently described (Petitmengin, 2006).

I hypothesise that the subject is then in contact with a very profound dimension of his experience, which is prediscursive, preconceptual, profoundly gestural, and prior to the separation into the five sensorial modes, in which the interior/exterior and I/others frontier is still permeable (Petitmengin, 2006, p. 257).

To summarize, the microphenomenological interview enables collection of fine-grained descriptions of a high level of reliability of the microdynamics of singular experiences, in their pre-reflective dimension. Based on this presentation of the microphenomenological interview, we are now in a position where we can start to consider whether this methodology can provide some tools with which we can work on improving the implicit mentalizing skills of psychotherapists.

### **How Can We Improve the Quality of the Therapist's Implicit Mentalizing Skills?**

The point of departure for this joint exploration of mentalization theory and phenomenology is the claim by Fonagy (2003) that our helpfulness as clinicians seem to be dependent on the quality of the implicit mentalization we offer through a general attitude. This general attitude is to be found in the mentalizing stance of the MBT-therapist (Allen et al., 2008). The challenges they described in refining this stance motivated the turn to phenomenology in order to explore whether this tradition can offer insights relevant for improving the quality of implicit mentalization. This turn was based on a suspicion that mentalization theory are constrained by its reliance on the ToM-framework in ways that make access to implicit processes of our understanding difficult.

## **Short-comings Of the ToM-Framework and Some Implications for Mentalization**

### **Theory**

Starting out, I want to acknowledge the comprehensive attempt in mentalization theory to tie together several aspects into a coherent theory. The mentalization based approach consist of a rich tapestry detailing the development of our cognitive apparatus by combining evolutionary science, neuroscience and developmental psychology, the actual working mechanisms of our mentalizing capacity based on cognitive science, a theory of psychopathology and mechanisms of change as well as detailing clinical interventions aimed at these various aspects. The way these areas have been brought together has allowed theoretical, clinical and research activities to mutually fertilize and sustain one another in an impressive way. As both a wide ranging and a continuously unfolding approach, one should perhaps not assume this to be a fully coherent theory without any inconsistencies. A serious admiration for the effort of combining all these different aspects under the umbrella concept of mentalization is at the heart of my interest in this approach. But in light of this admiration, I find it unfortunate that the theory is constrained by its reliance on representationalist and cognitivistic approaches.

**Theoretical Constraints of the ToM-framework.** One of these theoretical constraints has to do with the lens through which we perceive. There is no God's-eye view (Putnam, 1981). Every perspective makes certain aspects become visible whereas others become hidden from view. I suspect the ToM-framework can contribute to a certain blind spot that may be relevant in relation to how one can go about improving the quality of the therapists' implicit mentalizing skills, that is a blind spot towards the very existence of the space opened up through epoche. Based on the way implicit mentalization has been conceptualized in mentalization theory, one implication is the dichotomy between conscious and non-conscious mental states, not allowing any in-between. Although it does not seem

plausible to attribute such a *strict* dichotomy to mentalization theory and to the clinical applications of this theory, it nonetheless may serve as a substantial hindrance to open up this space for a fuller thematic exploration. I suspect the ToM-framework, largely influencing substantial parts of cognitive science and neuroscience as well, serves as an additional hindrance for exploration of this space of epoche, on top of the inherent difficulties regarding access to this space. One first has to be mindful of this dimension in order to explore it and systematically cultivate the associated skills. If one's theoretical foundation does not acknowledge the existence of implicit knowledge that can be made thematically available for consciousness, systematic exploration of this domain is made difficult, even though one's clinical insights tell you otherwise. And not knowing what to look for, makes it hard to recognize what is right in front of your eyes. The striking lack of research on clinical intuition in psychology (Marks-Tarlow, 2014) may serve as an illustration of how the main theoretical paradigms of western science concerned with objective knowledge, makes exploration of subjective experience difficult - even though subjective experience obviously is acknowledged as what is closest to all of us.

The cognitivist and representationalist assumptions of ToM may further constrain our access to the implicit aspects of our knowledge. The assumption that our mind is essentially closed off results in the forgetting of embodiment. If one understands the mind as a cartesian interiority, invisible from the outside, we end up with a dichotomy between the internal subjective and the external objective realms. And we need inferences or imagination to bridge the gap between the others outward visible appearance and their hidden interiority. This dichotomy miss the phenomenon of expressivity (Køster, 2017). Behavior expressed through the body is saturated with meaning, most of which is understood directly, in an automatic and pre-reflective way. Embodiment plays an essential role in experience and cognition, and implicit mentalization captures this aspect (Davidsen & Fosgerau, 2015). The boundaries

between oneself and the other are not hermetically closed as the cartesian perspective will have it. Our embodiment connects us at a fundamental level. But missing this perspective, access to our implicit knowledge becomes problematic. Relying on a framework that doesn't acknowledge the fundamental role of embodiment, makes a thorough exploration of the concept of implicit mentalization difficult.

The assumption that social cognition is rooted in inner mental representations makes one forget the enacted dimensions of social understanding (Køster, 2017). When one understands social cognition basically as an individual task of inferring or simulation the other's mind through inner representation one miss out on how social understanding is situated in interaction (De Jaegher & Di Paolo, 2007; Fuchs & De Jaegher, 2009). Social cognition is constituted in the in-between as a result of the reciprocal interactions between subjects. Patterns of interaction create, or enact, a world. Interaction imbues the world with significance, meaning and value and in this way structure our environment. The world and the other is not something pregiven it is our task to decipher on the basis of inner mental representations. They rather emerge from our ongoing patterns of interaction, embodied as implicit knowledge in the situational context.

These blind spots and constraints point to limitations in the theoretical framework of mentalization theory. That is not to say that the ToM-approach dominates every aspect of mentalization theory. But even though mentalization theory draws on different theoretical approaches, ToM is essential for the very conceptualization of the concept of mentalization. And ToM is, as I will argue, poorly suited to grasp the processes of implicit mentalization crucial for therapeutic effectiveness. Through my parallel reading of phenomenology and mentalization theory I have often been struck by the apparent similarity when it comes to clinical insights and implications. This makes me all the more puzzled about why

mentalization theory still favor the ToM-framework. The aporia between the theoretical implications and the clinical insights becomes increasingly noticeable.

### **A Potential Role for Phenomenology**

Phenomenology offers a framework that does not seem to be caught up in the same aporias regarding implicit mentalization as the ToM-framework does. First and foremost, it may serve as a guide in what to look for. It directly directs our attention to aspects of our experiences we otherwise naturally tend to miss. Phenomenology opens a space, through the *epoche*, that seems to correspond to what is understood as the processes of implicit mentalization. By giving detailed accounts of the unfolding of these pre-reflective processes that structure our conscious experience, phenomenology offers a potential way to work on improving the quality of this dimension of our relational knowledge. Phenomenology may also be a guide to the depth or the level of description that is actually possible to attain regarding these processes. Through the microphenomenological interview, fine grained descriptions of a few seconds of lived experience take about an hour to obtain. To do this one both need a framework that makes it possible for these elusive aspects to become visible as well as detailed knowledge in how one may go about doing this. Microphenomenology is one method to use. A somewhat similar approach that may be better known to clinical psychologists is the method of focusing or experiencing developed by (Gendlin, 1962, 1978). This approach is also deeply embedded in the phenomenological tradition, but it seems to be aimed, to a larger degree, at explicating the *content* and meaning of our implicitly felt sense. I would not in any way deny the potential fruitfulness of this approach as well for improving the quality of one's own implicit mentalizing skills. But I suppose the microphenomenological approach directed at the structures of our experiences is more in line with the basic tenets of mentalization theory. The explication of implicit knowledge, making the pre-reflective aspects of social cognition conscious, seems to be fully in line with what

Fonagy et al. (2019) recently have said about the essential interpersonal component of salutogenesis:

in recognizing and jointly considering the subjective experience of the individual, it becomes endowed with a conscious significance. This recognition by consciousness is valuable because it creates the conditions for epistemic trust and the possibility of adaptive social communication and learning with others (p. 98).

Explicating the implicit is thus crucial for the patient in MBT. To be able to assist in doing this, the therapist similarly needs to be aware of how these interpersonal components play out, in general as well as specifically in the therapeutic relationship. Phenomenology offers a way through which this knowledge can be attained.

Detailed descriptions of implicit aspects of subjective experiences may serve different functions. They can have a cognitive function in that it provides better understanding and gives knowledge about how experiences unfold as well as knowledge about different varieties of experiences. They can have a heuristic function in that they allow processes that has earlier gone unnoticed to be scientifically explored and they can have a pedagogic or therapeutic function (Petitmengin, 2006). This last function is highly relevant in this context. When we become conscious of and can describe the implicit aspects of our own subjective experiences, we have a better understanding of our own functioning. Under certain conditions we may then also transform our experience and in that way improve the quality of our own implicit mentalizing. This means we will be captured to a lesser degree by our own processes. We will have more freedom in professional settings, as in the therapy room. Knowledge and transformation of how we intersubjectively engage is especially important in circumstances where we are faced with hindrances and difficulties. In treatment with borderline patients, where the relational and emotional tension fluctuates rapidly and the temperature may rise very quickly, having the know-how to stay open and attentive is of utmost importance.

Through knowledge gained this way, we can for instance change the sequence of inner operations that may lead to the loss of patience in encounters with clients.

### **How Can Microphenomenology Aid in Improving the Therapist's Implicit Mentalizing Skills?**

One can imagine different ways this approach may be implemented in systematically working with improving implicit mentalizing skills. One can for instance address different challenges often faced in therapy and thoroughly describe the implicit processes involved. In this way one can attend to how we empathically tune in to others, how we create and hold space or what happens when this collapses. Or we can for instance attend to experiences of emotional contagion or the processes involved in having a hunch, visual imagery or other aspects of clinical intuition. Whatever the gateway into this work, the processes involved in the basic cycle seem to be a fruitful place of departure for addressing the *quality* of implicit mentalizing. Improving the quality can be addressed at different levels, or with different aspects or phases of the basic cycle as well as the experience of explication itself. One can attend to, become familiar with and work with the aspects of suspension, of what we do when we refrain from making judgements or which processes are involved in re-directing our attention and the precise way in which we do it. Similarly, one can become familiar with the aspects of letting-go and of intuitive evidence. In becoming familiar with the way we perform these processes that usually goes unnoticed, we similarly improve our clinical intuition. Clinical intuition may be described as being hit by a lightning-bolt, as having a sudden insight, a visual image, a hunch or as a bodily counter-transference (Marks-Tarlow, 2014). Marks-Tarlow (2014) claims our intuitive repertoire seems to be unique, even idiosyncratic, and also claims that therapists tend to keep this aspect of their practice secret, perhaps in fear of judgements. By knowing that the different aspects of clinical intuition can be explicated, by knowing how these experiences unfold and what they entail, by knowing the criteria of

validation of intuitive evidence and the validity of their explication, therapists may to a larger extent rely on and make use of their intuitions and their implicit knowledge.

Patients referred to MBT are assumed to have, among other things, problems with biased implicit mentalizing. If the therapists become more aware of their own implicit mentalizing, they will increase their understanding of and empathic attunement to the client. They will be more attuned to nuances in how these processes unfold, both in themselves and in the client, and thereby have the opportunity to give more fine-tuned responses. As a result, this is assumed to increase the mentalizing skills or the epistemic trust of the client through the way one is met and held by the therapist. This would be in accordance with what Petitmengin (2006) says about the importance of the interviewer's meta-knowledge. Knowledge of one's own inner gestures is required in order to guide the interviewee to their own experiences (p. 252).

The different temporal scales introduced at the beginning of the description of the process of becoming aware, makes it possible to situate the work of improving implicit mentalizing skills in a wider context. The experience of becoming aware of something is usually quick and instantaneous, it hits us as a sudden insight. The subsequent explication of this sudden insight usually takes an hour using the microphenomenological interview. Improving the skills involved in this endeavor, does however require sustained training. These skills are not easily cultivated, and we cannot expect to use the knowledge obtained through a theoretical elaboration to be directly used in therapy with patients. There are many levels between beginner and expert, and practice is the only way to move forward.

It seems possible to distinguish two different aspects of this training that needs to be combined. The cultivating of attentional skills and suspension may be done through various methods, e.g., through meditation or mindfulness practice. In this way, substantial parts of it can be conducted on one's own, given sufficient discipline. In the process of explication there



may however be a stronger need for a skilled mediator to scaffold the process and give external indicators so you can more easily pick up on the indicators of interiorized activity (Depraz et al., 2000). A supervision context could be used to explicate experiences, to repeat and to reflect on the descriptions provided. It is through explication we can deliberately work on transforming the way we implicitly relate and understand. These two aspects of the training should be done simultaneously with working with patients. In a way, one must learn on the job. And as Allen et al. (2008) reminds us, this is not easy:

We forget to monitor our own mentalizing capacity, and we need to do this while interacting with the patient—and not while we are sitting back listening quietly, but rather when we are actively engaging in the relationship (p. 185).

Working with patients is what gives raw material to the training, to the explication of implicit experiences. Even though parts of the training may be conducted outside the therapy, the refinement of our implicit mentalizing skills must be embedded in actual therapeutic interaction. We have to dive in and get our hands dirty, learning as we go along.

One should not be expecting results at once. But by gradually increasing awareness and noticing the internal micro gestures involved in our understanding of and communication with the other person, we feed this into our next practice of the process of becoming aware and into our next interaction. In effect, our implicit mentalizing skills will gradually improve.

How such training is to be implemented more specifically, remains an open question for now. One line of future research could be to investigate whether and to what extent this approach needs skilled mediators to be successfully implemented, or whether it is possible to develop a self-administered method to guide the explication of implicit knowledge. Besides two PhD-theses, only one study seem to have been conducted in this area. This study concludes that a self-inquiry form administered once to untrained participants is not sufficient to ensure reliable and fine-grained descriptions (Sparby et al., 2021). This is thus not entirely relevant in our context, where the need for repeated training over a period of time is assumed.

Research directed at ways self-inquiry may complement the use of skilled mediators may provide fruitful insights in how we can improve therapist's implicit mentalizing skills based on a phenomenological approach.

### References

- Allen, J. G., & Fonagy, P. (2006). *Handbook of mentalization-based treatment*. Chichester: John Wiley.
- Allen, J. G., Fonagy, P., & Bateman, A. (2008). *Mentalizing in clinical practice*. Washington, DC: American Psychiatric Publishing.
- Barlassina, L., & Gordon, R. M. (2017). Folk psychology as mental simulation. *The Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/folkpsych-simulation/>
- Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In *Joint attention: Its origins and role in development*. (pp. 41-59). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Bateman, A., & Fonagy, P. (2010). Mentalization based treatment for borderline personality disorder. *World Psychiatry*, 9(1), 11-15. doi:10.1002/j.2051-5545.2010.tb00255.x
- Bateman, A., & Fonagy, P. (2012). *Handbook of mentalizing in mental health practice*. Washington: American Psychiatric Publ.
- Bitbol, M., & Petitmengin, C. (2017). Neurophenomenology and the Micro-phenomenological Interview. In S. Schneider & M. Velmans (Eds.), *The Blackwell Companion to Consciousness* (pp. 726-739). Chichester, UK: John Wiley & Sons, Ltd.
- Choi-Kain, L. W., & Gunderson, J. G. (2008). Mentalization: ontogeny, assessment, and application in the treatment of borderline personality disorder. *The American Journal of Psychiatry*, 165(9), 1127-1135. doi:10.1176/appi.ajp.2008.07081360
- Colombetti, G. (2011). Varieties of Pre-Reflective Self-Awareness: Foreground and Background Bodily Feelings in Emotion Experience. *Inquiry (Oslo)*, 54(3), 293-313. doi:10.1080/0020174X.2011.575003
- Daly, A. (2014). Primary Intersubjectivity: Empathy, Affective Reversibility, 'Self-Affection' and the Primordial 'We'. *Topoi*, 33(1), 227-241. doi:10.1007/s11245-013-9206-7
- Davidson, A. S., & Fosgerau, C. F. (2015). Grasping the process of implicit mentalization. *Theory & Psychology*, 25(4), 434-454. doi:10.1177/0959354315580605
- De Jaegher, H. (2018). The intersubjective turn. In L. D. B. Albert Newen, and Shaun Gallagher (Ed.), *The Oxford Handbook of 4E Cognition*: Oxford University Press.
- De Jaegher, H., & Di Paolo, E. (2007). Participatory sense-making. *Phenomenology and the cognitive sciences*, 6(4), 485-507. doi:10.1007/s11097-007-9076-9
- De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10), 441-447. doi:10.1016/j.tics.2010.06.009
- Depraz, N., Varela, F. J., & Vermersch, P. (2000). *On Becoming Aware: A pragmatics of experiencing*. Philadelphia: Philadelphia: John Benjamins Publishing Company.
- Duschinsky, R., & Foster, S. (2021). *Mentalizing and Epistemic Trust: The Work of Peter Fonagy and Colleagues at the Anna Freud Centre*: Oxford University Press.
- Fodor, J. A. (1983). *The modularity of mind*: MIT press.
- Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *The International Journal of Psychoanalysis*, 72(4), 639-656.

- Fonagy, P. (2003). Clinical implications of attachment and mentalization: Efforts to preserve the mind in contemporary treatment-Part II-Epilogue. *Bulletin of the menninger clinic*, 67(3), 271-280.
- Fonagy, P., & Bateman, A. W. (2006). Mechanisms of change in mentalization-based treatment of BPD. *Journal of Clinical Psychology*, 4,62(4), 411-430. doi:10.1002/jclp.20241
- Fonagy, P., Gergely, G., & Target, M. (2007). The parent–infant dyad and the construction of the subjective self. *Journal of child psychology and psychiatry*, 48(3-4), 288-328. doi:10.1111/j.1469-7610.2007.01727.x
- Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Developmental Psychopathology*, 21(4), 1355-1381. doi:10.1017/S0954579409990198
- Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2019). Mentalizing, Epistemic Trust and the Phenomenology of Psychotherapy. *Psychopathology*, 52(2), 1-103. doi:10.1159/000501526
- Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G., & Target, M. (1995). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In *Attachment theory: Social, developmental, and clinical perspectives*. (pp. 233-278). Hillsdale, NJ, US: Analytic Press, Inc.
- Freeman, C. (2016). What is Mentalizing? An Overview. *British Journal of Psychotherapy*, 32(2), 189-201. doi:10.1111/bjp.12220
- Froese, T., & Gallagher, S. (2012). Getting interaction theory (IT) together: Integrating developmental, phenomenological, enactive, and dynamical approaches to social interaction. *Interaction Studies*, 13(3), 436-468. doi:10.1075/is.13.3.06fro
- Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the cognitive sciences*, 8(4), 465-486. doi:10.1007/s11097-009-9136-4
- Gadamer, H.-G., & Jordheim, H. (2003). *Forståelsens filosofi : utvalgte hermeneutiske skrifter* (Vol. 45). Oslo: Cappelen.
- Gallagher, S. (2007). Simulation trouble. *Social Neuroscience*, 2(3-4), 353-365. doi:10.1080/17470910601183549
- Gallagher, S. (2010). Movement and emotion in joint attention. . In D. M. a. J. S. S. Flech (Ed.), *Habitus in Habitat I: Emotion and Motion* (pp. 41-54). United States: Oxford.
- Gangopadhyay, N. (2017). The future of social cognition: paradigms, concepts and experiments. *Synthese*, 194(3), 655-672. doi:10.1007/s11229-016-1162-5
- Gendlin, E. T. (1962). *Experiencing and the creation of meaning : a philosophical and psychological approach to the subjective*. New York: Free Press of Glencoe.
- Gendlin, E. T. (1978). *Focusing*. New York: Everest House.
- Goldman, A. (2005). Imitation, mind reading, and simulation. *Perspectives on imitation: From neuroscience to social science*, 2, 79-93.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*: Mit Press.
- Heidegger, M. (2002). *The essence of truth : on Plato's cave allegory and Theaetetus*. New York: Continuum.
- Husserl, E. (1977). *Cartesian Meditations: An Introduction to Phenomenology*. Dordrecht: Springer Netherlands.

- Hutto, D., & Ravenscroft, I. (2021). Folk psychology, as theory. *The Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/folkpsych-theory/>
- Karmiloff-Smith, A., Klima, E., Bellugi, U., Grant, J., & Baron-Cohen, S. (1995). Is there a social module? Language, face processing, and theory of mind in individuals with Williams syndrome. *Journal of Cognitive Neuroscience*, 7(2), 196-208. doi:10.1162/jocn.1995.7.2.196
- Køster, A. (2017). Mentalization, embodiment, and narrative: Critical comments on the social ontology of mentalization theory. *Theory & Psychology*, 27(4), 458-476. doi:10.1177/0959354317711031
- Koubová, A. (2014). Invisible excess of sense in social interaction. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.01081
- Legrand, D. (2007). Naturalizing the Acting Self: Subjective vs. Anonymous Agency. *Philosophical psychology*, 20(4), 457-478. doi:10.1080/09515080701441736
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, 8(12), 528-533. doi:10.1016/j.tics.2004.10.001
- Liljenfors, R., & Lundh, L.-G. (2015). Mentalization and intersubjectivity towards a theoretical integration. *Psychoanalytic Psychology*, 32(1), 36-60. doi:10.1037/a0037129
- Luyten, P., Campbell, C., Allison, E., & Fonagy, P. (2020). The Mentalizing Approach to Psychopathology: State of the Art and Future Directions. *Annual Review of Clinical Psychology*, 16(1), 297-325. doi:10.1146/annurev-clinpsy-071919-015355
- Luyten, P., & Fonagy, P. (2015). The Neurobiology of Mentalizing. *Personality Disorders: Theory, Research, and Treatment* 6(4), 366-379. doi:10.1037/per0000117
- Mangan, B. (2007). Cognition, fringe consciousness, and the legacy of William James. In M. Velmans & S. Schneider (Eds.), *The Blackwell Companion to Consciousness* (pp. 671-685): Blackwell Publishing.
- Marks-Tarlow, T. (2014). The interpersonal neurobiology of clinical intuition. *Smith College Studies in Social Work*, 84(2-3), 219-236. doi:doi.org/10.1080/00377317.2014.923712
- Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology, Vol. 1. Theory and methods* (pp. 357-390): John Wiley & Sons.
- Newen, A., De Bruin, L., & Gallagher, S. (2018). *The Oxford handbook of 4E cognition*: Oxford University Press.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological review*, 84(3), 231. doi:10.1037/0033-295X.84.3.231
- Petitmengin, C. (2006). Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the cognitive sciences*, 5(3), 229-269. doi:10.1007/s11097-006-9022-2
- Petitmengin, C. (2017). Enaction as a lived experience: Towards a radical neurophenomenology. *Constructivist Foundations*, 12(2), 139-147.
- Petitmengin, C., Navarro, V., & Le Van Quyen, M. (2007). Anticipating seizure: Pre-reflective experience at the center of neuro-phenomenology. *Consciousness and cognition*, 16(3), 746-764. doi:10.1016/j.concog.2007.05.006
- Petitmengin, C., van Beek, M., Bitbol, M., Nissou, J.-M., & Roepstorff, A. (2019). Studying the experience of meditation through Micro-phenomenology. *Current opinion in psychology*, 28, 54-59. doi:10.1016/j.copsyc.2018.10.009

- Putnam, H. (1981). *Reason, truth and history* (Vol. 3): Cambridge University Press.
- Ratcliffe, M., & Bortolan, A. (2020). Emotion regulation in a disordered world: Understanding borderline personality disorder. In C. Tewes & G. Stanghellini (Eds.), *Time and body: Phenomenological and psychopathological approaches* (pp. 177-200). Cambridge: CUP.
- Reuter, M. (1999). Merleau-Ponty's notion of pre-reflective intentionality. *Synthese*, 118(1), 69-88. doi:10.1023/A:1005144911619
- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132(3), 617-627. doi:10.1093/brain/awn279
- Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., & Fonagy, P. (2020). Learning to mentalize: A mediational approach for caregivers and therapists. *Clinical psychology (New York, N.Y.)*, 27(3), n/a. doi:10.1111/cpsp.12334
- Sparby, T., Lumma, A.-L., Edelhäuser, F., Glaser, R., Schnitzler, L., & Weger, U. W. (2021). First-person access to decision-making using micro-phenomenological self-inquiry. *Scandinavian Journal of Psychology*, 62(6), 798-805. doi:10.1111/sjop.12766
- Thompson, E. (2007). *Mind in life : biology, phenomenology, and the sciences of mind*. Cambridge, Mass: Harvard University Press.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind : cognitive science and human experience*. Cambridge, Mass: MIT Press.
- Vermersch, P. (2018). *The explicitation interview*. Retrieved from [https://www.researchgate.net/publication/324976173\\_The\\_explicitation\\_interview](https://www.researchgate.net/publication/324976173_The_explicitation_interview)
- Volkert, J., Hauschild, S., & Taubner, S. (2019). Mentalization-Based Treatment for Personality Disorders: Efficacy, Effectiveness, and New Developments. *Current Psychiatry Reports*, 21(4), 25-25. doi:10.1007/s11920-019-1012-5
- Wampold, B. E. (2013). *The great psychotherapy debate: Models, methods, and findings*: Routledge.
- Zahavi, D. (2011). Empathy and Direct Social Perception: A Phenomenological Proposal. *Review of Philosophy and Psychology*, 2(3), 541. doi:10.1007/s13164-011-0070-3