



Università
Ca' Foscari
Venezia

**Second Cycle Degree Programme (D.M. 270/2004) in
European, American and Postcolonial Language and Literature**

Final Thesis

**Theory of Mind and the Brain: A View from the
Literary Framework and an Insight into the
Phenomenon of 'Mind-blindness'**

—

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Supervisor

Prof. Flavio Gregori

Co- Supervisor

Prof.ssa Roberta Cimarosti

Graduand

Francesca Pavan

838531

Academic Year

2015/ 2016

Table of contents

<u>Abstract</u>	5	
<u>Riassunto</u>	7	
<u>Introduction: Literature and the Brain</u>		
1. Anatomy of the brain in function of the literary experience.....	9	
2. Language and its relation to the brain: metaphor, parable and projection.....	10	
3. Consciousness and the attempt to bridge the gap self-others as represented by fictional characters in literature.....	18	
<u>Chapter 1: Affective Response to Literature</u>		
1.1 Role of emotions and their phenomenology: nature of the relation cognition-emotion.....	32	
1.2 Universality of emotions across cultures: prominence of both biology and culture in the field of cognitive cultural studies.....	37	
1.3 Some hints at narratology: rhetorical control of readers' attention through specific techniques mastered by the authors.....	44	
1.4 Dynamic aspect of feelings during the reading experience.....	51	
1.5 Empathy and the pragmatic value of literature.....	53	
<u>Chapter 2: Theory of Mind and the Issue Mind- Body</u>		65
2.1 Theory of mind and its implication in narrative.....	71	

2.2 Some hints at the history of cognitive science and the re- discovery of the body in the studies on cognition.....	73
2.3 The problem of other minds and the Cartesian dualism.....	77
2.4 How to deal with other minds: Theory- Theory and Simulation Theory accounts...	80
2.5 Empirical studies on narrative and Theory of Mind: how do we make sense of others in everyday life through the narrative process.....	88

Chapter 3: Debates on Theory of Mind and Alternative Accounts to the Dominant Approaches to Social Cognition

3.1 The Theory of Mind debate.....	104
3.2 Against Theory of Mind: Inter-subjective Relation theory, Intentional relation theory, Narrative practice Hypothesis.....	112
3.2.1 Alternative accounts of ToM.....	113
3.3 Further developments of the debates on theories of mind.....	121

Chapter 4: The Communicative Strategies of Joint Attention and the Phenomenon of ‘Mind-blindness’

4.1 Joint attention.....	138
4.2 Autism and mind-blindness.....	140
4.3 Empirical studies on mind-blindness.....	142

<u>Conclusion: Do we really need a Theory at all?.....</u>	151
--	-----

<u>Bibliography.....</u>	155
--------------------------	-----

Abstract

Studying literature means to study how the human brain works since literature cannot exist without a mind that, through the powerful tool of the creative imagination, give birth to the unique world of the literary artifact. For this reason, my aim is to clarify the peculiar and in some way controversial relationship between the brain, considered as a physical organ and thus studied by scientific, empirical disciplines such as neuro and cognitive sciences and literature, which is one of the highest form of art ever elaborated by the human mind since the beginning of the time. By doing so, I will take into account both some aspects of the neuroanatomy of the brain and its functioning in relation to the phenomenon of literary aesthetic reception and the pragmatic role that features typically defined as literary, such as metaphor, projection and foregrounding have in the everyday mind, trying to bridge in this way the gap between real and fictional world. In particular, I will focus on the phenomenon of character identification through an analysis of one of the most prominent feature of the reading experience, namely the empathetic response of readers toward fictional characters with an attempt to understand the neurological reasons that underpin readers treating fictional characters as they were real people that inhabit their real worlds. Accordingly, a brief overview on the role of emotion in the process of narrative understanding will be provided, insofar the link between emotion and cognition in relation to both biology and culture in the field of cognitive cultural studies is concerned. Moreover, I will enter into the debate about the various hypothesis on the “Theories of Mind” (ToMs), which are the set of dominant theories that have been employed to explain the way in which people make sense of other’s behaviors, intentions and desires in both real and fictional worlds. By taking an insight into the different accounts of ToM, I hope to help clarifying what is the core of these theories and why we really need a Theory of Mind to successfully relate to others in a social context as well as to empathize with fictional characters in order to better understand human nature as depicted by the works of art. Additionally, in order to shed light on the various hypothesis

on ToMs and how they work, some instances of empirical studies on the brain by using the recent discovered scientific empirical techniques will be given. In conclusion, to support the thesis in favor of the necessity of a Theory of Mind in order to correctly bridge the gap self- others, some cases of dysfunction of the ToM mechanisms in autistic children (known as “mind-blindness”) and people which underwent strokes that have inhibited their ToM abilities will be analyzed in light of the recent discoveries made by the empirical studies on the brain using the functional magnetic resonance imaging (fMRI), the positron emission tomography (PET) and electroencephalogram (EEG).

Riassunto

Lo studio della letteratura è strettamente legato allo studio del funzionamento del cervello umano dato che non ci può essere creazione artistica senza una mente che, attraverso il potente strumento dell'immaginazione creativa, dia vita al peculiare mondo della creazione letteraria. Per questo motivo, il mio obiettivo è quello di cercare di chiarire la peculiare e in un certo senso controversa relazione tra il cervello, considerato come organo fisico e come tale studiato da discipline scientifiche ed empiriche quali sono ad esempio le scienze neuro- cognitive, e la letteratura, intesa come una delle più nobili arti mai elaborate dall'uomo sin dall'antichità. Con questo fine, prenderò in considerazione nel mio elaborato sia alcuni aspetti della neuroanatomia del cervello e il suo funzionamento in relazione al fenomeno della ricezione letteraria attraverso i sensi, sia il ruolo pragmatico che alcune figure tipicamente considerate letterarie come metafore, parabole e la focalizzazione su alcuni personaggi e situazioni narrative piuttosto che su altre rivestono nella mente di tutti i giorni, cercando in questo modo di unire il divario esistente tra mondo reale e fittizio. In particolare, mi concentrerò sul fenomeno dell'identificazione letteraria attraverso l'analisi di uno dei più importanti fenomeni derivati dall'esperienza di lettura, ovvero l'empatia che porta i lettori a considerare e a trattare come reali i personaggi narrativi pur riconoscendo il loro ruolo fittizio. Per questo motivo, cercherò di offrire un'analisi delle ragioni neurologiche che spingono i lettori ad identificarsi con tali personaggi; per fare ciò, presenterò una breve panoramica del ruolo che svolgono dei sentimenti durante il processo di comprensione narrativa, concentrandomi maggiormente sul legame esistente tra cognizione e aspetto fisiologico e culturale delle emozioni in relazione agli studi cognitivi culturali. Inoltre, porgerò una breve panoramica delle varie ipotesi fino ad ora note denominate "Theories of Mind" (ToMs) ovvero la serie di teorie dominanti riguardo il modo in gli esseri umani sono soliti percepire ed interiorizzare il comportamento, le intenzioni e i desideri altrui, sia che si tratti di persone reali appartenenti al mondo quotidiano sia a quello fittizio. In questo modo, fornendo un'analisi delle varie ipotesi di ToMs, spero di poter aiutare a chiarire

quali siano i punti fondamentali di ciascuna teoria ed, inoltre, mi focalizzerò sull'importanza di avere una Theory of Mind per quanto riguarda le relazioni interpersonali, data l'estrema rilevanza di questo processo cognitivo non solo per poter fruire a pieno dell'esperienza letteraria, ma anche per poter stabilire delle sane ed equilibrate relazioni sociali nella vita di tutti i giorni, prerequisito fondamentale per poter capire meglio la natura umana, così come ce la presentano le opere d'arte. Inoltre, per chiarire ulteriormente la relazione tra le varie ipotesi di Theories of Mind e il cervello umano, verranno forniti degli esempi di studi empirici condotti utilizzando le recenti tecniche scientifiche non invasive che permettono di visualizzare variazioni di attività cerebrale mentre il cervello viene sottoposto a stimoli di varia natura. In conclusione, a supporto dell'ipotesi della necessità di avere una Theory of Mind per ridurre il divario sé-altri, verranno forniti degli esempi di disfunzione di tali meccanismi che portano al fenomeno denominato "mind-blindness", quali esemplificati da casi di persone affette dalla sindrome di autismo o che hanno subito danni a parti del cervello correlate ai meccanismi di ToM, casi che verranno analizzati alla luce delle già citate tecniche non invasive quali la risonanza magnetica funzionale (fMRI), tomografia a emissione di positroni (PET) e l'elettroencefalogramma (EEG).

Introduction: Literature and the Brain

1. Anatomy of the brain in function of the literary experience

How does the two cerebral hemispheres differ one from another in the process of elaboration and reception of language and how each hemisphere is involved in the process of narrative understanding? In order to give an answer to this relevant question, it is useful to first consider the anatomy of the human brain in light of the recent scientific discoveries about its working. As far as the processes of narrative and linguistic understandings are concerned, it is worth mentioning Broca's area, which is located in the inferior frontal gyrus (IFG). The IFG has been proved as a critical area for language production, even though this region of the human brain is involved primarily in competencies beyond language, whereas Wernicke's area, located in the posterior regions of the temporal lobe, is involved mainly in the encoding of vocal signals into meaningful words and sentences. In other words, Broca's area functions primarily in the planning of thoughts and execution of speech, whereas Wernicke's area functions are apt to make sense of the speech that a listener perceives.

Because of their role in language and other complex cognitive and motor functions, historically there has been considerable interest in the cerebral organization and lateralization of Broca's and Wernicke's areas in great apes, notably in chimpanzees since it seems that in this class of primates Broca's area is similarly located in the frontal operculum and is bordered anteriorly by the front-orbital sulcus and posteriorly by the pre-central inferior sulcus. Moreover, it has been proven that individual differences in joint attention abilities are associated with variation in gray matter in the posterior temporal cortex. In contrast, asymmetries in the gestures used in the initiation of joint attention are associated with gray matter asymmetries in the IFG, which is considered the homolog to Broca's area. These collective findings suggest that the non-verbal foundation of communication in chimpanzees overlaps with the neural basis of language in humans. Thus many of the behavioral phenotypes in non-verbal communication observed in

chimpanzees parallel those found in developing human children. For this reason, chimpanzees can serve as an important model species for understanding the biological factors influencing the typical and atypical development of joint attention and related socio-communicative processes.

2. Language and its relation to the brain: metaphor, parable and projection

Since language is a useful feature for social interaction and a tool that humans share with other primates, the discovery of mirror neurons in the 1980's changed the previous theories about language as prerogative of humans and explains their role as far as the process of narrative understanding and empathy are concerned. Is language innate and therefore are its mechanisms already inscribed in the genes or is it a direct consequence of the projection of metaphor, parable and other figures of speech? According to Chomskian linguistic theory, all speakers possess an ideal grammatical competence in the form of an internally represented set of linguistic rules. Yet, in processing linguistic expressions according to these rules, we have also to rely on other cognitive systems, such as short-term memory or systems for focusing our attention.

On the other hand, according to Mark Turner (1998), our mind is most and foremost literary as the stories we tell to ourselves everyday as part of making sense of the world demonstrate, so that narrative imagining is the fundamental instrument of human thought. Grammar, in his view, is not the starting point of language but instead it is the rhetorical figure of parable since it is the projection of stories that requires the invention of a grammar so that language arises directly from the projection of stories. As Turner points out "the backbone of any language consists of grammatical construction that arises by projection from abstract basic stories" (Seemann, 2012, pp. 257-8). Story and grammar have similar structures because grammar comes from story through parable; as a consequence, the origin of a rudimentary grammar happens before any hypothetical genetic specialization for a grammatical system. Since narrative imagining partitions and categorizes wholes into related elements and combines finite, grammatical elements into

infinitely many products, it seems that grammar arises from a dynamic system of projections of story structure to create an active system of grammatical constructions that successively can be adapted in many ways. Therefore, it is not important that all the members of a community share the same system that underpins grammatical construction; rather what matters is that the external products of that system fulfill the relevant communicative needs in local situations. Following this view, it is grammatical structure, rather than single lexical items, that arises from the projection of story. Grammar is based on everything involved in narrative imagining such as sensory modalities, motor capacities and perceptual categorization. Accordingly, the development of grammatical construction can mutually reinforce the improvement of the story structure.

What is the relation between the brain's fixed, inherited characteristics and its plasticity, its capacity to change, adapt and develop? To what extent are certain cognitive functions localized in particular cortical areas and to which degree is brain functioning distributed over different regions that reciprocally interact and are malleably arranged? How is the brain organized? Is it like a computer with a central controller with a parallel processing network or is it, instead, a chaotic, unstable array of ever-changing interactions? According to Turner's view, since the brain appears to be imprinted by experience with language, it has been proved that it is properly this early experience with language that affects cognitive operations that go beyond language, creating at a distance a virtual, distributed brain among all the participants of a culture. Therefore sentences are thought as small stories and, as a consequence, grammatical structure arises from a conceptual structure, while syntax derives from the projection of semantics onto phonology, rejecting any notion of a genetically programmed language module typical of the Chomskian view.

Nevertheless, it is impossible to explain where language comes from since we can only postulate the existence of a special genetic code that is responsible for creating an universal grammar organ in the brain, which someday might be located and understood. If human beings have some specialization for grammar at all, is still an open question as our understanding of the cognitive architecture is already too poor to comprehend its

functioning. Moreover, the ambiguities of evolution suggest interesting and important questions about what is genetically fixed in the brain and what may vary as a result of learning, experience and cultural differences. The neuroscience of vision, and in particular the work of the neuroaesthete Semir Zeki, helps to distinguish between inherited and acquired neurological traits so that “an inherited brain concept, or program, is not modifiable with the acquisition of further experience throughout life” whereas “acquired brain concepts are synthetic, a result of how neurons interact, and therefore change with experience.” As he explains, “We are not at liberty to discard, ignore, or disobey the brain’s inherited functions and features” (Armstrong, 2013, p.35).

Notably, grammar contains many devices such as tenses, moods, verbs that are useful to indicate how mental spaces are built and how they are connected so that it can be considered as a set of instruments useful for guiding the building and blending of spaces as well as the projecting of generic entities: in spatial perception, like in imaginative stories, there are various foci and temporal viewpoints and the projection of certain relations between them constructs tense. In English, for instance, the grammatical construction of present tense corresponds to the narrative category in which the temporal focus and spatial view point coalesce; the grammatical construction of past tense correspond to the narrative category in which the focus precedes the viewpoint and the grammatical construction of the future tense corresponds to the narrative category in which the view point precedes the focus.

Blending is fundamental at the very basic level of perception, understanding and memory as our perception of a thing is not the thing itself since there is no anatomical place where a perception of a thing or a concept resides and no point in which all these perceptions are anatomically brought together (for instance, an house appears to us as a unique thing but our visual perception of it is fragmented across the brain). Thus objects are perceived as having an unitary appearance and these unitary perceptions are the consequences of parallel mental activities in the brain, each with different functions, blended through an high order, complex mental mechanism. Specifically, it is worth to say that narrative structure is more complicated than grammatical structure since it varies very much from culture to culture. Nonetheless, what seems to be universal is the

stability of the repertoire of abstract basic stories and, among these, blending is an important feature since it is a fundamental activity of the literary mind, which can be defined as the mental set of abilities employed for connecting and combining concepts, that has always been considered an highly literary capacity.

As a consequence, the condition of the multitude of perceptions arises the binding problem, that of integration which tries to respond to the question of how we perceive a thing as unitary. Form, color and motion are all processed differently and in different places in the brain, so how is it possible that they are integrated through a perceptual and conceptual integration? Gerald Elderman in his book (1990) has proposed a model of “convergence”, according to which the brain contains records of the combinational relations of fragmentary accounts, so that the recall of events or entities arises from the reactivation of the fragmentary records contained in sensory and motor regions. According to this theory, “reentry signaling” depends upon the coordination of distributed fragmentary operations in the process of response to the visual stimulus to the world and this signal of reentering:

acts to coordinate inputs and resolve conflicts between the responses of different functionally segregated maps. Each mapped region use discriminations made by other regions for its own purposes [...] these various responses remain segregated and distributed among multiple brain areas and still constitutes a unified representation (Elderman, 1990, p.72)

There is a theory that is meant to unify the study of the brain with the study of the mind, which is the assumption that meaning is dynamic and active, although we perceive the sensory perceptions and their mental images as the same thing when they are in actual facts partitioned into two different concepts. As previously mentioned, sensory activity and conceptual activity seem to us a unitary whole but actually they involve a complex process of blending. Specifically, this act of blending abstract image schemas together is a basic conceptual operation whose mental processes of unification are mirrored and exploited by writers in the process of literary creation since there seem to be certain literary process pre-coded and embedded in the mind; thus the brain seems not to work in the ways we expect, based on our concepts of unitary and stable notions. This belief seems to be misleading to neurobiology since the blending process is not only a matter of biology but instead constitutes a basic feature of the everyday mind.

According to Norman Holland's account (2009), the study of unusual events, those which demands more attention, are useful to understand our mental functioning and are analyzed by scientific test experiments which employ elaborate and sometimes innovative techniques to study the fundamental mechanism of the mind; for this reason, cognitive scientists give special attention to the study of unusual events. Moreover, literary processes like blending make the everyday mind possible because features that has always been considered literary such as blending, parable, projection and metaphor are in actual facts the foundation of the mind and the tools through which it perceives everyday reality. Thus, how can we define these figures of speech? Aristotle is credited to have provided the first definition of metaphor in his Poetics:

Metaphor consists in giving the thing a name that belongs to something else; the transference being either from genus to species, from species to genus, or from species to species, or on grounds of analogy (Aristotle, 1997)

As it is clearly shown, Aristotle defines metaphor as a linguistic rather than cognitive matter. Over the past half century, metaphor has moved from being a peripheral topic in the philosophy of verbal art to assuming the status of a major philosophical, linguistic, psychological and scientific issue in the theory of cognition. For centuries prior to this dramatic change of status, metaphor was regarded as nothing more than a figure of speech serving various rhetorical purposes, but not as an essential part of human thought.

In sharp contrast, contemporary empirical research on language and cognition gives metaphor a central position in the processes of abstract conceptualization and reasoning, defining it as an essential cognitive instrument since metaphors have been proved to increase our ability to see deeper into some important aspects of reality. Moreover, the significant rise of cognitive neuroscience, with remarkable progress in the methods of neural imaging (i.e. PET and fMRI), has recently begun to make possible the exploration of neural models of metaphorical competence. These developments taken together have opened up an entirely new way of addressing questions about how we process metaphors and, thus, understand life. To provide a few instances of studies using fMRI, Tim Rohrer (2001) has shown that both literal and metaphoric sentences using hand terms such as "she handed me the apple" (literal meaning) and "she handed me the

theory” (metaphorical meaning) activate primary and secondary hand regions within the primary and secondary sensory-motor cortical regions/maps. Additionally, George Lakoff maintains that what constitutes the inference patterns that are used to reason about some abstract domain is spatial or bodily logic of this sort, preserved in neural maps. Lakoff summarizes the core idea of the neural theory of conceptual metaphor as follows:

In situations where the source and target domains are both active simultaneously, the two areas of the brain for the source and target domains will both be active. Via the Hebbian principle that “neurons that fire together wire together”, neural mapping circuits linking the two domains will be learned. Those circuits constitute the metaphor. (Lakoff, 2009, p.60)

How does the brain make the neural patterns in its nerve-cell circuits and manage to turn those neural patterns into the explicit mental patterns that constitute the highest level of biological phenomenon, called images? Solving this problem encompasses necessarily to address to the philosophical issue of the so-called qualia, which can be defined as the simple sensory qualities that compose images. By doing so and by referring to the concepts currently discussed in phenomenology, developmental psychology and cognitive neuroscience, it is possible to distinguish the minimal or core self from the extended or narrative self. As Damasio and other scholars noticed (Damasio 1999; Gallagher 2000; Rochat 2004; Zahavi 2005), the minimal self is characterized by an implicit, pre-reflective self-awareness that is present in every experience without requiring an act of introspection. Thus, any sensation, any perception or action directed towards an object implies a tacit self-awareness, an activity given immediately and non-inferentially. This first-personal “givenness” of all experiences may be regarded as a general medium in which specific modes of experience are articulated. As the most basic form of selfhood, it may also be called “mineness” or “ipseity” (from the Latin ipse = ‘self’). “Ipseity” is preserved even when autobiographical memory is lost, as in amnesia or dementia, or when a long-term sense for the future is missing, as in certain frontal brain damages(Gallagher et al., 2009, pp.549-51).

On the other hand, the extended or narrative self begins to shape in the second year of life and it is based on a number of emerging capacities that are closely interrelated, such as:

- a. The capacity for a higher-order awareness of one's conscious states, i.e. introspective or reflective self-consciousness.
- b. The capacity to understand others as intentional agents and to take their perspective, i.e. self-transcendence.
- c. The capacity to understand and issue verbal reports about one's own or others' feelings, thoughts and intentions, i.e. narrativity.
- d. The capacity to form a conceptual and autobiographical knowledge of oneself, i.e. a self-concept.

The extended self emerges in the course of early socialization, depending on the acquisition of autobiographical memory, concepts and language; thus its fundamental structure is inter-subjective and reciprocal since it is constituted through the ongoing relation to others, namely the social self. Moreover, taking the perspective of others implies a shift from an ego-centric to allo-centric space as well as assuming a concept of oneself and others as intentional agents who are responsible for their actions. In Damasio's view (2000), this is not only a cognitive achievement, but also it gives rise to a number of self-reflective emotions such as shame, embarrassment, feelings of guilt or pride, which all depend on the internalization of the evaluating gaze of the others. Accordingly, only a being with a constant sense of "mineness" is able "to form concepts about herself, consider her own aims, ideals and aspirations as her own, construct stories about herself and plan and execute actions for which she will take responsibility" (Gallagher and Zahavi, 2005). Disturbances of basic self-awareness are therefore bound to affect the extended self as well. Both dimensions of self-experience, however, are not present a priori, but have to be constituted by a complex interaction of biological, psychological and social processes. Additionally, the coherence of the self is a constant achievement of self-constitution and self-affirmation, an attainment which may be disturbed in manifold ways during life time.

Accordingly, at least two levels of inter-subjectivity should be distinguished: a primary inter-subjectivity, which develops during the first year of life and consists in imitating and recognizing facial expressions and other bodily movements (a resemblance

which is supposedly made possible by the presence of the mirror neurons system in the pre-motor cortex area) and a secondary inter-subjectivity, when around the age of one year infants, by noticing how others interacts in the world, start to recognize their goals and intentions (meta-representation ability).

As far as the disturbances of primary inter-subjectivity are concerned, autism is worth to be mentioned since in the last decades it has become a topic of discussion in both phenomenology and cognitive neuroscience. Autism is still dominated by a cognitive and modular approach, assuming a faulty development of ToM-modules that leads to a disturbed capacity to attribute mental states to others. In recent years, however, criticism has been raised by phenomenological psychiatrists arguing that the deficit already involves failures of early interaction and inter-affectivity. This idea is supported by the fact that many autistic symptoms such as lack of emotional contact, anxiety or agitation are already present in the first years of life, long before the supposed age to acquire a ToM. Moreover, between 15% and 60% of autistic individuals are able to pass false belief tests successfully, providing an additional evidence that the disorder can hardly be due only to a lack of ToM (Reed and Paterson 1990). From a phenomenological approach, autism should rather be conceived as a disorder of primary or embodied inter-subjectivity. This includes disturbances in sensory-motor integration, imitation and affect attunement as well as in the holistic perception of the self. Moreover, there is evidence that autistic children show a variety of basic sensory-motor abnormalities on the neurological level and a particular aspect of disturbed integration concerns the sensory-motor feedback loops involved in imitation.

Based on the mirror-neuron system and shared self-other representations of movement (Decety and Sommerville 2003), imitation serves as a major instrument for early social cognition and literature shows a consistent finding that people with autism do not readily imitate the actions of others (Smith and Bryson 1994; Hobson and Lee 1999) and there is also an increasing evidence for a mirror-neuron dysfunction in autism spectrum disorders (Oberman et al. 2005; Dapretto et al. 2006). Problems with imitation might then lead to a vast series of impairments in early inter-corporeality, affect attunement, joint attention, pretend play and in the acquisition of a successful ToM

ability. Moreover, autistic children show problems in establishing perceptual and situational coherence since they tend to focus on single parts or elements rather than perceiving the Gestalt of objects (i.e. the general quality of an object) and they are inclined to treat things as decontextualized, thus missing the particular meaning provided by the situation as whole (Frith 1989; Happé 1995). Therefore, according to this point of view, what autistic children lack is not properly a theoretical concept of others' minds, since ToM-like strategies of mentalizing and inferring from social cues are manifestly employed by high-functioning autistic individuals as a compensation for the lacking capacities of primary and secondary inter-subjectivity (Zahavi and Parnas, 2003).

3. Consciousness and the attempt to bridge the gap self-others as represented by fictional characters in literature

Consciousness is an entirely private, first-person phenomenon which is part of the private, first-person process that has been defined as "mind". Consciousness and mind, however, are closely tied to external behaviors that can be observed by third persons, thus the sciences of the human mind and behavior are based on this incontrovertible correlation between the private and the public, first person mind, on the one hand, and third person behavior, on the other. According to the Russian literary scholar Mikhail Bakhtin(1895-1975), consciousness never gravitates toward itself but it is always found in intense relationship with another consciousness so that everyday experience is dialogic and individual consciousness is always embedded in a social context. In his *Problems of Dostoevsky's Poetics* (1984), he claims that fictional characters are individuals with as much freedom and autonomy of movement as real people, so they are subjected to the same type of social interactions that guide human life. According to this view, the mind is always distributed among the participant of the discourse and it is always extended to the social environment, so that to include the physical story world.

In his analysis of Dostoevsky's characters, Bakhtin underlies how those characters are free people, capable of standing alongside their creator, having the power of not

agreeing with him and even of rebelling against him. In this way, they give rise to a plurality of independent and unmerged voices and consciousness, so that a genuine polyphony of entirely true voices is the chief characteristic of Dostoevsky's novels. What unfolds in his works is not a multitude of characters and destinies in a single objective world, manipulated by a single authorial consciousness, but rather a plurality of consciousness, each with equal rights and each with its own world, which combine but are not merged in the unity of the event. Typically, the consciousness of a character is given as someone else's consciousness; yet, at the same, time it is not turned into a close object, rather it is open to various interpretations. Consequently, as Bakhtin notices:

all the elements of novelistic structure in Dostoevsky are profoundly original; all are determined by that new artistic task that only he could pose and solve with the requisite scope and depth: the task of constructing a polyphonic world and destroying the established forms of the fundamentally monologic (homophonic) European novel. (Bakhtin, 1984, p. 7).

The history of each individual soul is presented in Dostoevsky not in isolation, but together with a description of the psychological tribulations of many other individuals; his world is the world of a multitude of objectively existing and interacting psychologies, which is an idea that stands at the core of the new defined genre called polyphonic novel.

To provide an example of dialogization and thus of polyphony, it is be useful to consider one of his latest work, *Notes from Underground*, a confessional novel published in 1864 which is characterized by an extreme and acute dialogization. The entire style of the *Notes* is subjected to the most powerful and all-determining influence of other people's words thus it does not contain a single word gravitating exclusively toward itself and its referential object; as a consequence, it is not represented a single monologic world. Moreover, it is worth to say that the intense relationship to another's consciousness is complicated by an equally intense relationship to his own self, which seems to be split as well as engaged in a continuous meta- analysis since the beginning of the confession:

You imagine no doubt, gentlemen, that I want to amuse you. You are mistaken in that, too. I am by no means such a mirthful person as you imagine, or as you may imagine; however, irritated by all this babble (and I feel that you are irritated) you think fit to ask me who I am—then my answer is, I am a collegiate assessor [...] My room is a wretched, horrid one in the outskirts of the town. My servant is an old country-woman, ill-natured from stupidity, and,

moreover, there is always a nasty smell about her. I am told that the Petersburg climate is bad for me, and that with my small means it is very expensive to live in Petersburg. I know all that better than all these sage and experienced counselors and monitors.... But I am remaining in Petersburg; I am not going away from Petersburg! I am not going away because ... ech! Why, it is absolutely no matter whether I am going away or not going away.

But what can a decent man speak of with most pleasure?

Answer: Of himself.

Well, so I will talk about myself. (Dostoevsky, 1994, p. 3)

The opening words of this passage make it clear that an internal polemic with the other is concealed, even though the other's words are just imagined and anticipated, determining in this way an attempt by the protagonist to read other people's mind, an attitude which he engages through all the narration. Furthermore, although the hero manifests indifference toward the other's opinion, precisely in this act of anticipating the other's response and in responding himself to the questions, he demonstrates to the other his own real dependence on him since he fears that the other might think he fears the other's opinion. Accordingly, through these fears he immediately demonstrates his own dependence on the other's consciousness along with his own inability to be at peace with his own demarcation of the self, being trapped in a vicious circle. As a consequence, it seems that what the man really wants to do with his own confession is to destroy his own image as it appears reflected in another's eyes, attitude which resolves in his ultimate desperate effort to free himself from the power and influence of other's consciousness upon him. In this attempt of striving to hold his own image and his own discourse as they exist in and for the other person, it is possible to hear the desire for a clear self-definition as well as perceiving the presence of others as mere interferences and as mere tools for better defining his identity, manifesting themselves as voices interrupting one another continuously, penetrating his entire body and even depriving himself of self-sufficiency.

Analogously, in Dostoevsky's novel *Crime and Punishment* (1866), what strikes most about its protagonists' dialogues such as those by one of its main protagonist, the murderer Raskolnikov and his monologic discourse, is its extreme internal dialogization. Moreover, the vivid personal address it makes to everything he thinks and speaks about, underlies that he seems not to think about phenomena but rather to speak directly to them. In this way, he frequently speaks within himself, sometimes using the 2nd person singular, as if addressing to someone else:

It shall not: be? But what are you going to do to prevent it? You'll forbid it? And what right have you? What can you promise them on your side to give you such a right? Your whole life, your whole future, you will devote to them when you have finished your studies and obtained a post? Yes, we have heard all that before, and that's all words, but now? Now something must be done, now, do you understand that? And what are you doing now? You are living upon them. They borrow on their hundred roubles pension. They borrow from the Svidrigalovs. How are you going to save them from the Svidrigalovs, from Afanasy Ivanovich Vahrushin, oh future millionaire Zeus who would arrange their lives for them? [...] So he tortured himself, fretting himself with such questions, and finding a kind of enjoyment in it. (Dostoevsky, 2008, p.48)

Notably, his inner speech is filled with other people's words that he has just recently heard or read: from his mother's letter, from things Luzhin, Dunechka, Svidrigailov had said that were quoted in the letter, from Marmeladov's speech which he had just heard, from Sonechka's words which he heard from Marmeladov so that his inner thought is constructed like a succession of living and impassioned responses to all the words of others he has heard. As a result, his discourse unfolds like a philosophical drama, where the dramatis personae are embodied points of view on life and on the world, realized in living situations so that all the voices that Raskolnikov introduces into his inner monologue converge into a peculiar sort of contact, one that would be impossible among voices in a real dialogue. In this way, both the discourse of the hero and the discourse about the hero are determined by an open dialogic attitude toward oneself and toward the other. Accordingly, authorial discourse cannot encompass the hero and his words on all sides nor it can lock in and finalize him but it can only address itself to him; all definitions and all points of view are swallowed up by dialogue, drawn into its becoming and making the novel a place for the on-going construction of the character's identity.

In conclusion, because all the voices sound within a single consciousness, they become reciprocally permeable and they are brought closer to one another, creating an overlap in which they partially intersect reciprocally, creating various interruptions in areas of intersection: everywhere there is an intersection, consonance or interruption of replies in the open dialogue by responses in the heroes' internal dialogue and everywhere specific ideas, thoughts and words are reported through several unmerged voices, sounding differently in each. Accordingly, the object of authorial aspirations is precisely

the passing of a theme through many and various voices, its rigorous and irrevocable multi-voicedness and vari-voicedness, so that the very distribution of voices and their interaction is what matters to Dostoevsky (Bakhtin, 1984, pp. 238-9).

The dialogic and polyphonic theory of novel is useful for the formulation of the hypothesis that the study of the mind and especially fictional minds, should involve the environment and above all the study of other minds, since the environment of human thinking is made foremost of other minds. Accordingly, interaction and inter-mental thinking (i.e. our knowledge and our skills are constructed by our dealing with other's minds) are our primary features as human beings and fictional characters share these properties with us. For this reason, the reading process is very creative in constructing coherent and continuous fictional characters toward which the readers are required to apply a complex series of mental strategies in order to suppose character's intention and inner lives. Among these, empathy, the use of real-world knowledge to fill the gaps in narration, the inclusion of both past, present and future information are prominent features since most of the information the reader acquire from a text is not explicitly written in it but rather inferred.

The embedded narratives are the story-like representation in the mind of a character and reproduced in the mind of the reader every time he/she engages with a work of fiction and, assuming that our minds are social, it is the relationship between minds that generates stories that are by nature aspectual as each character perceives the story world from his/her own peculiar point of view. Consequently, in order to interpret correctly a fictional world, a reader has necessarily to modify his own encyclopedia (i.e. the store of knowledge that readers possess and bring to the text in order to comprehend it) and enter in the fictional encyclopedia, which can be defined as the totality of knowledge about a story world.

The cognitive relationship between a character, a story world and the reader are key elements in the narrative process. Even though in the narration there are gaps of knowledge, the reader is able to fill them because he is accustomed to do so also in real life with real minds, by continuously reading them and making suppositions about their workings. In this way, another parallel shall be made between the real and the fictional

mind: real and fictional people exist only in a social context, thus the mechanisms to understand the beliefs, thoughts and intentions of others are the same both in real and story worlds. Moreover, as previously mentioned, because human consciousness is not an unified entity, but, rather is always in a conflict-ridden state between different mental states, any discourse is dialogic by nature since it is the manifestation of a dialogic nature of the world so that, according to Bakhtin, the internal dialogism of the world find expression in a series of peculiar features in narrative semantic, syntax and stylistic. The fictional mind is always distributed among the participant of a discourse, being extended to the social environment so that to include the physical story world; language is never unitary in the actual social life and neither is the literary language, which is always stratified and heteroglot in the forms that convey its meanings. As Bakhtin claims (1975), dialogism is the characteristic epistemological mode of a world dominated by heteroglossia, which is a term that refers to the social diversity of speech types. Following this view, everything is the result of the interaction with other beings, even language structure.

Indeed, what are the parts of the brain involved primarily in the process of narrative understanding? It seems that it is the right hemisphere the site in which literary language is processed since this part of the brain has been proved to be specialized in the understanding metaphorical meanings while the left one processes ordinary, plain language. When we are transported with a literary work, explains Norman Holland (2009), the right hemisphere language centers become more active than in its usual processing and, as a consequence, a wider range of poetic and literary meanings can be achieved. Each hemisphere has its own semantic system for decoding meanings so that both parts are indispensable for a correct interpretation of the word meanings in context. Interestingly, while engaging with literary works, the right hemisphere's language centers are dominant in receiving and interpreting ambiguous words, connotations, metaphors, jokes and sarcasm. Moreover, some recent studies suggest a regionally distributed network for the creation of narrative, which is a complex, multi-level process that involves the collaboration of various parts of the brain: episodic and semantic memories are generated in the amygdalo-hippocampal system, while the formulation of narrative

language involves left perisylvian region and the organization of persons and things into real and fictional temporal narrative frames take place in the frontal cortices and their sub-cortical connection to the limbic system, where the emotion are first perceived and elaborated.

Since the ultimate goal of the artistic experience is pleasure, the leading emotion in life that guides all human actions is seeking, defined by Holland as the search for desire-satisfying objects of their will. It is properly this mechanism the primary force that leads our will to live so that, in his view, cognition and emotion are part of the same process of narrative understanding since both the thinking and the emotional parts of human brain interact and take part in the process of literary understanding as well as in enjoying the artistic experience; notably, in this process emotions seem to modulate the brain activity and affect the rational understanding of the world. Accordingly, while we are dealing with a work of art, because of the inhibition of the reality-testing systems of the frontal lobes, the emotional systems become stronger so that we respond more emotionally than cognitively to literature. Humans can only do what their biology allows them to do so that, as part of their biology, the brain creates and re-creates literature and defines our response toward it.

A neuroscientific approach to literature has become increasingly prominent in the study of the literal phenomena and the ways the brain responds to fiction, since it seems that the same cognitive tools employed in real life are also involved in the process of narrative understanding. Accordingly, it is worth mentioning that historians of neuroscience, biology, psychology and neurology all agree in viewing the late eighteenth and early nineteenth centuries as a crucial period for the emergence of an unprecedented series of hypotheses and discoveries concerning the brain and the nervous system. It seems that only in the Romantic era the brain was definitively established as the organ of thought so that the innovative brain science of the late eighteenth and early nineteenth centuries can neatly be thought of as a direct consequence of the discoveries of the great thinkers of the Romantic period, one of which is certainly William Wordsworth.

Notably, Wordsworth's psychological poetry frequently roots mental growth in embodied experience, most notably in his masterpiece, *The Prelude*, an autobiographical

poem aimed at describing the growth of his own mind from the youth to mature life firstly published in 1798 and then republished in 1799, 1805 and posthumously in 1850. Wordsworth, in his poetry and poetic theory of the late 1790s and early 1800s, was deeply engaged in the new naturalistic and biological approach to mind which considered it an active organ shaped by experience rather being a static and passive element. To provide an example, these are the sensory and cognitive organs celebrated in the 1799 *Prelude*:

Thus day by day
Subjected to the discipline of love,
His organs and recipient faculties
Are quickened, are more vigorous; his mind spreads,
Tenacious of the forms which it receives.
(Wordsworth, 1998, 2.280-4)

As Alan Richardson notices in his work *British Romanticism and the Science of Mind* (2001) there is a growing interest for a kind of organic sensibility, terms that in the context of the late 1790s overlaps significantly with related terms like Darwin's "sensorium" and Cabanis' "sensibilité" since it implies a mind shaped by and realized in bodily organs, though not entirely defined by them. Accordingly, the poet of superior organic sensibility is not necessarily the one with the most susceptible sensory organs but sensations felt (or rather, produced) by internal as well as external organs provide the material out of which mind is constructed, an active sensibility stimulated by emotions, giving in this way an affective tone to cognition. As a result, Wordsworth perceives that: "all my thoughts / were steeped in feeling" (*Prelude*, 2.447-448). Thus, a genuine poetic sensibility start to gain importance in the process of poetic creation and the Lake poet in particular continues to register the permeation of thought with feeling and remains in touch with the sensational, bodily and emotive origins of mind, paying a special attention to "the lovely forms / and sweet sensations" and the "passions" that "build up our human soul" (*Prelude*, 1.134, 461, 462).

The 1799 *Prelude* is only one among several key texts from this period that suggest Wordsworth's interest in a naturalistic, biological approach to mind, as it is well shown in another passage from the same work:

Blest the babe
Nursed in his mother's arms, the babe who sleeps
Upon his mother's breast, who, when his soul
Claims manifest kindred with an earthly soul,
Doth gather passion from his mother's eye.
Such feelings pass into his torpid life
Like an awakening breeze, and hence his mind,
Even in the first trial of its powers,
Is prompt and watchful, eager to combine
In one appearance all the elements
And parts of the same object, else detached
And loath to coalesce (Wordsworth, 2: 269-80)

In the lines of the poem just quoted, Wordsworth is attempting at framing the naturalistic accounts of infant development that balance the empiricist stress on sensation with the infant's active participation in shaping the objects of that sensation, depicting a process of cognitive unfolding that give prominence to reason and emotion and that places the infant in a world of passionate social interaction from the moment of birth onwards. This new approach to the psychological development is a remarkable departure from the more passive, mechanistic accounts of Locke and Hartley and it looks forward to recent cognitive discoveries of neuroscience as backward to analytic philosophy.

Since the "nursing mother's heart," celebrated in the 1799 Prelude for its "awakening" role in subject formation will recur in later versions as the model for "feeling intellect", it can be said that The Prelude provides an even richer source for images that fuse the mental and the corporeal, as it explores the "growth of mental power" and models how the "mind of man is fashioned and built up" (1.67, 257). The mind is thus active and emotive from the beginning since the "infantile desire" (2.24) impelling the growth and organization of mental life through impassioned interactions with the natural and social world and guided by the innate activity of the developing mind, which makes the infant's first object of connection with the external world linked to the perception from the outset. In his way the mother's breast becomes the stimulus in the Prelude for the "first / poetic spirit of our human life" (305-69 and the poet who recognizes a "grandeur in the beatings of the heart" and celebrates the "hallowed and pure motions of the sense" acknowledges the importance of the role played by the body

and its organs in the formation and continual reformation of an active subject of perception which is “creator and receiver both” (1.383, 141; 2.303). The developing subject’s intercourse with the world can be therefore unconscious as well as consciously directed, another idea in harmony with the embodied psychologies of the time so that by “drinking in / a pure organic pleasure” from natural forms (I 395-396), the growing child maintains the active, emotive, bodily relation to the object world inaugurated in its “first trial” at the mother’s breast. Together, these images of organic growth and interaction convey an extraordinary image of embodied sensibility whose emblem is certainly the “beating mind” (Richardson, 2005, pp. 71-3).

Accordingly, historians of neuroscience have not hesitated to give a key role to Romantic conceptions and ideals: Edwin Clarke and L.S. Jacyna, for instance, in their 1987 study *The Nineteenth-Century Origins of Neuroscientific Concepts*, showed in detail how the foundations of modern neuroscience were established in the early nineteenth century by establishing new models of description of the functioning and structure of the nervous system stimulated by a kind of romantic biology and romantic philosophy of nature. Consequently, neuroscientific literature, which consists of a series of experiments aimed at describing what happens in the brain when we are doing literature with the help of behavioral techniques, is a useful tool for trying to explore how the brain both enables and limits people in creating and responding to literature, providing that studying literature means to study how the human mind works because literature is most of all an human activity, processed differently in the two cerebral hemispheres.

In particular, it has been demonstrated that the brain has found a way to produce neural states that captures features of the world, absorbing elements from the environment as sensing objects belonging to a space outside and having learned to distinguish them from its own ego, an activity indispensable for granting survival. As a consequence, in Holland’s account, the perceiving of the external world is just a projection of the internal expectations based on previous experiences, so that the world around us seems to be constructed rather than perceived. The result of this type of perception is what derives from the sense organs’ neurochemistry along with a good deal of cognitive messaging, so that sensations take place in the brains and neural entities have

the role of bridging inner and outer worlds. Regarding that, the reading experience is an extremely important human activity since it allows us to disintegrate the boundary between our world and the work of art: while reading we become part of the story we are dealing with, even forgetting our own body and the sense of time, being transported, as a result, into another timeless dimension.

This process of losing track of self and other is possible because we are perfectly aware that, while engaged in fiction, we will not try to change neither the characters nor the ending of the story, accepting it like it is. This is the great potential of art in general and especially of verbal art since of our agreement of treating a text as such, adopting a behavior of “disinterestedness”, as Kant calls the inhibition of the readers’ will to act upon the work of art he is dealing with, makes the experience of enjoyment possible. Poems and works of art are built in such a way we immediately recognize them as artifacts and thus different from reality; nonetheless, every time we come to a work of art, we consciously decide to believe. Hence, in Holland’s account, literature is only pretend and our “willing suspension of disbelief”, in Wordsworth’s terms, is what makes possible the enjoyment of art as well as our fully appreciation of it. This deal is what allows us to submit ourselves to the “poetic faith” so that as long we are transported in a work of art, we have no disposition to act upon the work of art itself, disengaging in this way all the mechanisms for testing reality and attributing judgments.

As it has been previously stated, literary stories are generally about people in social situations as human life and culture is incomprehensible without inter-subjective processes: human beings create story worlds together because they are motivated from birth to experiment with the exchange of fantasies with the aim of finding useful meaning in them and human consciousness, specifically, has the special gift of allowing imaginative travels through times and places. According to Colwyn Trevarthen, the stories of meaning we create are built on the mimetic skills we have inherited from our highly sociable animal ancestors, but at the same time we are born with new motives for fictional elaboration of rituals into new stories. In his words, our own personality, the “who” we are and the narrative of “what” we have done and known, grants us the role of

protagonists in a social drama where significant others live as supportive allies or contentious rivals (Zlatev, et al., 2008, p.8).

On the same subject, Alan Palmer (2008) claims that it seems likely that readers will draw on their normal modes of understanding people and encoding social episodes while they are reading stories, so that if the understanding of the self and the perception of social episodes involves a central role for affect, it seems likely that affect also plays a significant part in the narrative response by governing the cognitive processes of comprehension.

After having considered such issues, it is worth to make a reflection on what are readers doing in affective terms when they are immersed in the reading experience. The phenomenon of narrative comprehension involves relating the text to the existing knowledge of the world, defined encyclopedia, but there is also the sequential, experiential aspect of reading which covers ambiguity, indeterminacy and conflict between narrative schemata, features that require the reader's interpretative activity when those schemata are shifted, transformed or superseded. According to recent studies in neuroscience, as far as the emotional organization of human mind is concerned, our thought is not primarily linguistic but rather multimodal, involving past memories and associations as well as present ones. For that reason, our multimodal memories can be reactivated by a language that stimulates us to recreate past experiences, as fiction normally does.

As for the processes of narrative comprehension, they are not likely to be found in a single region of the brain but rather in its heterogeneous architecture since the nervous system is made both by computational processes and emotional ones which are likely to reside in diverse regions of the brain that are different but equally involved in the process of narrative comprehension. In this way, narrative is an universal feature as it integrates these two fundamental aspects of the nervous system, namely the path- dependent cognitive processes (the computational one related to the transmission of information through the nervous system) and state- dependent (i.e. the emotional component), rendering the narrative process a mid- level process. Accordingly, fictional characters are the greatest practical- schemas ever invented since they are employed to illustrate our

moral problems or to practice new emotional situations without requiring our direct involvement.

According to Blakey Vermeule (2010), the reason why we care about fictional characters is because we want to know more about ourselves and about the people around us, thus we exploit literary characters to create similar stories to the one we need to understand in our real lives. Following this view, fiction can be described as a feature which allow us to enter into a character's consciousness for understanding its workings; he goes further claiming that the tools used by writers to proud us to care are the same that govern our everyday mind. The issue of why we care about characters we know that do not exist is thus related to why we care about people who are not us since they are external, marginal to our lives but at the same time are part of our social environment. In this way, fiction become a feature which allows its readers to enter into a character's consciousness, asking them just to suspend their disbelief and paying them back with a large amount of social information that will be too costly, dangerous and difficult to obtain with direct experience. In this manner, literature provides the most intense cognitive stimulation imaginable since the inference system is affected by the literary process because humans normally tend to respond to situations they know are not real but might become real in the future.

Thus, where lies the boundaries between fact and fiction and which techniques are employed by the authors to stimulate our inference systems? As already mentioned, humans are most interested in social information and interaction and therefore prefer among other types of information, social information. Literary characters are more flexible than other pieces of literary code such as plot line or time sequence framing, so authors can experiment more with their lives in order to make them more similar to every day characters. Moreover, writers of fiction seem to continually experiment with our mind reading ability, which is the most important cognitive mechanism of human sociability: by stimulating it, they can explore human mind deeply and at various level of intentionality.

Generally speaking, it can be summarized that at the center of high mind- reading novel, there is a narrator who adopts the position of an agent with full access to strategic,

social information and successively distributes that information across the story and among the characters in order to make the plot more interesting and to appeal to our mind-reading capacities. In this way, the arts could be considered one of the principal means through which humans organize their complex motivational habits and, consequently, structure their own evolved emotive inclinations into a functional program of behavior.

Chapter 1: Affective Response to Literature

1.1 Role of emotions and their phenomenology: nature of the relation cognition- emotion

In his 1872 masterpiece, *The Expression of the Emotions in Man and Animals*, Darwin taught us that most of our emotive reactions and in particular those known as the primary emotions (fear, anger, disgust, pain, surprise, joy and so on) consist in a collection of responses that have been preserved during the course of evolution due to their original adaptive utility. Moreover, he claims that it is not surprising that these emotions occur in the same form in different species and, especially in humans, in different cultures. As previously mentioned, most interactions with the environment and the emotive conduct related to them depend on our capacity to perceive and understand the emotions of others. According to Paul Ekman (1999), who revisited Darwin's account on the primary emotions, it is interesting that among all these, disgust is one of the emotions most investigated in neurophysiological studies since brain imaging studies showed that when an individual is exposed to disgusting odors or tastes, there is an intense activation of two particular zones in the brain, which are the amygdala and the insula, the parts considered the emotional regions of our brain.

By the end of the nineteenth century Charles Darwin, William James, and Sigmund Freud had written extensively on different aspects of emotions giving to them a privileged place in scientific discourse so that, throughout the twentieth century and until quite recently, both neuroscience and cognitive science has given emotions an important turn: Darwin had conducted an extensive study of the expression of emotion in different cultures and different species, William James has produced an extremely important account on how our body affects our feelings and vice versa while Freud has focused himself into the pathological potential of disturbed emotions. Additionally, Hughlings Jackson has gone a step further underling a possible trajectory for the neuroanatomy of emotion suggesting that the right cerebral hemisphere of humans was probably the dominant one for emotion, much as the left one was the dominant for language. Unfortunately, during most part of the twentieth century, emotion was left apart from the

scientific studies since it was thought to be too much subjective to be studied in a laboratory. In this way, a clear division between mind-body, typical of the Cartesian view, was established since the mind remained linked to the brain in a somewhat equivocal relationship and the brain was classified as a consistently separated organ from the body rather than perceiving the two as part of a complex living organism. Only recently, the opposition emotion- reason is no longer accepted and the work of both neuroscience and cognitive studies has started to recognize the central role of emotions in the process of reasoning and decision making (Damasio, 2000).

Emotions are important in guiding human goals and for this reason they are essential for fictional minds, being the vehicle through which its characters normally obtain their ends. For this reason, emotions can be considered as some of the most complex phenomena of subjective experience, conceived by neurobiologists as states of bodily arousal that sometimes lead to conscious feelings. Interestingly, consciousness, according to Antonio Damasio, begins as a feeling of knowing and his account of emotion is an elaborated version of William James' theory of consciousness (1890): according to James, emotions regulate our behavior by turning us to the world in certain ways. Following this path, Damasio claims that emotions consist of changes in both the body and the brain and he goes further in classify between primary and secondary emotions, where the former are a sort of pre-organized, innate mechanisms whereas the latter are acquired associations between stimuli from the outside world and states of bodily arousal. As he puts it:

All emotions use the body as their theater (internal milieu, visceral, vestibular and musculoskeletal systems), but emotions also affect the mode of operation of numerous brain circuits: the variety of the emotional responses is responsible for profound changes in both the body landscape and the brain landscape. The collection of these changes constitutes the substrate for the neural patterns which eventually become feelings of emotion (Damasio, 2000, p.66).

As well as offering a theory of the nature and role of emotions and feelings, Damasio suggests a detailed account of the brain areas that are most likely involved in the process of emotional elaboration, claiming that there is not a specific emotional system but rather a number of areas are implicated in the phenomenon of emotions:

Sites such as the amygdala are part of multi-region systems that trigger emotions. On the other hand, sites in the hypothalamus, the basal forebrain (e.g., the nucleus accumbens), and in the brainstem (e.g., the nuclei in the perinuclear gray), are the principal executors of an emotion. These are structures that directly signal, chemically and neurally, to the body and brain targets whose changes will come to constitute an emotional state. Thus phenomenology can contribute to the description of feelings (Damasio, 2000, p.67).

Novels are stories with an high emotional content and therefore they are often memorable for the reason that emotions are the vehicle through which we obtain our goals. As mentioned before, emotions can be socially distributed among individuals in a group so that mental events, processes and states are crucial to the concept of embedded narratives, where action is considered as the result of a series of reasons, motives, intentions and purposes specified by the narrator and reached by the characters through the story. It is interesting to observe how this mechanism is developed in the discourse by the narrator and, in this regard, the neurophysiology and neurochemistry come to our help, suggesting that the mammalian brain has a seeking system that leads organisms to pursue what the environment offers to them, be this language, food or knowledge. Accordingly, it is worth to mention that appetite leads life since it is our instinct to search for the desire- satisfying objects that allows us to carry on life with the aim of satisfying our desires and needs.

Seen in this way, emotions are part of an homeostatic regulation and are poised to avoid the loss of integrity of the individual. The seeking mechanism, in Holland's view, is thus the one that allows us to enjoy reading or watching a film since it make us continuing to deal with a specific work for seeking pleasure. Following this line of thought, cognition and emotion are part of the same process of narrative understanding since the thinking and the emotional parts of the brain both interact and take part in the process of literary understanding and enjoying, a cognitive activity that draws on verbal skills as well as on our knowledge of the world and personal experience. All this activities rely on cortical systems, located in the frontal and temporal lobes while the sites of emotion seem to be the sub- cortical areas of the brain, making in this way a clear separation between the cognitive cerebral areas and the emotional ones.

The relation between cognition and emotion has always been a problematic issue in cognitive studies. After the discovery of the mirror neurons at the beginning of the 1990's, which have been defined as groups of neurons that fires both when an human being or an animal acts and when the animal/person observes the same action performed by another so that their neuron "mirrors" the behavior of the other as though the observer was itself acting, the process of empathic reading started to be better explained in light of our neurological architecture. After this break-through discovery, it has been declared that while watching or reading about an action, the motor regions of the brain experience an impulse to imitate that action but then, simultaneously, the brain inhibits the impulse to act. As a consequence, mirror neurons are thought to be responsible for the process of empathy and identification at the neural level and are present in all mammals as they confers advantages in survival and reproduction as knowing what others are thinking or feeling can determine a better outcome of a situation; interestingly, it has been stated that mirror neurons can also be related to our emotional response to poetic language.

According to Mark Tuner (1998), metaphor and poetic language in general involve a great deal of brain activity, requiring a great effort in connecting the target (something that is not yet understood) and the source (the already known concept). This cognitive process is modulated by emotions which regulate the brain activity and affects our rational understanding of the world so that, as previously mentioned, while we are dealing with a work of art, because of the inhibition of the reality- testing systems of the frontal lobes, the emotional systems become stronger so that we respond more emotionally than cognitively to literature. In other words, when involved in literature, the cognitive system inhibits those actions that the emotions would normally prompt, so that action is separated from emotion and the emotional system is engaged while the action system is disengaged. Moreover, because of the awareness that what is represented cannot be changed, we can feel more intensely the emotions that we would feel if the situations were real. This lead us again to the seeking process firstly identified by Holland since we turn to literature mainly to have our emotion stimulated, whether in pleasant or unpleasant ways.

Literary characters are the target for representing human emotion and interaction with the world, thus they can be defined as useful tools to better understand human nature since in novels characters' intentions and desires can be recognized perfectly while in real life we can know other people only in an approximate way. A possible explanation for this phenomenon is that while in ordinary life we do not separate between "the what" of semantic knowledge and "the where" of visuo-spatial perception, we do separate the two pathways in literature and from this differentiation derives our confused perception of the character's identity in nature, that is to say, if they are real beings or just fictional artifacts. As a consequence, we feel for literary characters the same feelings that we have for real people, so that they are like an optical illusion for our brains since when we imagine things, the "what" and "where" contradict each other¹.

According to Lars Bernaerts (2013), the representation of character's consciousness has played and still play a key role in narratology (i.e. the theory and the study of narrative and narrative structures and the ways these affect our perception) and it involves the already mentioned experience of qualia, that is the quality of "what it" to be a certain character or being in a certain situation, ability that requires an active role of the reader in inferring character's intentions and feelings through the activation of his Theory of Mind system. In this way, the reader is not a passive spectator but rather he is involved directly both physically, through the mediation of mental simulations and psychologically, by activating his mind reading abilities during the narration.

Generally speaking, it can be said that narrative simulates the intrinsic, non-intentional quality of an experience, defined as qualia, which is the sense of "what is like" to be in a certain mental state, such as that of a specific fictional character: through metaphorical language, the reader is invited to imagine some experiences without actually experiencing them, so that narrative fiction, with its power to allowing the reader to investigate indirectly what there is in other people's mind through the literary representation of his/her consciousness, is what distinguishes novelistic discourse from other kind of discourse. This is the peculiar value of verbal art: the capacity to provide the

¹ For a fuller account, see Semir Zeki's essay on vision and visual perception in Frackowiak, R. et al., *Human Brain Function*, pp.161-9, USA: Academic Press, 2004

mimesis of the mind, showing the dynamic interplay between assimilation and estrangement, naturalization and denaturalization through the quasi- mimetic evocation of real life experience, where are depicted various models of human mind at work.

Literary narrative is a multi- level cognitive performance, in which consciousness is the dominating function that can surface at several levels assuming different shapes; notably, the literary construction of experience disturbs in some ways the attempts to naturalize the minds of the protagonists since, although their minds are private, these minds reflects features of the literary communication. In this way, narrative presuppose the activation of a mind schema since the encounter with fictional worlds and minds represented in novels, plays or films allows readers to engage emotionally as well as cognitively with other cultures, stimulating the adoption, even though just temporarily, of other's point of view and to engage with other's perspectives, preparing the ground for the rise of a more opened, intercultural mind.

1.2 Universality of emotions across cultures: prominence of both biology and culture in the field of cognitive cultural studies

The nexus between narrative and minds seems to rely on emotions since the reaction to certain situations tell us something important about the way people of a certain culture perceive events in life. Especially, emotional reaction to literary works explicitly tell us something about what moves people of a precise culture; thus literary response is an important feature to understand more deeply human nature and the difference in taste in the various cultures.

According to Patrick Colm Hogan (2003), there are extensive narrative universals which mirror the universality of emotion across cultures. The study of literary universals is thus important for cognitive studies since it gives prominence to the cognitive tools responsible for the process of understanding literature which are shared by all people around the world from ancient time since the present days. In Hogan's view, the nature of narrative is prototypical and so is emotion, with standard situations which reflect

universal feelings or conditions shared by people belonging to various cultures and having different mind-sets (for instance, young lovers overcoming obstacles to reach an happy ever after marriage is a typical, cross-cultural story). In his account, emotions are a biologically determined processes, depending on innately set brain devices, laid down by a long evolutionary history, so that the prototypical stories can be seen as the expansions of the micro narratives that defines our prototypical emotions.

As Chris Danta noted (2014), the way in which our mind constructs stories is part of an universal human process by which we learn to distinguish, organize and adapt to our environment. Literature is structured in such a way that its emotional effects are founded upon the priming of emotional memories and our reading process is important since it reinforce the procedure of identification between what we read on the page and our personal, laden memories which come to the surface while we proceed in the reading experience. In Danta's description, emotion is a function of narrative so that emotional inference through the qualia is a matter of sensations and of finding the appropriate words to express them.

Patrick Hogan (2011) goes further in developing this idea by sustaining that the structure of stories is inseparable from the structure of human mind, being it strictly related to emotions, so that affection is strongly dependent to the making of stories since feelings affect and explain the stories we create. The story structures are fundamentally shaped and oriented by our emotional system which governs not only goals but also the way in which stories are developed, so that emotion defines the standard features of all stories. As a consequence, in order to formulate a systematic theoretical account of stories, we need to turn first to affective neuroscience, the science responsible of studying emotions, which claims that an emotional experience involves both physical reactions such as changes in vocalization, posture, facial expressions and an active response toward the source of a particular emotion. This conative sensitivity to the emotional expression of others appears to be innate and thus it is a defensive mechanism that humans share with the other primates as there seems to be a set of genetic instructions that are responsible for producing an emotion system, which are related to a complex system of interactions between one another and with the environment. While we are examining a

work of verbal art, Hogan claims, we may have an interest in interpreting the particularity of the work (hermeneutic interest) but also we may be interested in its commonality across works (studies on universals). In order to unify the two fields, the isolation of patterns that recur across works in different traditions and different historical periods is important for understanding the various sorts of particularities that are salient for an hermeneutic study.

The link between story structure and emotion has been studied since ancient times by Aristotle who, in his *Poetics*, has stressed the emotional force of recurrent themes in stories such as the recognition of stock characters and the reversal of the initial situation of conflict, so that story structure appears to be deeply influenced and bound up with emotions. Recently, various researches in cognitive science and neuroscience have enriched the study of emotions in relation to stories, contributing to redefining their peculiar relationship. Among these, David Herman (2013) strongly sustains that there is not only divergence but also convergence among cultures as far as the group dynamics are concerned, as the likelihood of resolving environmental problems or the struggle for food; storytelling practices are in his view inextricably interlinked with ascriptions of reasons for acting, thus interpreting narrative strictly relies on the analysis of our cognitive tools. Furthermore, it is possible to understand other minded creatures because of the existence of emotional states that are not simply qualities related to subjective experience but rather are given through bodily gestures and actions and therefore become visible to others.

It is worth to say that consciousness and emotion are not separable, nor are they a monolithic constructions, since consciousness, for instance, can be divided into simple, core consciousness (i.e. the part that provide the organism with a sense of self in a precise and unique moment) and a more complex, extended consciousness (i.e. a biological, stable phenomenon across lifetime that do not depend on memory or external circumstances). The extended type of consciousness, in particular, is interesting for literary studies since it provides the organism with a complex sense of self (such as an identity) and a progressive sense of time; it evolves during life time, is organized in several levels and it depends on memory. Accordingly, Antonio Damasio in his account

of consciousness in relation to story structure, claims that the two types of consciousness are equally important for the perception of the self as a whole since the core consciousness corresponds to the core self (i.e. the set of unique characteristics that defines an organism's identity) while the extended consciousness is related to the autobiographical self, which can be described as the organized record of the main aspects of an organism's biography across time and space.

At the beginning of the twentieth century, linguistic models of studying narrative, defined by the formalist narratological approaches of Gérard Genette, Algirdas Julien Greimas, Roland Barthes, Tzvetan Todorov, undervalued the role of emotions since their methods were more scientific than affective, with the result that classical literary theory has been influenced more by psychoanalysis than cognitive science. On the other hand, the approaches that started to consider affective science and neuroscience in relation to literature has been defined post-classical, which have been foreseen by the theories of Mikhail Bakhtin, who has been one of the first critic to recognize the fundamental role of emotion and subjectivity in the process of narrative telling.

As far as the re-evaluation of the role of emotion in the study of narrative across cultures is concerned, Hogan's account of the relation among story structure, emotions and cultures is prominent: for him, there seems to be three predominant story prototypes that recur across various cultures, namely the romantic, heroic and sacrificial one that derive respectively from the particular emotional systems shared by all cultures. Accordingly, some narrative patterns that can be found in various texts belonging to different periods and countries support the idea of the existence of a cross- cultural system of emotions which are reflected in the stories told by its people. In this way, narrative structure can tell something important about human mind and its relation to the particular society in which it is embedded, so that the analysis of story structure in relation to the emotional system has important consequence also for the understanding of the workings of the mind. For this reason, it is important to include both western and non-western stories in the analysis of literary reception as well as ancient and contemporary ones in the analysis of cultural universals. Particular features which have been considered as emblems of subjectivity in narratives are spatiality and the sense of time, which can be

considered as emotional experiences differently perceived by the various characters of a story.

There are neurobiological tests supporting the phenomenon called “place attachment” which is strictly related to social attachment, as observed by Jaak Panksepp (2004). In his account, the spatiality of human being is a sort of emotional geography that develops out of a fundamental human propensity of organizing the world along the axes of normalcy and attachment. The emotional encoding of experience seems to occur repeatedly and at different levels in the brain and subcortical and cortical areas seems to be primarily involved, so that the experience of an emotion seems not operate on an absolute scale but it involves our expectations and anticipations. Consequently, attachment to a certain character involves a special likelihood of emotional contagion since we generally empathize more with characters that reflect our feelings and that are similar to us.

Accordingly, Hogan claims that there are a great variety of theories regarding the emotional response to a literary work, the most important among these is called appraisal theory, which claims that we respond emotionally only to facts with a certain meaning or significance for us. Thus appraisal can be defined as the process of evaluating the ways in which a particular situation affects one’s ability to achieve one’s goals and some appraisal models appears to be modular (i.e. they posit distinct mental systems that govern different types of goals). Following this theory, it seems that we do not respond emotionally simply to physical or abstract configurations but rather to the meaning that certain situations have for us since our perceptual systems are configured in such a way as to be sensitive only to certain types of stimuli.

Among all the emotion elicitors, namely perception, recollection and imagination, the last one is particularly relevant as far as the study of human inferential abilities are concerned. Since it is parallel to perception, imagination seems to involve the same areas of the brain that become active both when something is imagined and when it is perceived physically. Hence imagination is a very important feature as it conjuncts current perception with memory through anticipation, helping in this way the brain to anticipate the sensations that will result from perceptions by transforming them into

action in few fractions of second. Moreover, according to appraisal theory, we respond emotionally to literature because we mentally simulate the experience of the characters by linking their emotional response with our own emotional memories. In this way, we process emotional incidents into events by extending them in both the past and the future, elaborating the past through causal attribution and the future by anticipations in working memory. As a consequence, story occurrences that are not routine but emotionally relevant for us, appear to be relevant and such episodes are the elaboration of events into complex unities, usually beginning with a shift away from normalcy through an incident that interrupts the linear development of the story, so that they seem to be highly significant as far as the mental conceptualization and perception of stories is concerned.

As for the neurophysiological aspect, empathetic response to other people's feeling has consequences at both the subcortical level, namely the region that responds to perceptual information and emotional memories with the activation of emotional systems, and the neocortical level that involves the appraisal-like elaboration of the full perceptual situations, like the outcomes. Notably, the first factor governing empathetic response is the intensity of the emotion, so that we are more likely to respond to deep feeling such as crying than to simple sorrow derived from seeing someone frowning.

Since the two main purposes of verbal art seem to be the emotional and the didactic one, emotions operate on actual experiences but also on simulation and imagination of that experiences; they contribute to our decisions and influence our future actions, so that emotions of this type are thought to be bounded up with our decision-making processes. Following this line of thought, literature can be defined as a feature that bears primarily on our emotional system and, as emotions are strictly related with our thoughts and actions in the world, it can be affirmed that literature can have the potential power of training human sensibility. Emotions have the power of affecting our decisions and behaviors because they are consequential and not already pre-set or pre-coded in our neural architecture, so that it is through feelings, which are inwardly directed and private, that emotions, which are outwardly directed and public, begin their impact on the mind. Since stories produce emotion episodes and these are likely to be stored as emotional memories, they can deeply influence our personality and behavior so

far that we relate to the world in function of the way we imagine it. Stories seem to affect our social and personal interactions in the world and literature can help or inhibit our willingness and ability to simulate other people's thoughts and feelings since the relation between fictional characters and real people seems to be really strong, being the latter a plainer version of the complex and sometimes extreme experience of fictional characters. Analogously, culture seems to play an important role in the process of narrative understanding since it influences our ways of seeing the world and language, as a cultural and inherited vehicle, is another strong means of expressing thoughts and feelings that has a fundamental cognitive value, as each individual consciousness is always embedded in a social context.

The question whether specifically linguistic and grammatical principles have evolved in the human species was first theorized by Noam Chomsky. According to Chomsky's theory of universal grammar, humans possess an universal set of grammatical features that contains such purely linguistic things as nouns, verbs and basic rules of European grammar shared by every human being in the world. Nonetheless, as it became clear that these features were not valid for many non-European languages, this hypothesis changed to include very abstract linguistic things, supposedly representing the universal computational structure of language. Despite that, the Chomskian hypothesis of an innate universal grammar was no more valid for explaining the phenomenon of universalism across cultures. Conversely, according to the new theory, languages have been created within the constraints of preexisting human cognitive and social tools, a view that clearly rejects the idea that of an evolution of some kind of innate syntactic features which sustains the hypothesis of the existence of an universal grammar.

As noted by Michael Tomasello (2010), although many aspects of human linguistic competence have evolved biologically, specific grammatical principles and constructions have not since universals in the grammatical structure of different languages are generated from a more general process that has emerged along the development of human cognition, communication and vocal-auditory processing, operating during the conventionalization and transmission of the particular grammatical construction of every single linguistic community. Such a view has been widely accepted so far, substituting

the previous account of universalism, even though there are still doubts and controversies among the linguists since currently there are no evidences that can prove these theories empirically.

1.3 Some hints at narratology: rhetorical control of readers' attention through specific techniques mastered by the authors

Cognitive psychology provides us with empirical methodologies to study the handling of attention by authors in different types of stories insofar readers are proved to be manipulated by controlling foreground and background details. Moreover, recent studies have shown that literary texts capture better than non- literary text reader's attention by employing particular linguistic devices to raise and lower attention. Speech representation and thought representation, by shifting the focus toward one character or another and by changing the focus of attention, are fundamental techniques to better understand character's consciousness. In particular, foregrounding is one method that influence reader's reaction and it refers to the range of stylistic effects that occur in literature, whether at the phonetic level (e.g., alliteration, rhyme), at the grammatical level (e.g., inversion, ellipsis), or at the semantic level (e.g., metaphor, irony). It enables literature to present meanings with an intricacy and a meticulousness that ordinary language does not normally allow and, by creating complexity of various kind, requires cognitive work on the part of the reader, a work initiated and partially directed by feelings.

Following this theory, when perception has been de-automatized, that is, has become a conscious rather than an automatic, unconscious activity, a reader employs the feelings that have been evoked by the work of verbal art to find or to recreate a context in which the de-familiarized aspect of the story can be located. In this way, readers take longer to interpret foregrounded passages, to savor their affective implications and to evaluate the contributions of those passages to understand the story as a whole. De-familiarization can be defined as a phenomenon that is central to literary experience as it

can be considered the hallmark of literariness, a process during which a reader uses prototypic concepts in a context where their referents are rendered unfamiliar by various stylistic devices, so that the reader is required to reinterpret such referents in non-prototypical ways or even to relocate them in a totally new perspective that must be created during reading.

The origins of the de-familiarization theory may be traced back in the Romantic period, especially in Coleridge's proposal that the purpose of literature is to overcome the automatic nature of normal, everyday perception in order "to give the charm of novelty to things of every day"(Coleridge, 1817, Preface). A similar position was undertaken by the Russian Formalist Victor Shklovsky in his essay *Art as technique* (1925) where he sustained that "habitualization devours things and art exists so that one may recover the sensation of life"(Lemon et al. 1917, p. 12). It is a matter of facts that literary texts complicate comprehension by challenging the familiar, prototypic concepts that readers initially apply to the text since one of the central functions of literary language is to loosen or to put into question, the normal relationship between the diction of the text and the referents of the words adopted by the writer.

This peculiar feature is the poetic function to which the Russian formalist Roman Jakobson refers to (1990) where he claims that "the poetic function deepens the fundamental dichotomy of signs and objects" (Jakobson, 1990, p.70). Nevertheless, feelings of appreciation (aesthetic pleasure or interest) are only an initial moment in readers' response to the formal components of literary texts that interact during reading to produce composite and interactive metaphors of personal identification that modify self-understanding. By doing so, literature take us away from the familiar world leading us to the unfamiliar and unexpected, in order to promote the understanding of the behaviors and social mechanisms that govern everyday life of which we are mostly unaware because we are accustomed to them in the everyday routine. In order to call attention to the work of art itself, writers uses metaphor, alliteration, consonance, metonymy and other literary features in order to prolong the reader's perception, making him/her focusing better on the thing itself. Indeed, by interrupting the automatic perception of the word, the he/she is forced to make extra effort in determining the meaning of the text.

This process can be identified with the peculiar value of verbal art, which is the capacity to provide the mimesis of the mind, showing the dynamic interplay between assimilation and estrangement, naturalization and denaturalization through the quasi-mimetic evocation of real life experience, where are depicted various models of human mind at work.

Stories are built in such a way that they recall attention to the structure and function of storytelling itself; this reflexivity explains why narrative constitutes a powerful resource for making sense of the world as well as making sense of experience so that aesthetic emotions differ from the everyday emotions they reenact, having an “as if” (qualia) dimension. Moreover, characters in narratives can be exploited to model the so called interpersonal complexity with the aim of exploring cross-category relationships and to revealing the transformations of a person over time.

As Teresa Bridgeman has noted (2007), time and space are more than background elements in narrative; indeed they are constitutive elements that forge the text and affect the basic understanding of it, deeply influencing the way in which we build mental images while we read. Since we tend to think of stories in terms of sequences of events, the category of time has always played an important role in theories of narrative. Space, instead, has often been set in opposition to time, being associated with static descriptions ending up slowing the narration. Nevertheless, time and space are strictly bounded up with each other in narrative discourse since, according to classic narratology, “story” (i.e. the basic sequence of events that can be abstracted from any narrative telling) and “discourse” (i.e. the presentation and reception of these events in linguistic forms) mutually cooperates in written narrative texts. Moreover writers, by mastering these story features through changes of time (flash back and forward) and changes of perspective (internal, external, zero) play with our Theory of Mind mechanisms by modifying, tricking or fostering our expectancies about the text we are dealing with. Thus time and space affect readers on different levels: firstly, since the process of reading is itself a temporally situated experience, we are involved and embedded in the process of narrative understanding in first person as active participant and observers of the workings of the various features of the discourse. Secondly, any world we construct when reading is just

partial since it still requires a spatio- temporal stability that readers constantly try to grasp by fixing events and sequence of events in a spatial and temporal position in order to give a meaning to the story.

Following Uri Margolin's point of view on characters in novels in the same book², a character in the widest sense can designate any entity, individual or collective, usually human or human-like introduced in a work of art, so that characters exist both within and outside the story world, being defined in relation to the story they belong to. The cognitive-psychological approach, on the other hand, view characters as text-based mental models of possible individuals, built up in the mind of the reader in the course of textual processing, being in this way a complex, mental representation created by the readers which has no connection with reality.

To provide a fuller picture of the hypothesis of the so called socially "distributed mind" (i.e. the idea that minds are always grounded in the discourse) it is worth to take into account David Herman's study of Ernest Hemingway's short story *Hills Like White Elephants* (1927) where he sustains that the fictional mind is an on-going construction through the text rather being a pre-existing entity and the same is true also for the reader's own construction of reality. In his account, Hemingway's story creates its two protagonists as embodied by implying the physical situatedness of their view of the white hills across the valley. Even though this story, as many others by the author, presents only overt and surface behaviors of the characters, omitting to report their internal states, thoughts and feelings, as the story unfolds it is possible to perceive character's inner states by inferring them in their exchange of dialogues, so that their mind do not reside beneath the surface of their attitudes but rather is presented as distributed in what they say or even do not say. Accordingly, an effort of mind reading skills is required to Hemingway's readers since the story offers little information about the setting and the appearances of the characters:

2 For further explanations, see *The Cambridge Companion to Narrative*, ed. David Herman, pp. 66-70, Cambridge: Cambridge University Press, 2007

The girl was looking off at the lines of hills. They were white in the sun and the country was brown and dry.

– They look like white elephants, she said.

- I've never seen one, the man drank his beer.

- No, you wouldn't have.

- I might have, the man said

- Just because you say I never had doesn't prove anything (Hemingway, 1987, pp. 211).

These are the opening lines of the story where it is displayed an exchange of ideas concerning the appearance of the hills between the unnamed male protagonist and a young girl named Jig. Both characters clearly rejects the storyline in which they seeks to position themselves reciprocally since each has a different, irreducible position regarding not only the landscape but also the decision they are bound to make (in this case, whether or not doing the abortion). Accordingly, they remain fixed in their position, without taking into account other's point of view or correctly interpreting other's mental states, inability that make them incapable to understand themselves and their true feelings. This failure of mind-reading inevitably leads the two protagonists to isolation, denying the possibility of a successful interaction with the environment in which they live.

The study mentioned above on Hemingway's narrative is part of a larger approach involving the structural theory and analysis of narrative texts, known as narratology, which is a discipline that defines narration as the telling of a story in a way that simultaneously respects the needs of the author and enlist the co-operations of its audience. Among the narrative features examined by narratology, focalization is an important category since it explains the ways in which narrative information is submitted to a certain perspectival filter. Normally, the narrator is the functional agent who verbalizes the story non-verbal matter, as well as manages the exposition and establishment of the relationship among the characters, thus deciding what it is to be told and in which order. Interestingly, one of the main aim of post-classical narrative theory, which distinguish itself from the classical approach for taking into account the embedded/extended mind, is to develop tools for studying the mind relevance in storytelling with the aim of understanding what cognitive processes can be lodged in modes of narration, types of perspective and spatio-temporal contexts of narrated

situations and events. Among the concepts that inform the post-classical narrative theory, the following features are worth to be mentioned:

- Positioning: the act of self and other positioning in the story, which is the first step for making sense of both our own and other minds;
- Embodiment: the idea that the mind should be viewed as the nexus of brain, body, environment; the most evident manifestation of this bound is the metaphor;
- Minds as distributed among characters belonging to a story, so that cognition should be viewed as a supra/trans individual activity distributed among groups functioning in specific contexts;
- Emotionology: the collective emotional standards of a culture opposed to the experience of the emotion itself. Stories are interpreted in light of models of emotion shared by groups of people;
- Qualia: the concept/sensation of being “in someone else’s shoes” . The idea of the concept even before the experience of that situation; stories enable us to know the experience of being someone else without actually be that persons.

Though, the figural text appears to be determined by the filter of the focalizer’s mind, while the reader, seeing the story world through it, becomes a witness rather than the narrator’s communicative addressee.

Gérard Genette (1983) famously claimed that in the zero focalization events are narrated from an omniscient point of view (as Henry Fielding’s *Tom Jones*) while in the mode of internal focalization the events of the stories are focused through one or more story- internal reflector, thus the events are seen from a single perspective, as in Joyce *The Portrait of the Artist as a Young Man*. In the external focalization, on the other hand, events are reported from various, external points of view, making this mode of reporting thoughts the most open to different interpretations. According to Dorrit Cohn (1978), just as narrative can use direct, indirect and free indirect discourse to present utterances of characters, fictional texts can use quoted monologue, psycho- narration and interior monologue to present the thought process of fictional minds. She goes further to explain the relationship between authorial narrator and the figural mind:

The more conspicuous and idiosyncratic the narrator, the less apt he is to reveal the depth of his character's psyches or to create psyches that have depth to reveal. It almost seems as though the authorial narrator jealously guards his prerogatives as the sole thinking agents within his novel, sensing that his equipoise would be endangered by approaching another mind too closely and staying with it too long (Cohn, 1984, p.25).

In this way, as far as the representation of characters' consciousness in third person narration is concerned, it is possible to notice that the more consonance there is between the authorial narrator and its characters, the less audible the story is; the more dissonant they are, the more the story is audible. Specifically, it seems that free indirect discourse has been the most employed narrative technique in modernism since it allows a deeper analysis of the character's consciousness both by the writer and by his readers, reporting directly what is going on in character's mind without removing completely the figure of the narrator, who still offers important details about the psyches and feelings of his characters which sometimes are difficult to understand and even remain unrevealed. To provide a significant example of how psycho-narration works, Peter Walsh's meditations about his life in Virginia Woolf's *Mrs. Dalloway* should be taken as a noteworthy example:

He was not old, or set, or dried in the least. As for caring what they said of him—the Dalloways, the Whitbreads, and their set, he cared not a straw—not a straw (though it was true he would have, sometime or other, to see whether Richard couldn't help him to some job). Striding, staring, he glared at the statue of the Duke of Cambridge. He had been sent down from Oxford—true. He had been a Socialist, in some sense a failure—true. Still the future of civilization lies, he thought, in the hands of young men such as he was, thirty years ago; with their love of abstract principles; getting books sent out to them all the way from London to a peak on Himalayas; reading science; reading philosophy. The future lies in the hands of young men like that, he thought (Woolf, 2014, p.50).

In this case, free indirect discourse (psycho narration) is employed for making the reader aware of Peter's false beliefs about both himself and other people, since he is not able to correctly interpret their desires and thoughts, making himself a mind-blind character who assumes constantly an angry and broken attitude, being in this way the exact opposite of Clarissa's joyful and optimistic takes on life. The psycho-narration brilliantly portrays all the flows and hesitations of the character's volatile psyches under the author's control

(for instance, Woolf often clarifies her character's thoughts by adding between brackets some additional information about her characters' real and sometimes unconscious feelings). Thus this particular narrative technique contributes at rendering the literary experience richer and more vivid, focusing precisely on emotional moments of character's lives which are critical for them, namely point of crisis characterized by an intense reasoning that makes the reader at the same time part of that mental swirl and apart from it.

In conclusion, the two major problems that narrative discourse informs seems to be those relating the events to one another in time and keeping track of the participants in those events when they are sometimes the same and sometimes different from the events. As far as the first issue is concerned, it is worth to say that keeping track of events in time leads to some incredibly complex grammatical structures. For instance, the simplest situation is when an event is located in time close to the present, even though narratives usually require us to situate events displaced in space and time which belong to the past or to the future. Thus the grounding of events in time in extended discourse and the tracking of referents across events are complex processes which the communicative function seems to drive, so that all linguistic communities who desire to develop some kind of narratives and to engage in other forms of extended discourse must create a set of grammatical conventions to which attend for understanding entirely the story world created by the narrator.

1.4 Dynamic aspect of feelings during the reading experience

To what degree feelings may be thought to guide the reader's interpretive activity? Feelings are culturally determined and, especially in literary response, they can be divided into various domains: evaluative feelings toward the text as a whole, such as the overall enjoyment, pleasure or satisfaction for reading, the narrative feelings toward specific aspects of the fictional event sequence, such as empathy towards a character or resonance with the mood of a setting, then the narrative feelings which are prompted by

events and characters in the imagined world of the text and the aesthetic feelings in response to the formal (generic, narrative, or stylistic) components of a text, such as being struck by a deep metaphor and finally self-modifying feelings that restructure the reader's understanding of the textual narrative and, simultaneously, the reader's sense of self.

Keith Oatley (1992) suggests that feeling plays a mimetic role since reading, in his term, is a simulation in which "the central process is that the reader runs the actions of the character on his own planning processes, taking on the character's goals, and experiencing emotions as these plans meet vicissitudes"(Oatley, 1992, p.66). Aesthetic feelings reflect heightened interest, thus in this sense are prompted by the formal (generic, narrative and stylistic) features of a text and, in response to such features, readers usually slow their reading and manifest greater uncertainty.

How feelings may play a determining role in modifying readers' self-understanding? The distinction between remembered emotions and fresh emotions seems to mark two contrasting modes of literary response: in remembering an emotion during reading, a similarity between the world of the text and some analogous situations scripted in memory seems to arise, so that whenever we recognize a setting as something familiar or whenever a character's behavior reminds us of something we have done, an implicit comparison is recalled in memory. In experiencing a fresh emotion, in contrast, we realize something in a literary text that we have not previously experienced. This tendency to the reinstatement of previously experienced feelings within familiar settings has given rise to the notion of affective scripts. During literary reading, aesthetic feelings, such as those derived from feeling struck, captured or held in response to foregrounding passages, provide the reading experience with a diffusely heightened feeling tone, an affective context within which narrative feelings are more likely to cross conventionally scripted boundaries. Among all the positive effects that feelings have on readers, it can be mentioned the anticipatory effects as they alerts us of the fuller significance of an event as that event begins to unfold and prepares us to interpret incoming evidence in a specific way; anticipation of this kind seems to be one of the fundamental properties of feeling as a reminders of our evolutionary struggle.

Moreover, when foregrounding accentuates aesthetic feelings, the readers' sense of self is sometimes challenged: although feelings during reading may be called into play mimetically, as remembered feelings, they may evolve in new directions as reading proceeds. Remembered feeling, in other words, does not remain merely replicative since what begin as remembered is gradually modified in the course of the literary reading. Accordingly, it seems likely that readers make often implicit comparisons between the world of fiction and what they know, believe or feel about their own lives, connecting intrinsically real and fictional minds.

1.5 Empathy and the pragmatic value of literature

The issue of how we understand other minds was one of the central philosophical concerns in the late nineteenth and early twentieth centuries. It seems likely that typical adult human beings have developed the ability to understand other persons as minded creatures by recognizing others' states of mind and by making sense of their behavior in light of their mind's powers. This ability has to be conceived as the psychological foundation of the nature of being social animals and to successively being full members of a society since the inability to read the minds of others would position us at a distinct evolutionary disadvantage in any social context where the success of our action depends in large part on the cooperation with other people. More importantly, severe deficiencies in the ability to read the others' minds (such as those exemplified in autism spectrum disorder) not only diminish the attitude of pursuing the individual interests in a social context but also is associated with fundamental restrictions in the ability to form basic social bonds as well as to join social practices such as the speaking of a common language.

The abilities and the repertoire of mental concepts such as beliefs, desires, and emotions that ordinary people without any specific, psychological training possess and use for understanding other people as minded creatures are collectively referred to folk psychology by contemporary philosophers and psychologists. The controversial point

about these supposedly capacities has always been whether our epistemic access to other minds proceeds in a manner radically different from that by which we acquire knowledge about other domains of investigation (theoretically or by simulating others) or whether it is structurally similar to such methods. The former view is taken by traditional proponents of the empathy theory and by current simulation theorists, while the latter view has been sustained by the naturalists theoreticians in the traditional philosophy of social science and by theory-theorists in the context of current philosophy of mind. According to the latter account, humans not only attribute mental states to other persons from a third-person perspective but also conceive themselves as minded creatures from a first person, egocentric perspective when trying to read other people's mind.

Although empathy has been regarded as the fundamental method for gaining knowledge of other people's mind since the beginning of the nineteenth century, in the last decades philosophers of social science have tended to regard it as a naïve method incapable of providing a satisfactory picture of human mind understanding. Only recently, with the discovery and study of the mirror neurons in the 1990's, empathy has started to regain importance in the field of cognitive science since it is considered the primary method for acquiring knowledge of other agents participating in the discourse. The renewed attention to empathy make the researchers considering it as the default method for understanding other agents within the folk-psychological framework³.

Taking a brief historical overlook on the role of empathy in relation to other minds, it is worth to mention that it has emerged as one of the primary concepts for solving the problem of other minds and, as a central concept in the discussion about the foundation of the human sciences, empathy is associated with the convergence of two rather independent philosophical traditions or discourses at the beginning of the twentieth century: the hermeneutic tradition of the "philological sciences" (Boeckh, 1886), which focused predominantly on the concept of understanding and the "philosophical aesthetics" which seems to be in charge of having introduced the concept of empathy. The debate about empathy was interested simultaneously in investigating the

³ For a full account, see Stueber K., *Rediscovering Empathy: Agency, Folk Psychology and the Human Sciences*, Introduction and 1st chapter, USA: MIT Press, 2010

psychological processes of our mindreading abilities, regarding questions of “descriptive epistemology” and in addressing the nature of the epistemic justification of our beliefs about other minds, namely issues of “normative epistemology”(Stueber, 2005, p.5).

The conception of empathy as a series of mechanism of inner imitation should be better being regarded as a form of basic empathy which is defined as series of mechanisms that underlie our theoretically unmediated, quasi-perceptual ability to recognize other creatures directly as minded creatures as well as implicitly identifying them as fundamentally similar to us. Because of the need to explain a person’s behavior entirely, it is necessary to turn first to our cognitive capacities in order to better explain other’s agency and his/her complex behavior and reasoning, applying in this way to a second order or enactive empathy. According to the enactive position, within the folk-psychological context, human beings tend to view each other first of all as rational agents. Since our capacity to consider other people’s mental states, including their thoughts, their true or false beliefs and their helpful or harmful intentions, helps us to navigate our social environment, mental state reasoning seems to function flexibly across the various domains (Seemann, 2012, p.315).

Can empathic emotion motivates altruistic action and an increased cooperation in social situations? According to neuroscientists, novel deeply stimulates the activity of the mirror neuron system so that empathy enters into the affective responses to fiction even though its role in the shaping of readers’ behavior and emotions has been debated for many centuries. According to cognitive literary studies, literary works have an emotional effect on both readers and the physical world in which they are located as each text stimulates differently an empathetic responsiveness: the fictional world seems thus to provide a comfort zone for readers feeling empathy without demanding them a reaction in the real world, a freedom from obligations that opens up the possibility for the development of various types of emotional reactions.

Empathy need not to be a conscious response, being a feeling that derives directly from the architecture of the human brain, equipped with mirror neurons; it is a feeling that arises from the social awareness of the presence of other human beings since birth and, as such, it is determined both by biological, hereditary features and by the influence

of the different cultures and external environment. Even though readers are perfectly aware that fictional characters cannot offer us help nor have any kind of interaction with our lives (so that there is no expectancy of reciprocation in the novel reading experience) there has been evidence, through empirical studies on empathy in a set of volunteers, that novel can influence pro-social behaviors and altruistic actions since responsiveness can be learned or developed and, as a consequence, novel reading might contribute to the cultivation of the empathetic response.

Since the 1980's, empirical studies on empathy has been conducted by psychologists who used short realistic fictions as techniques to study subject's actual or possible reactions toward a fictional world; by submitting the target to these texts, it is possible to measure the strength or weakness of the subject's empathy and successively classifying it in a special surveys of empathy scales. Recent evidence suggests that specific brain regions support multiple distinct cognitive components of mental state reasoning for moral judgment such as the initial encoding of the agent's mental state (Young and Saxe, 2008), the use and integration of mental states such as outcomes for moral judgment (Young, 2007), the emergence of spontaneous mental state inference when mental states are not explicitly provided in the scenario (Young and Saxe, 2009) and even post-hoc reasoning about beliefs and intentions to rationalize or justify moral judgments (Kliemann, Young, Scholz, 2008; Young, Nichols and Saxe, 2010; Young, Scholz, and Saxe, 2011). In particular, the functional magnetic resonance imaging (fMRI) was used in testing subjects while they were exposed to other's experience, such as pain. According to the results, a person while watching another person receiving a shock or traumatic experience, show an active response in the affective part of the brain (anterior part) which has been proved to respond to both real and imagined pain. Accordingly, a key brain region for moral judgment seems to be the RTPJ (temporoparietal junction) as individual differences in moral judgments were significantly correlated with individual differences in the RTPJ response.

Neuroimaging, lesions and behavioral studies both with humans and animals have been increasingly capable of characterizing the neural basis of empathy, providing new insights into the question of how we understand others' minds. Interestingly, recent

evidence supports a model of two separate brain systems for empathy: an emotional system and a cognitive system. The capacity to experience affective reactions toward the observed experiences of others or to share with them an associated feeling has been defined as emotional empathy. Emotional empathy may involve several related underlying processes, including, among others, emotional contagion, emotional recognition and shared pain, an assumption might be taken as the core of the so called Simulation Theory (ST). On the other hand, the term cognitive empathy describes empathy as a cognitive role-taking ability, which refers to the capacity to engage in the cognitive process of adopting another's psychological point of view, recalling in this way one key concept in the Theory-theory (Frith and Singer 2008). Suzanne Keen claims that the question of how and why empathy works in the bodies and brains of human beings can only be answered with theoretical speculations about its physiological substrate, though nowadays the fMRI- based research and recent advances in the study of mirror neurons have made researchers closer to the solution of the dilemma than they have been before (Keen, 2007, p.14).

As far as the neural representation of empathy is concerned, in recent years researchers have created a new set of experimental paradigms to study the neuroscience of empathy. Particularly, research on the cognition mechanisms underlying action has consistently demonstrated that watching another person executing an action can interfere with the planning and execution of an another action previously initiated by a subject. Findings of this type suggest the existence of a common coding both for one's own actions and for the perceived actions of others (Prinz, 2005). Additionally, cognitive neuroscience has demonstrated overlaps between regions representing one's own and others actions. The results of these experiments have lead researchers to speculate that social cognition is built on the automatic simulation of others' behaviors as the brain perceives others' actions and, through their simulation, infers their meanings.

A similar position is shared by the sustainers of the shared network hypotheses, according to which empathy for another person's pain, for instance, should activate components of the pain matrix and this activation happens to emerge in contrasting neural activity during trials depicting painful vs. non-painful situations: interestingly, in

experiments that include direct pain trials, it is possible to find overlaps between self and other pain representations. Although cue-based and picture-based paradigms furnish distinct patterns of data, recent meta-analyses provide strong evidence for the existence of a core network for empathy for pain: one image-based meta-analysis representing separate 9 studies (Lamm, Decety, and Singer, 2011) and two coordinate-based meta-analyses representing 32 (Lamm et al., 2011) and 40 studies (Fan, Duncan, Greck, and Northoff, 2011) found significant bilateral anterior insula (AI), dorsal anterior cingulate cortex (ACC) and anterior mid-cingulate cortex (aMCC) activity during empathy for pain across a variety of experiments conducted by different research groups. Participant self-reports of empathic states and traits support the role of these areas during the representation of indirect pain. Human beings are basically similar to one another and this assumption suggest structural similarities in both people's physiological and cognitive response system. Even though similarity alone is not enough to guarantee an universal empathic response among readers, an important point is that our survival depends on the effective functioning of our inter-personal mechanisms in a social context and thus empathizing with others confer an advantage in the struggle for living (Baron- Cohen et al., 2013, p.198).

As already mentioned, readers are perfectly aware that fictional characters cannot offer us help nor have any kind of interaction with our lives. Nevertheless, there has been evidence that novel can influence pro-social behaviors and altruistic actions. Thus, are there real effects of narrative in the lives of their readers? Which is the relationship between empathy and morality? According to the psychologist Martin Hoffman's account (2010), empathy can influence the way we treat others since empathetic response begins as an automatic, involuntary feeling that is aroused by hearing others speak or by seeing them acting in a precise way and then leads people to undertake a kind of simulation of the behavior/words perceived from the target, an empathetic response that seems to be the direct consequence of living in a social world and thus of being completely embedded in it. Accordingly, novels and fiction train us to care one another by portraying a series of social behaviors and their direct consequences, projecting outside, in a movie or in a novel, a depiction of social interactions and reactions to certain behaviors.

The relationship between reading and empathy justifies the cultural centrality of novels for the development of an healthy society and, since human beings have the capacity to engage intellectually and emotionally with imaginary worlds, narrative empathy is a phenomenon which stands halfway among aesthetic, psychology and philosophy. Particularly, the discipline of aesthetic has always been interested in empathy as a way of perceiving feelings and their manifestations as immersed in the matrix of social relations among humans. Since human beings long for safe places that allow them to experiment feelings and thoughts of various kind, fictional characters described as they were real are extremely useful since they open up the path for an emotional response to a scenario with a significant relation to reality. The fictiveness of character's mental states invites reader's participation in the empathy phenomenon in a playful way so that they can enjoy the literary experience and obtain pleasure while in real life other's actions and behaviors activate a series of defensive mechanisms. Thus the experiments of modern fiction with their emphasizing techniques aimed at rendering the characters' consciousness more vivid appears to be dramatically useful insofar the grasping of character's feelings and consciousness is concerned.

James Joyce (1882-1941) experimented stylistic features aimed at rendering the internal, bodily and psychological states of mind of his characters entering in this way more deeply into their consciousness. To provide an example of such techniques, the last chapter of the novel *A Portrait of the Artist as a Young Man* (1916) seems to contain all the hallmarks employed by the modernists to represent character's flow of consciousness as well as the shifting of perspective from one consciousness to another, making in this way the narration more fluid and providing a deep and multi-perspective analysis of the every psyche as involved in the narration:

-Look here, Cranly, he said. You have asked me what I would do and what I would not do. I will tell you what I will do and what I will not do. I will not serve that in which I no longer believe whether it call itself my home, my fatherland or my church: and I will try to express myself in some mode of life or art as freely as I can and as wholly as I can, using for my defence the only arms I allow myself to use- silence, exile and cunning.

Crainly seized his arm and steered him round so as to lead him back towards Leeson Park. He laughed almost slyly and pressed Stephen's arm with an elder's affection.

-Cunning indeed! He said. Is it you? You poor poet, you!

-And you made me confess to you, Stephen said, thrilled by his touch, as I have confessed to you so many things, have I not?

- Yes, my child, Cranly said, still gaily.

-You made me confess the fears that I have. But I will tell you also what I do not fear. I do not fear to be alone or to be spurned for another or to leave whatever I have to leave. And I am not afraid to make a mistake, a lifelong mistake and perhaps as long as eternity too.

Cranly, now grave again, slowed his pace and said: -Alone, quite alone. You have not fear of that. And you know what that word means? Not only to be separated from all others but to have not even one friend.

- I will take the risk, said Stephen.

And not to have any one person, Cranly said, who would be more than a friend, more even than the noblest and truest friend a man ever had.

His words seemed to have struck some deep chord in his own nature. Had he spoken of himself, of himself as he was or wished to be? Stephen watched his face for some moments in silence. A cold sadness was there. He had spoken of himself, of his own loneliness which he feared.

-Of whom you are speaking? Stephen asked at length.

Cranly did not answer (Joyce, 2000, pp. 208- 9).

This passage could be regarded as the core of the book itself since it contains Stephen's ultimate choice, which has been matured after years of internal struggle between what he feels and wants (namely to dedicate himself to art and to free expression) and what the external environment, which is constituted of his family and the institution of the Irish government and the Church expected him to become and to do, finally deciding to commit himself to art and freedom, leaving behind both his homeland, Ireland, and his family, even though he will always remain preoccupied for them during his voluntary exile. The dialogue between Stephen and his best friend Cranly is presented by Joyce through the mode of free indirect discourse: character's thoughts and feelings are displayed as they were really speaking within each other, so that the ultimate effect is that of having a portrait of their conversation as well as a report of what is going on in their minds while they are engaged in the discourse. In this way, both Stephen's and Cranly's feelings are represented so vividly that they appear to be as real persons to the reader, a reality effect which is strengthened by the shifting of focus between one consciousness to the other, an ability brilliantly mastered by the author.

For Virginia Woolf(1882-1941), as a modernist writer, fiction is the most democratic place where everything can be discussed and represented so that literature ought to challenge its readers by providing them different issues and materials upon

which they can actively reflect. The novelist's task, indeed, is to show human minds and feelings at work, requiring an empathetic effort to both the reader and the writer. As a consequence of this position, it can be said that character identification lies at the heart of readers' empathetic response, a condition that lead directly to the idea of the transformative power of reading fiction which stands at the core of the hypothesis of the pragmatic value of literature: knowing how to inhabit stories is the essential requirement of a mature life as well as the prerequisite for the recognition of our own identity.

Particularly, Woolf seems to be one of the authors that have experimented more with reader's Theory of Mind and the ways in which fiction builds on and tests human cognitive propensities to theorize about other's mental states at the beginning of the twentieth century. Significantly, her prose requires an active role of the reader in inferring character's feelings and thoughts, as it is exemplified by one of her masterpiece, *Mrs. Dalloway* (1925), in which she takes for granted in the opening pages of the novel that one of its protagonist's trembling, Peter, is not due to an illness but rather is a consequence of his deep feelings for Clarissa. In fact, as Lisa Zunshine notices (2007), our collective past histories as readers allow us to attribute immediately the cause of Peter's trembling to his excitement for seeing Clarissa again after a long time, enabling us to empathize directly with him by engaging our Theory of Mind device. In other words, the ToM abilities seems to underlie our interaction with the literary texts in a profound and complex way, so that every attempt to isolate each aspect of the interaction reader-text appears to be reductive and incomplete as far as the understanding of human mind-reading is concerned.

To provide an example of the ability of the readers to navigate the text at various levels of intentionality and, by doing so, also to grasp the representation of multiple layers of consciousness, a brief passage from *Mrs. Dalloway* should be provided. Halfway through the story, Richard, the protagonist's husband, meets an old acquaintance of his, Hugh Whitbread, at Milicent Bruton's house, a woman deeply interested in politics. In an cleverly written passage, Woolf reports the exchange of thoughts between Lady Bruton, Hugh and Richard who have been requested to help the woman to write a

letter to the editor of The Times which is meant to be very influential for Great Britain's future:

And Miss Brush went out, came back; laid papers on the table; and Hugh produced his fountain pen; his silver fountain pen which had done twenty years' service, he said, unscrewing the cap. It was still in perfect order; he had shown it to the makers; there was no reason, they said, why should ever wear out; which was somehow to Hugh credit, and to the credit of the sentiments and to the credit of the sentiments his pen expressed (so Richard Dalloway felt) as Hugh began carefully writing capital letters with rings round them in the margins and thus marvelously reduced Lady Bruton's tangle to sense, to grammar such as the editor of the Times, Lady Bruton felt, watching the marvelous transformation, must respect (Woolf, 1925, p.109).

As Zunshine noticed, in this passage there are various levels of intentionality expressed:

1. The makers of the pen think it will never wear out (1st level).
2. Hugh says that the makers of the pen think it will never wear out (2nd level).
3. Lady Bruton wants the editor of the Times to respect and publish her ideas (2nd level).
4. Hugh wants Lady Bruton and Richard to believe that because the makers of the pen think it will never wear out, the editor of the Times will respect and publish the ideas recorded by this pen (4th level).
5. Richard is aware that Hugh wants Lady Bruton and Richard to believe that because the makers of the pen think that it will not wear out the editor of the Times will respect and publish the ideas recorded by this pen (5th level).
6. Woolf, by saying "so Richard Dalloway felt" wants us as readers to recognize that Richard is aware that Hugh wants Lady Bruton and Richard to think that because the makers of the pen believe that it will never wear out, the editor of the Times will respect and publish the ideas recorded by this pen, which corresponds to the 6th level of intentionality (Zunshine, 2012, pp.31-3).

In this particular passage, Woolf requires her reader to make an effort not only in keeping track of the various levels of intentionalities distributed among the participants of the discourse but also she offers these various levels of embedded intentionalities through the description of body language, presenting Richard watching Lady Bruton observing Hugh

constructing his pen, unscrewing the cap and beginning to write, making the reader aware of the character's state of mind since Hugh cares about his writings and Lady Bruton admires the letters he writes. A depiction of each character's mental state is provided by offering not only a portrait of their internal feelings but also a simultaneous depiction of the external scene, as though the reader is at the same time inside the character's minds and in the physical environment in which the action develops. In this way, we are told that Lady Bruton feels that the editor of *The Times* will respect a letter written so beautifully; Hugh, while unscrewing the pen, thinks that because the pen will never wear out, the sentiments it produces would not and, interestingly, Richard seems to be able to represent both the minds of Lady Bruton and Hugh's. Thus this scene is challenging because the reader has to process an effort of keeping track of five- six order intentionalities. It is clear that, for understanding these complex activities, he/she must employ his mind-reading abilities since he/she is asked to oscillate rapidly between the simulation of different mental states, sometimes stretching the ToM abilities as well as the ability to process various embedded intentionalities beyond a cognitive comfort zone (i.e. beyond the 4th level) in order to understand whether the narrator whose mind he/she is simulating is reliable or not. This scene could be taken as example of the great effect that fiction has on us as readers but also as inhabitants of the everyday story world.

Given empathy's enormous importance, psychologists of various types have developed new tools and techniques, ranging from capturing data in infants to cross-species comparative studies, to explore empathic abilities and the same has done recently also the discipline of neuroscience. As far as the neuroscientific approach to empathy is concerned, since the spread of tools like fMRI for measuring task-related brain activity, which have been employed for the first time two decades ago, researchers have claimed that these instruments can be used for attempting to characterize the neural bases of empathy. The resulting avalanche of data deriving from the commitment of many researchers on the field has clarified some of the innumerable ways in which perceivers represent the experiences of social targets.

As a consequence, over the last 20 years neuroscientists have characterized several regions of the human brain that exhibit properties related to experience sharing, which

can be studied by applying to neural resonance. Interestingly, particular neural regions happened to respond to both perceivers' experience of a state and to their observation of targets experiencing the same state, engaging in this way the supposed "mirror neuron system" which encompass premotor, inferior frontal and inferior parietal cortex (Rizzolatti and Craighero, 2004). As previously mentioned, it was found that when executing and observing motor acts and when experiencing and observing non-painful touch, perceivers engage somato-sensory cortex while when experiencing pain and observing targets in pain, perceivers also engage somatosensory cortex but additionally manifest activity in regions related to the interoceptive and affective components, including the anterior insula and anterior cingulate cortex. Nonetheless, it is worth to say that empathy involves sharing not only at such low-level states, but also at higher, affective states, such as the understanding of the cognitive sources of those states responsible of provoking the neural response.

Briefly, since Premack and Woodruff's pioneering work with chimpanzees, reported in the famous article *Does the Chimpanzee have a Theory of Mind?* published in 1978, scientists have studied the ability of humans and some other animals to ascribe unique mental states to others and to utilize inferences about mental states during social interactions. Following the result of these studies, it has been observed that mental state attribution, in its various forms, has been one of the major topic of research for decades, with special attention being paid to its developmental trajectory (Flavell, 1999) and its breakdown in autism spectrum disorders (Baron-Cohen, Leslie, and Frith, 1985).

Chapter 2: Theory of Mind and the Issue Mind- Body

Theory of Mind, which is the ability to explain behaviors in term of the underlying states of mind, as defined by Lisa Zunshine (2010), is a mechanism that help us to empathize with other human beings. This ability, known also as mind reading, is thought to be the tool responsible for our interaction with the literary texts. Hence this capacity of attributing mental states to the people which are around us constitutes our genetic and normal way to engage and interact with the environment. Since fictional characters are potential containers of a great variety of thoughts, feelings and desires with which we engage in order to understand their behavior and read their mind for predicting their actions, they are extremely useful for further investigating our ToM abilities. Fiction engages actively our mind reading since it manipulates our tendency to interpret thoughts, feelings and aims of fictional characters which mirrors our own every day tendency to interpret other people's intentionality and attitude, a faculty defined as meta-representation.

Meta-representational ability, as described by Lisa Zunshine in one of her prominent book about the recent developments of cognitive studies (2012), is the ability of the mind that represent itself and its working which is a necessary feature of a normal brain since there is no self- awareness without meta-representation, which can be described as the cognitive mechanism that enables us to be aware of our goals as well as the intentions of other people. The failure in keeping track of oneself as the subject of representation of other people's mind and also of interpreting correctly their goals is a typical feature of a pathology called schizophrenia, even though it has to be clarified that our Theory of Mind is not perfect since we routinely misinterpret and misread other people's mind. Assuming that Theory of Mind is the feature that enables us to interact and empathize with literary characters leads to considering the works of fiction as engaging meta-representation par excellence since they are the vehicle through which human emotions are explored and showed, in all their greatness and peculiar aspects.

To provide an example of a literary character who is not able of keeping track of himself in the process of monitoring the sources of his representations in relation to his

fantasies about the world, the male protagonist of Samuel Richardson's epistolary novel *Clarissa* (1748), Richard Lovelace, seems to be worth a special attention. The entire story, being made of an exchange of letters among friends, strongly stimulates the reader's ToM since the events are represented "as if" they were a real interchange of dialogues among the characters aimed at interpreting and reading each other's intentions, thoughts and beliefs, even though these attempts systematically fail through the novel leading to the death of the female protagonist of the story, Clarissa, a beautiful and virtuous young lady manipulated by her greed family and then imprisoned and raped by the vile rake Richard. The failure of correctly interpreting other's thoughts and feelings in Lovelace's case seems to be due to a lack of meta-representation abilities since he, unlike other literary characters, seems not to be entirely aware of the games he makes on Clarissa as well as his perpetual cheating on her. Moreover, since the story is told in a first-person narrative, split between the two main protagonists, there is not an omniscient narrator which intervenes in the course of the narration for clarifying the characters' intentions and feelings as well as directing the events from the outside as a *deus ex machina*. As a consequence, the reader needs a great amount of time for realizing that Lovelace is dishonest not only toward Clarissa but he is also misleading about himself and his own intentions.

In this way, Richardson foreshadows the modernist attitude of employing an unreliable narrator in their stories, so that readers have necessarily to rely exclusively on themselves and their own ToM inference abilities to make hypothesis about characters' behaviors as well as making them continuously wonder whether they can or cannot believe in what it is told in the story. This uncertainty about what is supposed to be taken as true and what instead is just employed for deliberately misleading our interpretation and understanding, is a typical feature of modernist writings aimed at making the readers unsure about everything, arriving at the conclusion of the entire unreliability of both fictional and real worlds; in this way, the clear boundary between fact and fiction that had characterized literature up to the nineteenth century is definitely blurred.

In order to better understand the feeling of meta-representational uncertainty which recurs through all the book, divided in nine volumes, two particular passages from

Clarissa in which the mental workings beyond Lovelace's stratagems for entering into Clarissa's bedroom are represented, are worth to be provided:

I have just now heard that her Hanna hopes to be soon well enough to attend her young lady, when in London. It seems the girl has had no physician. I must send her one, out of pure love and respect to her mistress. Who knows but medicine might weaken nature and strengthen the disease? As her malady is not a fever, very likely it may do so- but perhaps her hopes are too forward. Blustering weather in this month yet- and that is bad for rheumatic complaints (Richardson, 1985, p. 604)

...But, what's the matter? What's the matter? What a double—But the uproar abates! What a double coward am I?—Or is it that I am taken in a cowardly minute? For heroes have their fits of fear; cowards their brave moments; and virtuous ladies, all but my Clarissa, their moment critical—But thus coolly enjoying the reflection in an hurricane!—Again the confusion renewed!—What? Where?—How came it? Is my beloved safe? Oh, wake not too roughly my beloved! (Richardson, 1985, p. 607).

In the first passage, Lovelace's plans for killing Hanna seems to be due to the man's loss of the ability to distinguish between his own fantasies (that of eliminate anyone who might help Clarissa to escape his plans) and the world outside his imagination, being in this way unable to meta- represent himself as well as correctly attributing mental states to the persons who inhabits his environment. The second fragment further reinforces the idea of Lovelace's inability to distinguish between his version of reality and the reality itself since he seems for a moment to doubt his own attitude ("Am I a coward?") then he immediately correct himself ("heroes have their fits of fear") and at the same time he seems to be really worried by the turmoil in the house even though he perfectly knows that the fire that has been lighted by his maids is a fake one, being it just one of his stratagems for sleeping with the untamable Clarissa.

According to Zunshine's interpretation of the scene, Lovelace needs to ally himself to Clarissa's state of mind, being as surprised and frightened by the fire as she is since this is the only way that allows him to carry on with his plan of seducing her by convincing first of all himself that they accidentally found themselves in the same room, rather than by a premeditated and cruel plan, so that he can free himself of his anxiety. Similarly, when he hears the uproar he fakes to be frightened too, simulating the natural reaction of a person being stuck by the clamors of a sudden burst of a fire and when the

confusion is renewed, Lovelace react again as if terrified (“What? Where? How came it?”). Evidently, this sequence of responses do not truly represent Lovelace’s real thoughts and feelings since he is aware of being the artificer of all the situation. Nonetheless, in both episodes as in many analogous ones through the story, there is no evidence that the man is consciously aware of his pretending and of his fictional attitude so that, by the end of the novel, the reader is left with the same sense of unsolved uncertainty about the actual workings of Lovelace’s meta-representational abilities, having the feeling that this is a narration guided and mastered by an unreliable narrator which remain so from the beginning to the end of the tale, offering no certainties to his readers in which to believe.

It is interesting to notice the definition attributed to novels such as *Clarissa* as analyzed by Blakey Vermeule (2010), where he sustains that Richardson, along with Henry Fielding, represents the beginning of the so called high mind-reading tradition in the English novel, which is a type of social novel characterized by the special way in which the various characters interact one with another with the aim of trying to understand the reasons behind other people’s intentions and behaviors. Significantly, at the center of the high mind-reading novel, there is a narrator who adopts the standpoint of an agent with full access to strategic social information and who distributes that information at different rates, positioning in this way some characters and sometimes the reader himself in a temporarily blinded and disadvantaged position, a technique masterfully developed by Richardson’s 1748 masterpiece and further improved by the modernists.

These examples taken together can be considered as an evidence that literature strongly stimulates the Theory of Mind mechanisms since novels and literary texts in general test the functioning of our mind-reading abilities, making us aware that even if we are trying to interpret literary character’s behavior, they will remain a fictional product of a human creative imagination. In the real world, these testing capacities are extremely useful to grant a social survival and narrative, on this purpose, seems to structure our sense of time, allowing us to predict, plan and retrospectively explain what has happened to us, a literary capacity that is indispensable to human cognition: whenever we attribute to a person certain thoughts or feelings on the basis of her/his

behavior, we are engaging our mind- reading abilities. According to cognitivist psychologists, this human capacity must have evolved in ancient times in our ancestors and it was adopted and improved over time in order to succeed in the struggle of survival. For this reason, Theory of Mind could also be described as the cognitive endowment of humans as a social specie.

As previously mentioned, the way in which our mind construct stories is part of an universal human process by which we learn to distinguish, organize and adapt to our environment. Literature is a vast, inexhaustible source of information about human nature, thus we read stories in order to see how the author and the text challenge and even re-shape our ideas about our own nature. By doing so, both literary writers and readers participate in a mutually, recognized thought experiment that interrogates the ways in which human mind engages with the world and with other minds. In order to comprehend the cognitive mechanisms that make imagination possible, it is useful to take into account human shared capacity to stimulate and develop imagination for understanding the mechanism that underlie our mental functioning, which is an highly specialized activity.

According to Bakhtin's theory (1919), as analyzed in Vermeule's book (2010), aesthetic perception must involve the reader shuffling to and fro between states of activity and passivity while engaging with a literary text. Writers write primarily with the aim of appealing to human mind and to represent its engagement with the world and its working in a complex society. In this way, according to recent study in neuroscience, our thought is not primarily linguistic but rather multimodal, involving both past memories and associations as present ones and our multimodal memories can be reactivated by a language that allows us to recreate past experiences, as fiction does. In fact, as we read, we can snap between imagining ourselves with the same experience of a fictional character and an outer view, changing continually our perspective and activating the power of our imagination. Thus, literature is keen to psychology in the degree that both desire to understand minds and what moves them, so that fiction can be perceived as a sort of thought experiment.

According to Paul Elkman (2007), it is surprising that humans can become emotional by reading, an activity which came later in the history of our species. He

hypothesizes that language changes into something else while we are reading, so that sensations, smells, sounds are treated by the brain like any other event that leads the automatic- appraisal mechanisms activated by the arousal of emotions. As a naturalist, he argues for the existence of an innate, biological emotion system that is more or less uniform across cultures. In particular, his studies on the observation of the faces of others expressing an emotion demonstrate an activation of the mirror neurons of the premotor cortex; successively, these neurons would then send a copy of their activation pattern to the somatosensory areas and to the insula, so that the activation of these areas, analogous to what occurs when the observer spontaneously expresses that particular emotion, can be considered as a phenomenon standing at the core of human understanding of the emotive feelings of others. As a consequence, this instantaneous understanding of the emotions of others, due to the mirror neuron system, is a necessary condition for the activation of the empathy mechanism which lies at the heart of our most complex, interpersonal relationships.

By contrast, the constructionists' approach argues that emotions are culturally specific, thus context and function determine emotional life, which vary over time. Following the concept of "emotionology", created by Stearns and Stearns in 1985, there is a collective emotional standard of a culture, as opposed to the experience of emotion itself. Thus "emotionologies" are systems of emotion terms and concepts deployed by a particular culture whose literary narratives contributes to form and reconfigure. Stories can serve as an instrument for distributing intelligence across time and space, so that narrative reinforce the supra individual nature of intelligence (i.e. minds as distributed among the participants of the discourse). Additionally, Daniel Hutto (2007) argues that it is through childhood engagement with narrative, which is built around belief- desire schemata, that humans learn the forms and norms of folk- psychology (ToM).

According to some philosophers, among which Richard Rorty can be mentioned, reading narrative fiction develops the ability to make psychological inferences about emotions, thoughts and motives others have in a certain situations so that individual and also cultural differences can be overcome. Thus empathy stimulates moral reflection and

the embracing of an habitual empathic attitude towards other humans which can be considered as the most evident example of a pragmatic, didactic role of literature.

Some narrative techniques that may contribute to the rise of empathy in novel reading have been classified by Suzanne Keen (2007), when she mentions some features of character's identification, which include depicted actions, roles in plot trajectories as well as various modes of representation of consciousness; then it is worth to mention narrative situation (including point of view and perspective), the nature of the mediation between author and reader and the position of the narrator which contributes to reader's identification with characters' goals and plans. Among all these, the already mentioned narrative techniques of free indirect discourse (and psycho narration) by providing characters' mental discourse as reported in the grammatical tense and person of the narrator's speech and by providing privileged information about characters' minds through the author's supposition about the character's thoughts and feelings, seems to be the narrative technique most likely to stimulate empathy in the readers.

The existence of an universal basis for emotion, shared by human beings and primates and supported by the study on facial expression of Paul Ekman, has helped psychologists to analyze and search for the rules that govern emotions across civilizations, discovering a common, universal structure of emotions that produces the same reactions among all the persons involved in a particular emotional experience, even though it is worth to underline that the theory of universalism seems to be an hegemonic one since it proposes a view of existence by which the experiences, values and expectations are in actual facts those of the Western civilization.

2.1 Theory of Mind and its implication in narrative

Our involvement in literature is both influenced by and influences various aspects of our "being-in-the-world", in Heidegger's terms. The relationship between literary studies and cognitive science has always been a problematic one and, over the years, various approaches to the study of literature at a cognitive level have been developed. Among

these, the so called “processual” and the “functional” approaches are the most famous and discussed ones. In the “processual” approach, the act of reading becomes the focus of attention instead of the classical focus on the interpretation of the text itself, often adopting empirical perspectives in order to understand how actual readers make sense of texts with the aim of discovering the underlying neurological and psychological processes involved in the process of reading literature. The “functional” approaches, on the other side, are aimed at explaining how the engagement with literary texts can play a role in a psychological way. Nevertheless, according to both theories, cognition and culture are intrinsically related and thus literary works can play an important part in the acquisition and development of various cognitive abilities. In this way, reading is perceived as a cognitive tool that enables readers to develop their capacities of organizing their present and past experiences, as well as improving their social capacity of ascribing complex mental states to self and others (ToM).

According to some researchers involved in the study of the relationship reader-text, such as Alan Palmer (2004), there is a basic affinity between actual and fictional minds. Literature is therefore seen as representing the workings of the mind and, in particular, literary works are useful to show the ways in which human mind engages with other minds in a fictional world. Hence, since fictional characters are understood as being similar to real people, scholars working in this sub-field believe that analyses of fictional minds and their behavior within a story-world can represent more clearly how real minds work. In this way, researchers gradually have become aware of the cognitive mechanisms that resides in their brain and arrived at the conclusion that the mind is always embedded in the literary text.

2.2 Some hints at the history of cognitive science and the re- discovery of the body in the studies on cognition

The term cognitive science did not appear on the scene of linguistic and psychological studies until the second part of the twentieth century as an umbrella term encompassing a large variety of disciplines such as neuroscience, psychology, artificial intelligence, linguistics and philosophy of mind. These disciplines were all united by a common aim, that of providing a rigorous, scientific account of cognition, which scholars active in these fields felt was lacking in the previous psychological and philosophical studies of the mind.

The cognitive sciences emerged in the 1950s as a critique of behaviorist psychology but until 1970s it is possible only to talk about a first phase in their development, called cognitivism. According to the cognitive perspective, the mind works in purely mechanistic/computational terms as an information-processing device according to which the brain transforms sensory inputs into abstract symbols which are internally processed and, subsequently, generate a motor output in a linear process. Mental activities were therefore understood as mere non-conscious symbolic computations. Since cognition was mainly conceived as the solving of abstract problems, the focus of the cognitivists' research was primarily on higher-level processes such as deductive reasoning and linguistic cognition.

The second phase in the development of the cognitive sciences has being defined connectionism and emerged in the 1980s. Here, the central metaphor of the mind is not the computer, but rather the artificial neural network in which a virtual system run on a computer and it is composed of layers of neuron-like units linked by numerically weighted connections, so that for connectionists cognition is still seen as computation but at a sub-symbolic rather than symbolic level; moreover, the importance of the relation between the cognitive systems and the environments in which they are embedded starts gaining an increasingly visibility in the connectionist account. Nonetheless, even if a sense of the dynamic interaction between the organism and the world slowly started to emerge, the role of the body in cognition continued to be ignored.

In the 1990s the importance of the body in cognitive studies started to acquire an increasing importance. As described above, in the previous paradigms the focus was on the brain, understood as an organ-machine capable of capturing aspects of the environment and subsequently recording them into internal or sub-symbolic representations, finally producing motor output based on these representations, making in this way cognition an intrinsically temporal phenomenon that can be described through a dynamic system theory. Consequently, an effect of focusing on the embodiment and situatedness of cognition is that the main processes to be studied in this view are not abstract reasoning as in cognitivism but rather subjective experience and consciousness, themes that started to rise interest among the scientific community at the end of the century.

One of the main issue which this approach attempts to clarify is the relation between organisms and the environments in which they live. Organisms are seen as autonomous self-organizing (autopoietic) systems coupled through recurrent sensorimotor processes with the environment and continuously striving to maintain their underlying structure under the pressure of external changing factors. Accordingly, the environment is not understood as a pre-specified, external domain, represented internally by the organism's brain but rather as a relational domain in which cognition is therefore conceptualized as an emergent process and as an active meaning-making procedure arising from the organisms' recurrent sensorimotor interactions with the world, instead of a passive tool gathering information from the outside world. Moreover, because of the self-producing or autopoietic organization of biological life already implied in cognition, this developing mind finds responsive expression in the self-organizing dynamics of action, perception and emotion integrating in this way a (neuro-)biological accounts of embodiment and cognition. The evolution of the human brain and the interpretations carried by particular cultures give us certain modes through which we have learned to see the world, thus the environment in which we are born and breed is extremely important in the process of cognitive understanding of how literary works are perceived by humans. Following this stand point idea, cognitive studies and cultural studies gradually started to mutually collaborate in order to arrive at a fuller understanding of human cognition.

Cognitive cultural studies is an interdisciplinary field that analyzes the relationship between the evolved human brain and the specific interpretations carried by the particular cultures since both aspects are relevant for the understanding of the literary process. Therefore, it is important to link the individual cognitive development with the functioning of the social institutions in which it is embedded. Culture undoubtedly influence our ways of seeing the world and language, as a cultural, inherited way of expressing one's thought and abilities, has a fundamental cognitive value. Literary experience feels so real because it engages the extremely complicated dynamics of social interactions in which everyone is continually engaged in the everyday life and, since minds are always grounded in the discourse and they have their own ways of articulating themselves, they can be rightly perceived as an ongoing construction in the discourse. Accordingly, the world is continuously created by learning new ways of seeing it because events constantly modify our point of view, so that the construction of reality is contiguous with the construction of fictional worlds: our way finding abilities depend on an emotional connection to the environment and people acquire their status of psychological beings just by participating in the discourse, being intrinsically social creatures by nature.

These embodied and affective interactions with the others emerge from the beginning of life and develop specifically during infant's first year, remaining crucial in all later social interactions. In fact, long before children reach the age of four (the supposed age for acquiring a Theory of Mind) several sensorimotor, perceptual, emotional and non-conceptual embodied practices, which constitute our primary access for understanding others, are already well developed. According to the phenomenological tradition, in most of our social interactions we have a direct understanding of another person's intentions because they are explicitly expressed in their embodied actions and expressive behaviors. In other words, before we are in a position to think about what other persons believe or desire, we already have specific perceptual understandings of what they feel. Thus, before theorizing, simulating, explaining or predicting others' mental states, we can already interact with others and understand them

in terms of their expressions, gestures, intentions and emotions; such understanding is fast, automatic and highly stimulus-driven.⁴

As already mentioned above, the role of the body in cognition was practically ignored in classical (cognitivist and connectionist) cognitive sciences. Only in the early 1990s, the body was rediscovered and its importance started to be reevaluated. Descartes's distinction between *res extensa* and *res cogitans* firstly originated an understanding of human beings as divided between a physical, visible body and an ethereal, invisible mind. Accordingly, Cartesian dualism intends to radically eliminate the body as Descartes argues that:

My mind, by which I am what I am, is entirely and truly distinct from my body, and may exist without it [...] I am not only lodged in my body as a pilot in a vessel. I am besides so intimately conjoined, and as it were intermixed with it, that my mind and body compose a certain unity⁵

Thus the following questions arise spontaneously: is the mind entirely and truly distinct from the body? Am I entirely and truly distinct from my body? While the first question concerns the mind and the body considered as abstractly-generically built and Descartes argues that they are conceivable in isolation from each other, which is the core of his dualistic conception of the *res cogitans*, the second question regards the mind and body of a concrete-specific subject and Descartes himself argues that the mind–body is a blended compound. From a physiological point of view, the constitution of one's own body as acting body necessarily brings into play mechanisms created for exploring the external world. In the traditional philosophy of mind, the mind is described as having two principal parts corresponding to metaphysics and epistemology. The epistemological part had, in particular, two critical issues: the problem of other minds and the problem of self-knowledge (Goldman, 2006, p.5). As far as “the other minds” problem, the main question seems to be how can one know that there are other minds at all. By trying to

⁴ For further explanations on the topic, see Elkman, P. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, USA: Holt Paperbacks, 2004

⁵ Descartes' account in Schmicking D; Gallagher S. *Handbook of Phenomenology and Cognitive Science*, pp.181-2, New York: Springer Press, 2010

answer to this question, Descartes focused on the respective essences of mind and matter claiming that the essence of mind is to be a thinking, un-extended thing while the essence of matter is to be extended but unthinking. Hence, mind and matter cannot be identical⁶.

In the last century, many philosophers inquired into the essence and nature of mind, even though in doing so they often continued to consider the concepts of mind and mental states as two separate aspects. The idea of the existence of a folk psychology, accordingly, seems to be deeply relevant since it is defined as a philosopher's label for the practice of making sense of intentional actions by appealing to the notions of an agent's beliefs and desires. This is an ability we have inherited from our ancestors and proponents of the hotly debated dominant theory for understanding folk psychological abilities, known as Theory-Theory (TT) and Simulation Theory (ST), typically sustain that our ancient cognitive endowment considers the physical mind and its mental states from the same perspective. To accept any of these views, or some hybrid combination of them, is to accept that our folk psychological abilities essentially have a basis of biological inheritance. As an evidence of this assumption, it can be observed that normal human children belonging to various cultures come to understand actions in terms of reasons using the same basic mentalistic frameworks (Hutto, 2008, pp.10-8).

2.3 The problem of other minds and the Cartesian dualism

The first grand wave of twentieth century cognitive science and philosophy of mind sought and still seeks to show how, against Cartesian dualism, cognition is meant to materialize in the brain. A new wave is arising from this view that studies cognition not merely as "embrained" (Damasio, 2000; Collins, 2000) but also as embodied in a much stronger sense, as inseparable from and shaped by the concrete extra-cerebral structures and dynamics of the body by taking into account its embeddedness in the natural and social world. Where traditional accounts thinly conceive the brain as a computational device, the new wave solidifies and enriches this view by showing how the body itself is

⁶ For further explanations see Descartes, R. Discourse on the Method and Meditations on First Philosophy, USA: Hackett Publishing Company, 1999

integral to the brain's controlling computations. Therefore, the mind not only extends itself to the external environment that do cognitive work but it also extends over and above the brain, to the morphology, dynamics and temporality of the environment and to our evolved and emotional attitude toward it (Gallagher et al., 2009, p. 235).

This link between body and cognition is supported by recent neurological literature, which shows that the higher cognitive processes involving the activation of higher brain structures, in fact also activate and are activated by lower brain structures known to be active in proprioceptive and visceral processes and in body control. Extended consciousness is, then, the capacity to be aware of a large range of entities and events (i.e. the ability to generate a sense of individual perspective, ownership and agency over a larger compass of knowledge) than those involved in the core consciousness, which is the generic sense of the self. Extended consciousness thus allows human organisms to reach the very peak of their mental abilities by sensing the minds of others.

How is it possible to have knowledge of other minds? In the twentieth century the problem of other minds moved to the top of the philosophical discourses since its repercussions are felt in many debates within philosophy of mind and cognitive science. Both the so-called Theory-theorists and Simulation-theorists generally assume that the mental phenomena of others are and will remain unobservable. The two accounts differ mainly in the explanations they offer about how we ascribe such unobservable phenomena. Generally speaking, we can say that most supporters of the Theory-theory of mind endorse the theory argument, whereas many Simulation theorists employ arguments from analogy: namely, in the Theory Theory's account, we infer and attribute these unobservable states by deploying a Theory of Mind which sustains that our ordinary psychological beliefs literally generate a theory that humans construct while they are children in order to explain their own behavior as well as the behavior of others. According to Simulation Theory, instead, we attribute mental states to others by projecting ourselves into their shoes and by creating similar states in ourselves intended to correspond to the states of the others. In other words, The Theory of Mind approach to social cognition stipulates that we need to read mental states in order to engage with

others as intentional subjects analogous to ourselves while the Simulation Theory claims that the knowledge of ourselves as human beings is obtained through the same cognitive tools that we use to understand others' behavior. In conclusion, because there is no way to access directly to other's mind, the relationship self- other minds continues to be a problematic one and no hypothesis can be taken as satisfactory since another's mental life remains essentially hidden and unobservable.

Speaking of the rediscovery of the body in cognition leads inevitably to considering that the specificity of our anatomy and skeletal structure can have influential consequences for our cognitive skills since it is our general anatomical constitution that defines our capacities to interact with the environment. Aristotle was one of the first thinkers to claim in his Poetics that imitating is a co-natural act that human beings perform since childhood, an ability that confers them an immense advantage for survival; moreover, humans differ from the other animals because they are the most imitative creatures among all the species and thus produce their first acts of understanding properly by imitation.

A neuroscientific explanation of this seemingly innate capacity for imitation started to emerge in the 1990s after the discovery of resonance system in the brain, which is in actual facts a cluster of the so-called mirror neurons located in the pre-motor cortex and Broca's area, which is activated both when we engage in specific motor actions and when we see other people performing the same actions. Nevertheless, understanding the nature of imitation in terms of the existence of an innate resonance system seems to be reductive and inadequate to provide a satisfactory explanation for human imitation which has proved to be different from instances of mimicry in other primates. The empirical studies on the brain that has been conducted since the 1990's had a great importance as far as the discovery of the mechanisms involved in the process of empathy and Theory of Mind are concerned, so that an analysis of the various hypothesis of ToM is useful to better understand the developments that cognitive science has made in the last decades.

2.4 How to deal with other minds: Theory- Theory and Simulation Theory accounts

- Theory- Theory of Mind (TT)

According to the Theory-Theory of mind (TT), in order to understand another person's mental state, a theoretical standpoint must be adopted which has its roots in the so called folk psychology (i.e. the set of expectations we develop during the formation of our person that concern how people behave and why they behave as they do). According to this theory, our complex mental states are implicitly defined by a common theory and developmental psychologists argue that such capacities for mindreading emerged only later in the development of the child since it has been proved that children under four, as well as autistic children, repeatedly fail the so-called false-belief tests; in other words, they have not yet developed a Theory of mental states (ToM) before that age. Moreover, the developmental research on mindreading started to develop only in light of two important discoveries in the 1980s: first, normally developing 2-year-olds are able to engage in pretend play and, secondly, it has been observed that children undergo a deep change in their understanding of the psychological states of other people somewhere between the ages of 3 and 4, as exemplified especially by the appearance of their ability to solve a variety of false-belief problems.

The finding that mentalistic skills emerge very early and in a way relatively independent from the development of other cognitive abilities, led some scholars (for example, Simon Baron Cohen, Jerry Fodor, Alan Leslie) to conceive that age as the end of the endogenous maturation of an innate Theory of Mind module. For instance, Wellman (1990) has argued that around age 4 children become able to pass the false-belief tests because they move from an elementary "copy" Theory of Mind to a fully "representational" Theory of Mind, which allows them to acknowledge the explanatory role of false beliefs. Another Theory-theorist who endorses a domain-general conception of cognitive development is Josef Perner (1991). In his view, it is the appearance of the ability to meta-represent that enables the 4-year-olds to shift from a "situation theory" to a "representational theory" and thus to correctly pass the false-belief tests (Gopnik and

Wellman, 2012). According to these studies, children both seem to have abstract structures of thought (theories) and at the same time to learn them through imitation.

Then, how could such highly structured abstractions be learned? In Simon Baron-Cohen's research, aimed at studying the disturbances of social relations in the autistic population and to respond to the question whether autistic subjects can read the mental states of others (namely, have a ToM mechanism), the use of the false belief task was decisive in settling this issue. The task typically involves a subject whose knowledge of a given state of affairs (e.g. the contents of a Smarties box) surpasses that of a third party who is temporarily absent from the experimental situation and whose beliefs regarding the given state of affairs is thus unknown even though he/she is asked to predict the outcome of the situation. It was found that autistic subjects systematically fail to predict false beliefs, so that this particular task has been regarded as a proof for sustaining that autistic children cannot access other minds or that their experience of other minds is characterized by a pervasive mind-blindness, while the normal functioning of social life depend on the ability to see other minds, namely to have a mind-sight.

Everyday psychology provides explanations and predictions of intentional actions by appealing to what the person thinks, knows and expects, activity that also includes reasoning about the origins of mental states and, specifically, children's developing understanding of beliefs, desires, perceptions, intentions and emotions is a complex matter. Accordingly, it is interesting to observe that two-years old seems to have an early theory including a conception of mental states such as desires and perceptions, but this theory is incorrect from an adult perspective, since it does not posit the existence of the representational mental states (i.e. prototypically beliefs). Hence 3-years-old is seen an intermediate phase where children develop a non-representational account of belief and they also begin to understand representational aspects of their well-developed notions of perception and desire.

During the first phase (up to 3 years), children begin to conceptualize mental representations as a kind of supplementary hypothesis while in a later phase, beginning around the age of 4, they start to reorganize their central explanatory theory when, having to realize what other participants in the discourse think, they understand that their

representation of the world is different from the world itself. More specifically, even at 2 years, psychological knowledge seems to be structured largely in terms of two types of internal states, which are desires, on one hand, and perceptions, on the other. Desire and perception alone provide examples of the two basic categories of folk psychology which are also the two types of theoretical constructs that Searle (1983) calls "world-to-mind" and "mind-to-world" states: while an understanding of desire incorporates an early knowledge that what is in the mind can change what is in the world, an understanding of perception, on the other hand, encompasses an early knowledge that what is in the mind depends intrinsically on what is in the world. Additionally, both desire and perception, as theoretical constructs, work to explain action but may also be dissociated from any particular actions that an agent may perform. By the age of three, children begin to show signs of a more sophisticated mental elaboration of situations and natural language can help us to explore these abilities since, before this age, children make extensive and appropriate use of terms for desire and perception but this understanding is largely non-representational, being it modeled on the basis of desires and perceptions. Interestingly, as the third year progresses, children begin gradually to realize the representational aspects of desire and perception which may be construed either non-representationally or representationally. As a consequence, in the adult Theory of Mind desire and perception are perceived mainly as representational, so that what people want and see is not the thing itself but rather the thing as represented.⁷

As far as the role of mindreading in adults is concerned, it seems to process in a specialized, cognitive module and it is also commonly supposed to have a critical role in fast-moving social interaction and competition enabling us, for instance, to understand what a speaker is talking about on the basis of their eye gaze, so that we should expect mindreading to show at least some key characteristics of a modular process (e.g. Fodor, 1983, 2000; Leslie, 2005) and to be relatively effortless since it requires little work from

⁷ For a complete account on the topic see also Hirschfeld, L. A; Gelman, S. A. Mapping the Mind: Domain Specificity in Cognition and Culture, pp. 341-50, USA: Cambridge University Press, 1994

memory or executive control. While children's ability to read other minds has been tested through the false belief tasks (also known as perspective-taking tasks), to analyze adult's mindreading seems to be more problematic. A possible solution to this problem is to test mindreading concepts that are more subtle or complex, where there might be plausibly more activity in the adults' ToM. Thus one useful approach to study mindreading in adults seems to be that of analyzing the nature of adults' reasoning by examining the heuristics and biases that are apparent in their everyday judgments and decisions. Such studies may pose mindreading problems where the expected answer is uncertain, such as judging how another person will make a difficult perceptual discrimination or how interpret ambiguous verbal messages. In tasks with a clear, straight answer, such as predicting the incorrect search of someone with a false belief about an object's location, researchers may ask participants to rate their certainty about their answer. Findings from these studies suggest that adults' judgments about others are likely to be conditioned by interference from their own perspectives, a phenomenon generally labeled "egocentric bias" (Nickerson, 1999) and "reality bias" (Mitchell et al., 1996). These studies taken together claim valuable insights into the cognitive basis of mindreading by suggesting that unbiased, non-heuristic mindreading may require time and cognitive effort.

Another approach for studying mindreading in adults involves the adoption of tasks that require simple judgments about beliefs, desires and visual perspectives that are conceptually similar to those used in the studies on young children. Following methods widely adopted in cognitive psychology such as questionnaires and cognitive tests, these tasks enable the measurement of adults' response time across many repeated trials. For instance, in one early study of this kind, German and Hehman (2006) presented adults with multiple trials of a belief-desire reasoning task, which showed adults to be slower to make judgments when a character had a false belief rather than a true belief and even more slower when she/he had a negative desire rather than a positive one. Nonetheless, it is worth to say that these tasks are simple and repetitive, thus they lack the subtleness, sophistication and uncertainty of much everyday mindreading. Even though, they appears to have two significant advantages: firstly, they make possible to posit deeper questions about the component processes of mindreading and, secondly, they need simple

mindreading concepts similar to those required in most developmental and neuroscientific studies. German and Hehman (2006) presented adults with short stories from which they had to infer a character's belief and desire in order to predict their action. Notably, they found that adults were slower on trials that required thinking about false beliefs and negative desires, compared with true beliefs and positive desires, which is the same pattern of relative difficulty observed in 3-6-years-old children while analyzed on developmentally sensitive tasks. This finding clearly suggests that psychologically relevant parameters, such as the strength of belief and desire, influence the effort adults must place in solving mindreading problems (Baron Cohen et al., 2013, 74-5).

Further studies have been conducted in order to find a way to isolate the different components of mind-reading one from another. For instance, Apperly, Warren, Andrews, Grant and Todd (2011) adapted the belief-desire paradigm obviating the need for participants to infer the character's mental states by asserting these states directly. In one of their study, participants read sentences describing which one of two boxes presented contained some hidden food, which box the character thought contained the food (his belief could be true or false) and whether he wished to find or avoid the food. All participants had to do was hold this information briefly in mind and then combine it to predict which box the character would open (e.g. if he had a false belief and a desire to avoid the food he would open the box containing the food on the mistaken belief that this box was empty). Although participants no longer had to infer the character's mental states, the valence of his belief (true vs. false) and desire (positive vs. negative) resulted to influence their performance.

By using a rather different paradigm in which participants made rapid judgments about the simple visual perspective of a character standing in a room, Samson, Apperly, Braithwaite, Andrews and Bodley (2010) were able to study the demands of mindreading inferences independently from demands associated with holding such information in mind or using it for further suppositions. They found that participants were slower to judge the character's perspective when it was different from the participants', suggesting

that, like young children, adults experienced an egocentric interference while making judgments about someone else's perspective.

In summary, recent experiments on adult mind-reading show that it is possible to separate the various mind-reading components while ToM mechanisms are active; these components include the inferring of mental states, holding the information in mind and using it at the proper time. Additionally, these findings suggest that each process may contribute to making mind-reading a cognitively effortful process. Although there is plenty of evidence that adults routinely and rapidly make inferences that go beyond the information given in a reasoning or comprehension task, it is equally clear that these inferences are not obligatory or stimulus-driven, but are instead dependent on participants' motivation for devoting cognitive resources to a particular task.

In another study, Apperly et al. (2010) presented participants with video scenarios involving a target character who is thought to have either a true or a false belief about the location of a hidden object. These stimuli clearly offered mindreading inferences about the character's beliefs, but the instructions only required participants to keep track of the location of the hidden object. Interestingly, critical data came from questions presented at unexpected points in the videos, which showed participants to be relatively fast at answering questions about the location of the hidden object (which they were instructed to track) but significantly slower to answer matched questions about the character's false belief. Generally speaking, the dominant view among researchers studying mindreading in infants appears to be that infants' mind-reading abilities would be essentially continuous with the full-blown mindreading abilities of older children and adults (Apperly, et al., 2010, pp.77-9).

As far as the neural basis for understanding mind reading is concerned, although a number of brain areas are commonly said to constitute a mindreading brain network, notably including the already mentioned areas of the medial prefrontal cortex (mPFC), temporal poles and bilateral tempo-parietal junction (TPJ), the debate has been reduced to the attempt at understanding which areas of the brain are mainly involved in this process. Saxe and colleagues (e.g. Saxe and Kanwisher, 2003; Saxe and Powell, 2006) firstly identified the brain areas that showed a very neat contrast between an activation observed

while participants have to respond to short stories concerning false beliefs vs. false photographs and those which were most selectively activated during other judgments about mental states, including personal preferences, personality and physical appearance. These studies consistently find that right-TPJ is the site in which the largest and most selective activation for mental states is recorded, whereas other areas of the “mindreading network” appear to show a lower activation or to show an activation just for other kind of judgments. This pattern has led the researchers to suggest that right-TPJ is the domain-specific neural basis of the mindreading process.

Accordingly, among the different areas in the neural regions implicated in mind reading, the prefrontal cortex seems to be the most prominent, as a study by Jenkins and Mitchell (2009) has further demonstrated. In this study, the authors presented participants with mindreading tasks that varied whether the scenarios concerned a character’s mental states or their preferences and whether a specific mindreading inference was relatively clear after having provided the context of a story. Supplementing other works, this study found that r-TPJ was selectively sensitive to the difference between scenarios involving mental states rather than preferences, whereas mPFC was not selectively sensitive to this particular differences. In sharp contrast, mPFC was sensitive to the difference between scenarios involving clearly-specified rather than ambiguous inferences, whereas r-TPJ shows no activation at all. These results are due to indicate that having the ability to take someone else’s perspective is not nearly a sufficient proof of a reliable mindreading mechanism, hypothesis that has received support from several functional magnetic resonance imaging (fMRI) and event-related potential (ERP) studies which exploited enterprises that manipulate demands on self-perspective inhibition within the context of a mindreading task.⁸

⁸ For the complete account of the research see Baron-Cohen S. et al. *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*, pp. 77-90, USA: Oxford University Press, 2013

- The Simulation Theory of Mind (ST)

According to Simulation Theory (ST), instead of drawing on folk-psychological inferences during the process of interpersonal understanding, we initially simulate others' behavior with the aim of inferring their mental states. The discovery of mirror neurons has provided an evidence that the very same cortical substrates are activated when actions are both executed and perceived in others; this data indicates that a neurally biological mechanism for coupling action and perception mediates our capacity to share motor goals and motor intentions with others. Moreover, neurophysiological and neuroimaging studies have shown that specific parieto-frontal circuits are devoted to the encoding of observed situational features in terms of action potentialities, notably both in monkeys and in humans. This process, in neurophysiological terms, implies that the same neurons must be able not only to encode the motor acts (e.g., hand-grasping) they control but also to respond to the situated visual features supporting it or even demanding those motor acts to imitate the action observed (Seemann, 2009, pp.435-6).

One alternative to Theory-Theory, then, is that the child's understanding of the mind is much more closely linked to evidence than to theory. In this view, understanding appears to be more similar to empirical typologies and generalizations than theories, thus the grasping of others' mental state is strictly related to the ability to imaginatively project ourselves into their situation rather than being built in a sort of innate, genetic theory pre-coded in the brains since ontogeny; such forms of projection seem not to exploit mainly our theoretical abilities but rather our motivational and emotional resources. There are some controversies regarding whether we are conscious or not of our peculiar cognitive mechanisms but nonetheless, in such a view, the actions, intentions and emotions of the other become much more transparent and much closer to the surface, rendering the gap between self and others as proposed in the Simulation approach less theoretical and more immediate, as it draws on neurological processes rather than theoretical ones.

According to ST, our third-person mindreading ability does not consist in implicit theorizing but rather in representing the psychological states and processes of others by

mentally simulating them and by attempting to generate similar states and processes in ourselves. Thus, the same resources that are used in reasoning about our own psychological states are employed to provide an understanding of the psychological states and processes of the simulated target. Notably, an empirical account of knowledge of this type is congenial with the classic philosophical view for the understanding of the mind. This latter position, originally identified with Descartes' theories but widely accepted up to the present days, further reinforces the idea that we have some form of privileged perceptual and introspective access to our own mental states as well as to that of others'.

2.5 Empirical studies on narrative and Theory of Mind: how we make sense of others in everyday life through a narrative process

Over the past two decades, work in cognitive neuroscience has become especially rewarding as the development of new techniques to observe the brain in terms of its structure and function nowadays allows researchers to link a certain behavior that is observed, clinically or in an experiment, not only to the presumed mental causes of that behavior, but also to specific indices of brain structure or brain activity. Moreover, the increased progress in neuroscience have enabled us to learn a great amount about how the brain works since technologies such as brain imaging have generated new experimental paradigms responsible of having allowed a clinical analysis of the specific regions of the brain that become active while exposed to stimuli of various kind.

For instance, circumscribed brain areas that has been damaged (i.e. lesions) can be observed while the patient is still alive through positron emission scan (PET) or a functional magnetic resonance imaging scan (fMRI). The striking thing about these new technologies is that such scans can be used not only in neurological patients but also in humans without brain diseases. A specific prediction concerning the activity of a circumscribed region during the performance of a particular mental task is useful to understand which regions of the brain are mainly involved in specific tasks, allowing us

to establish a progressively more detailed account regarding the relationship between certain aspects of mind that correspond to precise behaviors.⁹

The science of brain imaging is complex and is certainly not just a matter of taking a picture of what is going on inside the head since the generation of images of neural processing using non-invasive technology has made possible a variety of experiments that depend on reports about the experience of the experimental targets. The results of neurological observations and of neuropsychological experiments have revealed many aspects that are considered the starting points for further research on the studies on the brain. For instance, some aspects of the processes of consciousness can be related to the operation of specific brain regions and systems, thus opening the way for the discovery of the neural architecture which supports consciousness. Accordingly, it seems logical to assume that the ideas of the phenomenological philosophy can significantly contribute to the development of the cognitive sciences.¹⁰

Since the discovery of mirror neurons in the premotor cortex in monkeys that respond to the goal-directed actions of others (Rizzolatti et al. 1996), a great amount of researchers have made various attempts at exploring whether evidence can be found for the existence of a similar neural structure in humans. A standard paradigm used in a number of case studies is to compare an observation condition where participants watch the activity of another person with a an execution condition, where participants perform the action on cue, and an imitation condition where participants perform the action while they observe another person performing that action. Taken together, the results of these studies have revealed that premotor and parietal cortices in humans show mirror properties similar to those examined in individual neurons of monkeys since these areas have been perceived as active both while performing the actions or observing the same actions on others and even more active than in either of these conditions when these actions are both observed and imitated (Zlatev, 2009, pp.50-4). Moreover, even though

⁹ For further explanations on the studies, see Damasio, A. *The Feelings of what Happens: Body and Emotion in the Making of Consciousness*, USA: Mariner Books, 2000

¹⁰ For a fuller account on the discussion see Gallagher S., Zahavi, D. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*, USA: Routledge, 2008

mirror neurons in the pre-motor area may be insensitive to the difference between self and other, focusing mainly on the goals of actions that both monkeys and young infants are able to represent, it can be speculated that left and right parietal regions represent agents in intentional relations and therefore these are the very areas responsible for the distinction self-others as intentional agents. In other words, when the participant is the original source of the action, the right hemisphere is dominant while when the participant is imitating the other, what is dominant instead is the left hemisphere (Decety and Grezes, 2006).

In considering imitative tasks, it should be noted that imitation of new actions requires skills that do not appear in monkeys but only in humans when they are in a full blown state during the second year of life, while the infant is forming its concept of an intentional agent. Indeed, two-years-old find particularly fascinating to engage in mutual imitation, where they enjoy leading and following each other in new intentional actions. This play behavior can be interpreted as the exercising of possibilities made available in this specific period of their life by developments in the use of the intentional schema, in order to both understand self and other and to discriminate the self from the other even in particular contexts, where both actors are performing similar actions (Zlatev et al., 2008, p.53).

Several studies conducted by Iacoboni and his colleagues indicate that mirroring of expressed affective states may be an important basis for understanding emotions in people: fMRI brain imaging of participants were taken while they were engaged in observing or imitating a variety of emotional expressions depicted in photos (Carr et al. 2003; Dapretto et al. 2006). According to the results, observing and imitating emotional expressions in others appears to activate regions involved in those emotional areas responsible for the perception of the self, namely the amygdala and the insula, but also the pre-motor area and STS. Since the observation and imitation condition had similar pre-motor relations to action, this evidence can suggest that an implicit if not explicit matching of emotional expression is involved in emotional empathy. Indeed, from a developmental perspective, the transmission of emotional states from one organism to another is the original basis of emotional sharing (Zlatev et al., 2008, p. 56).

As shown above, a considerable amount of research has been devoted to establishing the neural basis for the understanding of the more complex intentional relations characteristic of Theory of Mind. Thus the focus of studies using ToM tasks is aimed at determining the brain regions functionally involved in the interpretation of complex stories of social interaction that are visually or verbally presented and in attributing mental states to the individuals involved in these stories. To sum up, two brain regions have been widely recognized to be the most active in the process of story understanding by using brain imaging studies when compared to control conditions involving comparable processing of non-ToM stimuli: the temporal/ parietal junction (TPJ), including the superior temporal regions (STS) as well as inferior parietal regions (Decety and Grezes 2006) and the medial prefrontal cortex (MPFC). The TPJ is supposed to be an area in which complex visual stimuli, often involving biological motion and social interaction, are analyzed or represented perceptually and semantically. In particular, it is worth to mention that while IRs focused only on intentional actions of a single agent, the TPJ is crucial for social interactions and for interpreting more complex mental states. Hence, in terms of IRT, the left side can be considered the responsible for representing the third-person information about IRs of one or more organisms involved in simple or complex objects and interpersonal interactions. Significantly, even in monkeys, this area appears to have individual neurons that are sensitive to both the eye direction of an observed target and in a target's behavior involving another object. Comparable findings with humans, concerning more complex IRs involving intentions, have been made using fMRI (see Pelphery et al. 2004). Thus this region is supposed to be primarily involved in emotion concerning multiple agents.

The second region of importance for the ToM tasks is the MPFC (medial prefrontal cortex), an area that appears to be devoted to “decoupling” (Leslie, 1994) or to create a second order representation of IRs that can be attributed to individuals. Some recent studies using false belief tasks found that damages to the left TPJ produce a specific deficit in false belief reasoning about others, so it appears that a functional TPJ on the left site is necessary for false belief reasoning. By contrast, it appears that the impact of brain damages in frontal regions is less specific and more diverse, including effects on

performance on tasks involving executive function but not on those that require a ToM tasks. Indeed, in one of the patients analyzed presenting a frontal damage, there was evidence that problems occurred only on false belief tasks that required the inhibition of first-person knowledge but not on those for which the participant was required to make suppositions about third-person actions (IRs). This result is congruent with other findings which suggest that the executive function associated with frontal activity may be necessary for differentiating between mental states of self and other and thus for attributing distinct mental states to various individuals.

Considering the two main regions described above which have been involved in research with complex ToM tasks, the results provide evidence for the Theory- Theory (TT) approach to social understanding since the TPJ provides third-person behavioral analysis of animate activity, while the MPFC decouples or represents abstractly IRs, presumably in a theoretical or conceptual system. Since the same behavioral analysis and conceptual representation could be applied to self and others, the main conclusion to be drawn is that complex ToM tasks involve two main regions of the brain, namely a posterior one associated with perceptual representation and an anterior one associated with meta-representation of the perceptual representations.

As far as the autistic children involved in these experiments are concerned, it was shown that they were as able to imitate emotional expressions as other children, but the imaging results suggested that the means that they used were different: the typically developing children replicated the results of the adult study, where the mirror neurons of the pre-motor and insula areas were involved in the observation and imitation of emotions along with other areas, while in autistic children these mirror neuron were not involved in the same degree. Moreover, the involvement of these areas during imitation was inversely related to the severity of autism by which they were affected. Conversely, other areas, such as the left anterior parietal and the right visual association areas, were more involved for autistic than for normal children. Accordingly, it was suggested that these areas served as substitute in the process of imitation in the autistic targets analyzed instead of the usual one involving the pre- motor and insula areas.

These results, combined with other findings, support the notion firstly proposed by Barresi and Moore (1996) namely that the main reason why autistic people have difficulty in ToM tasks as well as in emotion understanding and imitation is that they do not match and integrate first-and third-person information through an intermodal and intentional schema; as a consequence, they acquire and develop an independent first-person (or egocentric) and third-person (or allocentric) Theory of Mind. The Dapretto et al. study probably provides the best confirmation for this theory since according to them it is the lack of matching of these two types of information, through an intentional schema, that stays at the heart of the problems of social understanding in the autistic individuals (Zlatev et al., 2008, pp.57-61).

From a developmental perspective, others are present from the beginning of life since when we first perceive objects, we learn to recognize them in joint attention as well being always involved in some purpose: the world that we perceive is not only physically contextualized (embedded) but also socially contextualized, so that when the “I” perceive someone, it is not only a matter of capturing the presence of another living body situated in a general “there”, but also it is perceiving him/her as a perceiver of the same world that the “I” inhabits. Jean Paul Sartre gives a dramatic formulation of this idea when he writes that the relationship among things undergoes a fundamental change when the “I” experiences somebody else observing the very same things he is observing. In his words:

Thus suddenly an object has appeared which has stolen the world from me. Everything is in place; everything still exists for me; but everything is traversed by an invisible flight and fixed in the direction of a new object. The appearance of the Other in the world corresponds therefore to a fixed sliding of the whole universe, to a decentralization of the world which undermines the centralization which I am simultaneously effecting (Sartre, J.P. 1993, p.255).

From this perspective, it is clear that as soon as the other appears on the scene, the world appears to the “I” as alienated so that the presence of the other has the mere function of revealing to the “I” the complexities of the external environment.

Literature is both a cognitive tool and a key to understand how the mind works. Following recent developments in cognitive cultural studies, Isabel Jaén (2012) mentions a series of methodologies borrowed from the natural sciences such as questionnaires,

brain imaging and computerized text analysis that have been employed for studying the relationship between literature and the readers and, in particular, the effects of literary reading on self and society. Neurological approaches to literature, such as neuro-imaging, have been intrinsically useful for studying the brain in the last decades and it was employed primarily for understanding the neural underpinning of reading stories. For instance, the functional magnetic resonance imaging (fMRI) was used in testing subjects while they were exposed to other's experiences (such as pain) or while they were reading a particular moving text. As mentioned in other sections, a person while watching another person receiving a shock or traumatic experience, shows an active response in the affective part of the brain (anterior part) which is sensitive both to real and imagined pain.

Similarly, Holland explores the neurobiology of being transported while we are reading literature, by describing the effects of the phenomenon called "poetic faith" sustaining that so long we are transported in a work of art, we have no disposition to act upon what it is represented, hence no mechanisms for testing reality are evoked. In this way, our tools for testing perception are more open to feel and to join more intensely the world as presented by the verbal arts (Holland, 2009, pp. 108-11). Joseph Carroll (2004) enriched this view by sustaining that the genetic mediated disposition of human nature interact with specific environmental conditions, including some particular cultural tradition, so that literary works reflect and articulate the vital motives and interests of human beings as living organisms. Thus we need to consider both biology and culture in order to postulate a correct theory of literature and the dynamics between biological universals and cultural specifics play an important role in the comprehension of the literary phenomenon since both the study of universal, cultural and historical particularities are necessary to correctly understand the way in which we engage with literature.

Theory of Mind is helpful to explain our involvement in the fictional world since it seems that while immersed in reading, we use the same cognitive tools we employ in everyday life to explain other people's behavior. In this way, literary characters appears to be useful since they tell us something about human emotions and the way they work and

affect us as living beings: through our engagement in fiction, it seems that we train the cognitive abilities to read minds and run simulations of social experience. Since the mimicking of literary characters is essentially emotional, readers by simulating characters' behavior or simply sharing their emotions, show an empathetic response toward literature, so that stories may be viewed as an human adaptive strategy to maximize their chances for survival. Accordingly, literature is widely said to have a social value since literary characters are part of a social network and, as real people, they exhibit a series of collective thoughts and behaviors. About that, it is worth to say that the brain is the site where biology engages culture to produce the mind and its manifestation, which is the text. In this way, a complex and still actual debate about the literary discourse and the mind conceived as the cultural product of a given historical period has been established.

Lisa Zunshine in the field of cognitive cultural studies considers the blending between cultural and cognitive studies in order to arrive at a more coherent understanding of the human embodied mind and the texts it produces as its natural and cultural manifestations. Accordingly, the dialogue between the empirical science and cognitive, classical approaches to literature which has been conducted in the last decades is meant to give a decisive turn in cognitive literary studies in order to achieve a multi-level and multi-approach attitude toward the engagement with the fictional world in order to arrive to a better understanding of the interactions among human beings.

Since literature is a complex, high order mental process involving belief, desires and aesthetic questions, the main aim of cognitive neuroscience is to understand how these mental phenomena are in relation to the physical processes in the brain. This achievement has been accomplished by using behavioral techniques such as PET, ERP and fMRI in order to observe the brain at work while it is involved in the literary process. The results have shown that both language and stories have universal patterns since stories are one of the most important methods for imprinting systems of values on large target audiences, so that the processes of narrative comprehension have been supposed to be universals and unique to humans. Moreover, the mechanism underlying these processes are not likely to be found in a single region of the brain but rather in its

heterogeneous architecture. According to behavioral techniques, various part of the brain seems to be involved in the process of story comprehension and each part seems to respond to specific narratological needs, such as:

- Closure, i.e. the need to have an ending;
- Time sequencing, i.e. the before- after sequences necessary for giving the sense of consequentiality;
- Personification, i.e. the attitude of attributing mental states to others in order to predict their behavior;
- Identification, i.e. sympathy for some characters with a consequent empathetic response toward them;
- Ritualistic cues, i.e. items that demands shifts in attention from the physical to the imaginative world and vice versa.

As a consequence, to represent the empirical approaches to literary reading requires a blending of concepts and theories from literary studies and the use of socio-psychological methods such as questionnaires and self-reports, so that reading literature can be defined as an high-level, complex activity involving cognitive mechanisms as well as neuro-aesthetic mechanisms of response; it is an highly specialized system that presupposes a lot of training and involves a relation between systems rather than involving a close module, making in this way the narrative phenomenon a process determined foremost by processes relating to an internal structure, which is supposed to be universal.

As far as the empirical studies on Theories of Mind are concerned, even though currently there are several research techniques that can be used to connect characteristics of EEG recordings with cognition, just two methods have been used in Theory of Mind research. The first and most common, is the event-related potential (ERP) technique in which the ERP is recorded in time related to the presentation of a particular stimulus that requires a Theory of Mind judgment; the ERP signals from each trial are then averaged to capture the stable, reliable characteristics of the brain response. The second method concerns the characteristics of resting-state or baseline recordings of EEG with Theory of

Mind performance: in this method, a primary strength of the ERP technique has the potential to capture the signal in fine temporal detail (typically a 1–4 millisecond resolution) and successively to capture the neurocognitive events that are associated with the processing of a particular class of stimulus. The two methods described above represent the challenge of applying brain imaging methodology to the study of Theory of Mind, even though it is not entirely possible to determine precisely when someone has made an inference about someone else’s mental state since the application of the ERP method to standard Theory of Mind tasks, as most cognitive neuroscience techniques, typically requires participants to endure at least 40 trials of a particular condition type to maximize the signal averaging benefits, condition which is rarely available.

Because the same recording methods and experimental paradigms can be used with children of different ages as well as with adults, EEG/ERP methods provide a clear opportunity for looking at how the neural mechanisms that are associated with various aspects of Theory of Mind reasoning change over time. Recently, Lindsay Bowman and colleagues (Bowman, Liu, Meltzoff and Wellman, 2012) adapted the paradigm for comparing the neural correlates of belief and desire reasoning with eight and nine years old children. These findings showed that, like the adults, there was a left lateralized late slow wave effect associated with reasoning about both beliefs and desires; the functional enrollment of right parietal areas for reasoning specifically about beliefs may be a later developing feature of the neural correlates of Theory of Mind reasoning.¹¹

A peculiar aspect of Theory of Mind is what has been generally defined as the decoding of the others’ mental states. Even though mental state decoding and belief/desire reasoning cooperate to render accurate judgments about others’ mental states in everyday situations, they may also rely on fundamentally distinct neurocognitive processes. For this reason, the ERP technique is extremely useful for the study of mental state decoding because it is relatively easy to control the moment in which someone engages in a precise mental state deciphering. There is one interesting study that was

¹¹ The complete description of these experiments is described in Baron Cohen, S. et al. *Understanding Other Minds. Perspective from Developmental and Social Neuroscience*, pp. 140- 41, USA: Oxford University Press, 2013

conducted by Sabbagh, Moulson and Harkness (2004) who employ an ERP paradigm while adults were shown a mental state term (e.g. “desiring”) or a sex term (e.g. “female”) that was followed by a picture of the eye region of a face. In this study, the participants’ task was to determine whether the mental state term or the sex term was an accurate description of eye picture. In the results, it was found that the cognitive operations that are associated with making judgments related to gaze direction appear later in the ERP record, having a more anterior- temporal frontal distribution.

A second literature that is relevant to mental state decoding concerns emotion recognition (i.e. the ability to identify other’s emotions based upon facial expressions) that can be thought of as a special case of mental state decoding insofar it involves making a judgment about someone’s mental state based upon visible perceptual information. It is interesting that there are many ERP studies of facial emotion recognition, most of which investigate differences in neural responding to the six basic emotions (fear, anger, happiness, sadness, disgust, surprise), with a particular interest in understanding whether the neural response to negative emotions differs from positive or neutral ones and how the emotionality of the face affects the early perceptual processing of facial features and configurations and how these effects are modulated by attention (Baron Cohen et al., 2013, p.124).

The studies that has been described so far have in common an experimental use of PET or fMRI techniques with the aim of trying to localize the neural regions in which some aspects of Theory of Mind reasoning reside. In other words, the attempt was to compare the neural activations elicited by mental state reasoning with those elicited in some control conditions (i.e. reality reasoning or photograph reasoning) with the aim of identifying special regions of the brain that are engaged in reasoning about mental states. Any given ERP effect concerns more than one cortical source even though methods for accurately localizing cortical sources of ERP effects are improving rapidly due to better recording techniques that allow an higher spatial density sampling of scalps and an increased understanding of the physical and anatomical constraints that govern how electrical signals propagate from neural dipoles to the scalp. EEG is readily applicable to a wide variety of sensitive populations (e.g. children, individuals with autism and adults)

and this property has made it a reliable technique for understanding the neurocognitive underpinnings of various cognitive processes. Since the studies on false belief reasoning and mental state understanding more generally have rendered such similar findings, they can be taken as further evidence of the importance of the role of EEG in understanding the neurocognitive systems underlying Theory of Mind.

An interesting study conducted by McCleery and colleagues (McCleery, Surtees, Graham, Richards and Apperly, 2011) adopted a task developed to examine the cognitive dynamics of visual perspective taking using an ERP paradigm. In this task, participants were shown a picture of a stage set in which three walls were visible (left, back, and right). On the walls of the room, there was some number of “disks” (black dots) disposed in some arrangement and all visible to participants; within the room, there was a character (avatar) who had either a full or partial view of the total disks. On some trials, the participants’ task was to say how many disks participants themselves see and on other trials they have to say how many the avatar sees; for some time, participants saw the same as the avatars, whereas other times there was a discrepancy in the results. This project allows to draw some comparisons between the neural mechanisms associated with perspective taking (through the self-other comparison) and the neural mechanisms associated with resolving conflicts between two response options (when participants and avatars see the same vs. different numbers of disks). Although there were some complexities in the results, the findings were clear in showing that right (and to some extent, left) posterior regions were the first to show a difference in making judgments about one’s own vs. another’s perspective, with differences emerging on a slow positive ERP component, similar to the posterior slow wave that was present in the ERP studies described above. Interestingly, this component was slower when making judgments about others’ perspective than when making judgments about one’s own. Later in the ERP, a late slow wave of 600–800 milliseconds post-stimulus was employed for analyzing the lateral frontal component, which showed differentiations in instances where individuals had to make a judgment while the participants and the avatar saw the same vs. different numbers of disks. Integrating these findings with others from the ERP and broader neuroimaging literature, the authors concluded that a posterior temporal parietal system is

associated with computing differences in visual perspective whereas the frontal system is associated with selecting the appropriate response from the conflicting options.

To the extent that there are individual differences in Theory of Mind skills, it might be expected that different patterns of tonic cortical activation would provide insight into the neurobiological bases of these differences, even though there has been only one attempt in doing so: Sabbagh and Flynn (2006) explored whether individual differences in healthy university students' resting state EEG might be associated with mental state decoding. Using the "reading the mind in the eyes" task (as did Sabbagh et al. 2004) the results from individual differences and group analyses showed a tonic activation at right mid-frontal part while there were no significant effects in the posterior parietal regions. These findings suggest that stable, individual differences in mental state decoding might be associated with tonic activation of the right frontal regions (Baron Cohen et al., 2013, pp.124-8).

In adults and children, a large body of work in affective neuroscience has used measures of EEG activation to better understand individual differences in various aspects of affective style, associated with social competence in everyday situations (see Davidson, 1998). This work might offer an opportunity to provide an unexpected link between affective style and Theory of Mind. To sum up, in the last decade the studies that have employed neuroimaging tools to investigate the neural basis of ToM has raised increasingly, reaching 400 studies up to 2013 even though most of the fundamental questions about how our brains allow us to understand other minds still remain unanswered.

Over the course of development, human children make a remarkable discovery: other people have minds at the same time similar and different from their own which means that people can see the world from different perspectives, having different desires and acquiring different knowledge and beliefs about the things in the world. Thus children learn very early that other people's minds contain representations of the world which are often true and reasonable but which may be strange, incomplete or even entirely false. Developmental psychologists historically focused on one key transition in this developmental process, namely when and how children come to understand false

beliefs. Assessing the understanding of false beliefs has been taken to be a good measure of ToM capacity because it requires a child to understand both that someone can maintain a representation of the world and that this representation may not find validity in the true state of things in reality.

The first neuroimaging studies of ToM is known for having used versions of false belief tasks: adults, lying in positron emission tomography (PET) or magnetic resonance imaging (MRI) scanners, read short stories describing a person's action and successively were asked to explain that action. Notably, the results revealed increased levels of blood oxygen and glucose uptake (indirect measures of metabolic activity) in a small group of brain regions (left and right TPJ and mPFC) and also medial parietal cortex and more anterior regions of the superior temporal sulcus (STS), down to the temporal poles. Activity during the false belief task is clearly far from sufficient evidence that these brain regions have any role in understanding other minds.

For example, in an early PET study, Fletcher Happe, Frith, Baker, Dolan, Frackowiak, et al. (1995) told participants that they would be reading different kinds of verbal passages. Just before each item, the participant was told what kind of story was coming next. If it was a "mental" story, participants were instructed that it was important to consider the thoughts and feelings of the characters and then shown a story revolving around someone's mental state. If the story was a "physical" story, participants were first instructed that thinking about thoughts and feelings was irrelevant and undesirable, and then it was shown a control story. After each story, participants were asked to silently answer an action-explanation question, such as "Why did the person say that?" The striking result was that glucose consumption increased more in the supposed Theory of Mind brain regions while people read the mental stories than when they examined the physical stories (Baron- Cohen et al., 2013, pp.130-4).

Thinking about thoughts and feelings can be also manipulated by changing the task while holding the stimuli constant. In their first experiment, Spunt, Satpute and Lieberman (2011) showed to the participants of the study some pictures of simple human actions (e.g. a person riding a bike), and instructed them to silently answer one of three questions: why the person is doing the action (e.g. to get exercise), what the person is

doing (e.g. riding a bike), or how the person is doing it (e.g. holding handlebars). These questions require successively less consideration of the mind of the person and, correspondingly, showed successively less ToM region activity. Spunt and Lieberman (2012) replicated the result using a similar paradigm with brief naturalistic film clips of facial expressions of emotions. Participants judged either how the person is expressing their emotion (e.g. “looking down and away”, a control task) or why she is feeling that emotion (e.g. “she is confused because a friend let her down,” a mental task). Again, ToM regions were recruited more when thinking about “why” than “how”.

Differences in ToM are easier to find in young children, who are still learning how to understand other minds. Interestingly, two recent studies suggest that getting older and thus getting better at understanding other minds, is not associated with more activity in ToM brain regions rather with more selective activity: children from 5 to 12 years old seem to manifest an adult-like neural activity when listening to stories about characters’ thoughts and feelings. What is different is that ToM regions in younger children show similarly high activity when listening to any information about characters in the story, including the characters’ physical appearance or social relationships (Saxe et al. 2009; Gweon et al. 2012), whereas in older children and adults, the ToM brain regions are recruited only when listening to information about thoughts and feelings (Saxe and Powell, 2006; Saxe et al. 2009). This developmental difference in the selectivity of the ToM brain regions is correlated with age, but also with performance outside the range of difficulty in the ToM tasks (Gweon et al., 2012).

The correlation between a sort of neural selectivity and behavioral task performance remains significant in the right TPJ, even after aging. The strongest evidence that some brain regions are involved in a cognitive task is to show that disrupting those regions leads to disruption in task performance. Trans-cranial magnetic stimulation (TMS) offers a tool for temporarily disrupting a targeted brain region and, as some studies show (see Young, Camprodon, Hauser, Pascual-Leone, and Saxe, 2010) this disruption appears to leave moral judgment overall intact, but impairs people’s ability to integrate considerations of the character’s thoughts into their moral judgments.

Another way to study the necessary contributions of a brain region to ToM is to work with people who have suffered permanent focal (i.e. local) damage to that region, typically due to a stroke. Samson, Apperly, and colleagues (Apperly and Butterfill, 2009; Apperly, Samson and Humphreys, 2005; Samson, Apperly and Humphreys, 2007) have conducted a series of interesting studies using this approach. These authors tested a large group of people, with damage to many different brain regions, on a set of carefully controlled tasks and found that these individuals, who have passed all the control tasks (e.g. measuring memory, cognitive control, etc.), still failed to predict a character's actions based on their false beliefs due to their neural impairments.

The last decade of studies with neuroimaging discovered a significant, replicable functional regularity in the human brain regarding regions that showed an increased activity when participants think about thoughts. These regions may be real cortical areas or parts of a larger topographical maps, but in either case, the understanding of other minds seems to be a major organizing principle of responses over different cortex regions. Additionally, the function of these regions is not due to complete a task, rather to transform some class of input into some output; notably, the class of input has something to do with thinking about minds instead of bodies or abstract representations: whenever we are thinking about thoughts, there are neurons gathered in spatial proximity (i.e. into a part called region) that fire because they have related computational properties that are distinct from the computation properties of neurons in the surrounding cortex areas. As a consequence, we may conclude that Theory of Mind involves a series of correlated brain areas rather than being inscribed in a single, circumscribed region (Baron Cohen et al., 2013, pp. 142-4).

Chapter 3: Debates on Theory of Mind and Alternative Accounts to the Dominant Approaches to Social Cognition

3.1 The Theory of Mind debate

Social psychologists have investigated mindreading since at least the 1940s. Heider's and Wilfred Sellars' book *Empiricism and the Philosophy of Mind* (1956) had suggested that our grasp of mental phenomena does not originate from direct access to our inner life, but rather is the result of what has been called Folk Psychology, which is an ability that we acquire during our lives and is influenced by some form of enculturation. Their speculations contributed to the formation of the so called Theory-Theory and, since the end of the 1970s, primatology, developmental psychology, cognitive neuropsychiatry and empirically-informed philosophy have been contributing to a collaborative inquiry into Theory of Mind.

Theory of mind (ToM), as previously mentioned, is our capacity to attribute mental states to ourselves and others and to interpret, predict and explain human behavior in terms of mental states such as beliefs, desires and intentions. This phenomenon has been studied extensively by developmental and comparative psychologists and, more recently, by neuroscientists and cognitive psychologists since it is a complex, second order activity. In this way, mindreading or mentalizing should be defined as the root of our social self, a capacity that is considered uniquely human and which is likely to reside at the heart of social interaction and communication. For this reason, since its inception, the modern study of mindreading has involved close collaboration among psychologists, philosophers, literary critics and neuroscientists.

When Premack and Woodruff (1978) asked whether the chimpanzee has a Theory of Mind it was the commentaries of three philosophers (Bennett, 1978; Dennett, 1978; Pylyshyn, 1978) that led to paradigms involving the use of perspective differences tests such as the false belief. Even though some scholars claim that our understanding of people is based on inferential and quasi-scientific mechanisms, others instead argue that

we initially simulate others' behavior in order to infer their mental states. According to developmental psychology, for instance, by engaging in storytelling practices, children learn to understand the goals and intentions of others as well as theirs as active parts of the discourse. In this way, a kind of inter-subjectivity starts to arise, which can be described in terms of the acquisition of a self-other meta-perspective or meta-representation.

As previously stated, the two most important divisions within the Theory of Mind research are the Theory- Theory of Mind (TT) and Simulation Theory of Mind (ST). Cognitive sciences gradually shifted from a classical, computational perspective of the mind to a recent understanding of the mind as intrinsically embodied in the discourse and in the environment. Much of the interest in and research on Theory of Mind over the last twenty-five years resulted from the hypothesis that this uniquely human capacity might be conceived as the foundational tool of human cognition. Wellman, quoting Peter Carruthers' words, notices that:

The mind contains a number of innately channeled conceptual modules, designed to process conceptual information concerning particular domains . . . [such as] a naïve physics system, a naïve psychology or "mind-reading" system, a folk-biology system, an intuitive number system, [and] a geometrical system for re-orienting and navigating in unusual environments (Wellman, 2014, p.663)

Following this line of thought, the mind is perceived as a set of innate mental modules that are responsible for the existence of an genetically pre-coded ToM. However, as a series of recent studies on the field have shown, Theory of Mind requires a more complex and more developmental picture of the human brain since it is a foundational cognitive system that is developmentally dynamic, not static as it was thought in the first formulation of the theory. Philosophers (e.g., Stich, 1983), psychologists (e.g., Wellman, 1990), and anthropologists (e.g., D'Andrade, 1987) now agree that our everyday mentalistic understanding is organized around three large categories of mind and behavior: beliefs, desires and actions. Therefore, Theory of Mind is a an organized system of interconnected concepts and implications that includes perceptions, emotions and urges intersecting with beliefs, desires and actions.

In a first phase, there were two alternative proposals of ToM: according to the first, our inference mechanism is based on an innate, mentalistic module that reads actions in terms of underlying mental states, a module that becomes available early in life and that can be dramatically damaged in certain individuals, leading to the social-communicative impairments seen in autism. This nativist/ modular position is frequent in the literature on Theory of Mind debate, as Peter Carruthers' quote shows well. Such findings reveal that typically developing children evidence ToM understandings early in life, even in infancy, and normal children engage in these practices across all cultures on a consistent, maturational timetable, despite wide differences in languages and social experiences. Additionally, individuals with autism are specifically impaired in ToM understandings, notably in false beliefs (in comparison to other individuals, including other impaired individuals) and these mechanism which are localized in certain specific brain regions are clearly impaired in autism.

On the other hand, a contrasting and different proposal suggests that ToM arises from the sheer accumulation of social information from others: adults tell and show children how to think about and interact with people as mental beings and children consequently absorb the ideas of those around them. According to such a position, human infants show an early attention to communicative interaction and, crucially, these interactions provide social knowledge about others. Exposed to the knowledge of others by observing and interacting with them, children collect increasingly complex, mentalistic understandings of persons, lives and minds. Although they approach the facts in different ways, the modular and experiential-learning alternatives agree that Theory of Mind changes over time, thus they define it as a developmental process (Wellman, 2014, pp.7-8).

Taking an historical overview on the term, the definition of Theory of Mind entered the developmental literature via at least two different routes. First, Wellman (1979, 1985) working in the area of meta-cognition, used it to refer to the child's conception of human cognition. Second, and probably better known, Premack and Woodruff's (1978) used the term in their investigation of primate cognition, when they defined Theory of Mind as a

system of inferences that can be used to predict behavior by attributing mental states to individuals (Astington, 2005, pp.3-5).

David Premack and Guy Woodruff in their famous article *Does the Chimpanzee Have a Theory of Mind?* defined Theory of Mind as an ability to attribute mental states to self and others, as well as to interpret, predict and explain behavior in terms of mental states such as intentions, beliefs and desires. In the most precise use, Theory of Mind is a domain-specific, psychological structure composed of an integrated set of mental-state concepts employed to explain and predict people's actions and interactions that is reorganized over time when faced with counterevidence to its initial predictions. This assumption marked the beginning of the experimental methodology on human ability of mentalizing about others beliefs, thoughts and desires. Following this view, without a mirror mechanism, we would still have our sensory representation since a certain pictorial depiction of the behavior of others is already pre-coded in our brains.

From the very beginning, the relationship between the self and others is understood as problematic, being characterized by an immense gap as the mental states of others cannot be directly perceived, but only inferred from their external manifestations (for instance in facial expressions, gestures or patterns of behavior). To sustain this hypothesis, various fMRI studies provided evidences that, as in monkeys, the human mirror mechanism located in the parieto-frontal circuits is involved in understanding the goal of the observed motor actors. Moreover, a direct evidence that this peculiar mechanism is involved in imitation was provided by the previously mentioned fMRI study by Iacoboni et al. (1999). Therefore, it would seem that the mirror neuron system is indispensable for that sharing of experience which is at the root of our capacity to act as individuals but also as members of a society; this form of imitation, which is both simple and complex at the same time and which regards both the learning of verbal and gestural communication, presuppose the activation of specific mirror circuits which are the same tools responsible for our capacity to appreciate the emotional reactions of others since it appears to be correlated to a particular group of neural areas characterized by mirror properties.

Emotions, like actions, are immediately shared so that the perception of pain or grief or of disgust experienced by others seems to activate the same areas of the cerebral cortex that are involved when we experience the same emotions ourselves. In this way, when an individual observes an action performed by others, a potential motor act is evoked in his brain areas which is to all effects similar to that spontaneously activated during the organization and effective execution of the action. Therefore, whereas some scholars claim that our understanding of people is theoretical and inferential, others argue that we initially simulate others' behavior in order to understand and predict their mental states.

Accordingly, as far as the ways in which mindreading is perceived and elaborated at neural level, there are three competing hypothesis: by theorizing, by rationalizing, or by simulating which, in turn, give rise to the different position on the Theory of Mind debate. The first approach, namely Theory-theory, argues that ordinary people construct or are endowed with a naive psychological theory that guides their assignment of mental states. The second approach defined as Rationality Theory sustains, on the other hand, that the ordinary person is a rationalizer; in the third approach, Simulation Theory claims that ordinary people grasp their targets' mental states by trying to replicate or emulate them. Thus this latter view strongly sustains that mindreading includes a crucial role for putting oneself in others' shoes and this role-taking may even be part of the brain's design to generate mental states and resonate with them so that mindreading can be considered as an extended form of empathy (Goldman, 2006, pp.1-6).

Michael Tomasello (2008) claims that, as far as the acquisition of a full Theory of Mind is concerned, apes and young children both perceive the external world in the same basic way, by comprehending the actor's goals as internal representations of the state of the world they wish to bring about and the content of the other's perception is something potentially different from their own (Goldman, 2006, pp.47-8). Thus he stresses that while reading other minds, we are deeply involved in a process of representation of the content of other's mental states, a process helped by the engaging with narrative practices, such as conversations and storytelling, which are activities rich in images that activate a first, naïve folk psychological attitude.

According to Teresa Bogdan (2009), a Naive Theory of Mind is best understood as an assembled cluster of abilities that enables the grasping and representing of the mental states of others such as gazing, seeing and feeling, the sorts of mental states that young humans understand prior to their fourth birthday. Concerning the more complex states of mind, such as opining, hoping and intending, Bogdan's proposal is that children engage in a prolonged process of need-driven negotiations with adults, using a tranche of Naive psychological skills, so that through the course of these dealings, the child's mind is eventually reshaped and converted into the sort of mind to which folk psychological categories genuinely apply. In this account, the child begins his or her inter-subjective career without the capacity for a truly predicative thinking. Yet, after interacting with linguistically competent adults, he or she becomes able to think autonomously in their early days and, because they are only operating with a Naive Theory of Mind, children are not representing the contents of other minds in fully propositional terms but rather they just register the adult's intent as some sort of image without attributing it any meaning (Hutto, 2007, pp.69-70).

Nevertheless, not all theorists working on research on Theory of Mind abilities believe that such principles exist in the minds of mind minders. In particular, Simulationists maintain that no Theory of Mind principles exist in the minds since all the relevant work of mind reading can be done by directly manipulating one's own mental states: when predicting actions, one draws primarily on his cognitive resources in order to get an insight into other's mind, so that the core process is not theory mediated and the similarities that exist in the reasoning processes and those of the targets are exploited; this vision clearly obviates any need to represent something before understand it, as for the Theory-Theory account (Seemann, 2012, pp. 326-7).

Inspired by this idea, Nichols and Stich (2003) posit an Early Mindreading System that provides a more explicit and detailed account of the rudimentary cognitive machinery needed for basic mind minding. The key assumption in this theory is that practical reasoning, planning and updating are all required for the ordinary navigation of the world. These mechanisms are co-opted for the purposes of mind minding and have been in place independently and presumably long before any ToM capacity emerged on

the scene. Moreover, because the features of the background psychological profile are generic and constant across the cases of such responding, there is no need to represent the complex psychological background of the target. Thus elementary mind minding can be understood in terms of interactive, unprincipled embodied engagements rather than in terms of theoretical ones. As far as a Full Theory of Mind is concerned, instead, its devices are supposed to be built into our species, so that a Theory of Mind is supposed to have been present in Homo Sapiens 100 or 150,000 years ago or at least before the dispersion from Africa. This hypothesis invites us to reject the idea that a Full Theory of Mind could have been the product of a biological evolution and lead to the postulation of the existence of an innate and inherited mindreading device (Seemann, 2012, pp.307-9).

Since the only ability that has been proved to be directly related to Theory of Mind competence is language, insofar that children without language or with impoverished language do not achieve a ToM at all, the acquisition of complex language features, including semantic and syntax and their use as a representational system in conveying and reflecting on knowledge, is a process worth to be studied as far as the general comprehension of the ToM mechanisms and its development from childhood to adulthood are concerned. In order to become a member of the community and to enter into a sort of mind exchange system, children must learn before all the language of the mind since to provide an effective interpersonal exchange is one of the prerequisites that language fulfills. Moreover, children must become capable of turning someone else's statement about belief into their own mental representation, that is, they have to use language as a representational means, an accomplishment that requires several years of experience with language in use. Accordingly, the beginning of Theory of Mind is widely thought to reside in the achievement of inter-subjectivity in infancy, signaling the onset of a concept of intentionality of self and other. It is worth to mention that inter-subjectivity allows the sharing of activities as well as the ability of assuming different roles within that activity, as in feeding; it also furnish children with a sense of sharing perceptions and actions that enable them to differentiate the self from the shared thing (Tomasello, 1999).

As far as the passage from a private, intimate mind to a social, cooperative one, the listening of stories plays a vital role since it requires the mastery of the representational

language of narrative, so that story understanding involves the use of language both as an internal, personal representation and as derived from the external representation of the world. Generally speaking, the story is about something that is removed from the present and thus from the child's own experience and it is possibly about a totally unreal event which is just a product of the author's imagination. Thus the child has to represent in his mind a reality that is sometimes in sharp contrast with the known and actual reality, ability that requires a lot of brain working. Mental representations of this kind are not possible without the implements furnished by a representational language and therefore children's relative degree of mastering the language skills, as measured by standard tests, should be related to their understanding of story themes and characters. As expected, a study conducted by Fontaine (2002) on 4 and 5-year-olds' story understanding concerning the representation of different story genres (both fantasy and reality), revealed that receptive language was a significant condition for both story understanding and the understanding of the relation of fantasy-reality in stories. This result has been taken as evidence of the existence of a particular bond between receptive language and the achievement of the higher levels of the cognitive functions relating to it.¹²

Another pioneering study of possible links between conversational input to the child and the child's Theory of Mind was conducted by Dunn, Brown, Slomkowski, Tesla, and Youngblade (1991) where they noted a correlation between aspects of family conversation and the child's performance on a measure of false-belief understanding. The authors appropriately warned that the correlation could be interpreted in various ways: one possibility is that conversational input to the child promotes the child's understanding of mental states (notably false belief tests) while a second possibility is that a child with an advanced understanding of mental states might be particularly likely to sustain a conversation about thoughts and feelings. In this second interpretation, the type of conversational input a child receives might be a consequence rather than a cause of the child's mental-state understanding.

¹² For a fuller account on the study, see Astington, J.W. *Why Language Matters for Theory of Mind*, p.39-46, New York: Oxford University Press, 2005

3.2 Against Theory of Mind: Inter-subjective Relation Theory, Intentional Relation Theory and Narrative Practice Hypothesis

Theory- Theory (TT) and Simulation Theory (ST), which have been previously defined as the standard and dominant approaches to social cognition, share the important supposition that when we attempt to understand the actions of others, we do so by making sense of them in terms of their mental processes to which we have no direct access. That is, while we are making an attempt at interpreting their beliefs, desires, and intentions, we are engaging our primary and pervasive way of understanding others. Both TT and ST conceptualize social cognition as a process of explaining or predicting what another person has done or will do with the aim of anticipating his/her behavior. Nevertheless, it is worth to mention that claims that such theory or simulation processes are explicit or conscious are dubious from a phenomenological point of view. If in fact such processes were primary, pervasive and explicit, they should show up in our everyday experience but it was proved they rarely do. The phenomenological critique also rejects the idea, clearly found in TT, that our everyday dealings with others involve an observational, third-person stance toward them, which lead to the observation and further explanation of others' behavior. Conversely, from the phenomenologist's point of view, our everyday encounters with others tend to be of a second-person type and thus strongly interactive.

Claims that the processes described by TT or ST are implicit and thus not explicitly conscious find a different set of objections. In the case of TT, for instance, there is no evidence that mind reading processes are implicit or clearly visible in every day experiences. Moreover, although TT appeals to false-belief experiments, these experiments are visibly set up for testing explicit rather than implicit Theory of Mind processes (Gallagher, 2001) where subjects are asked to clearly consider the meanings of an observed third-party's behavior. On the other hand, implicit approaches to ST appeal to the neuroscience of mirror neurons and shared representations, but again there is no justification for calling these sub- personal processes "simulation", since according to ST, simulation involves the instrumental use of a first-person model to third-person "as if" or "pretend" mental states. In addition to these phenomenological and logical objections to

TT and ST, there is good evidence from developmental psychology that our ability to understand others emerges much earlier than TT or ST accounts predict.

Another objection can also be raised against the idea that a general theory (folk psychology) is sufficient for explaining the particularities of a great diversity of behaviors found in everyday life. Similarly, it has been underlined that running a first-person simulation routine, a process which is likely to be based on one's own mental states, seems inadequate to give a satisfactory explanation to the diversity of behaviors found in the world, objections that cast doubts both on TT and ST approaches (Zlatev et al., 2008, p.18-20).

3.2.1 Alternative accounts of ToM

- Inter-subjective relation theory

In situations of social interaction, we have a direct perceptual understanding of another person's intentions because these intentions are often explicitly expressed in their embodied actions and their expressive behaviors. An understanding of this type does not require necessarily the postulation or inference of a belief or a desire hidden away in the other person's mind since it is already manifest in his/her overt behavior. In contrast, as far as the interpretation of the neuronal resonance processes that would involve cognitive processes as implicit simulation is concerned, Gallagher (2005) has argued that these neural resonance process in fact instantiates a form of enactive social perception. In his account, it is worth to say that a primary, perceptual sense of others is already implicit in the behavior of the newborn: in neonate imitation, which depends not only on a contrast between self and non-self and a proprioceptive sense of one's own body, but also on a responsiveness to the fact that the other is similar to oneself (Bermúdez, 1996; Gallagher, 1996; Gallagher and Meltzoff, 1996), infants are able to distinguish between inanimate objects and people. This statement implies that from early infancy onward humans, and perhaps some others animals, have the capabilities for a primary inter-subjective

interaction with the others. Thus, before we are in a position to theorize, simulate, explain or predict mental states of others, we are already able to interact and to understand others in terms of their expressions, gestures, intentions and emotions, having in this way a first idea of how they act toward themselves and others. Thus the well-established Theory of Mind approaches, which involve theory (seen as an application of folk psychology) or simulation, by focusing on the acquisition of the concept of mental states (like belief) at the age around 3 or 4 years, seem to miss some basic and important points indispensable for the understanding of social cognition.

As far as the inter-subjectivity and Theory of Mind is concerned, the hitherto dominant approach in psychology, cognitive science and philosophy has been to analyze what has come to be known as social cognition in terms of a mentalizing that allegedly solves the philosophical and developmental problem of other minds. Despite the important empirical findings and hypotheses generated by the Theory of Mind (ToM) approach, it has been observed that in various occasions the research have significantly obscured rather than clarified what needs to be explained about these theories. In order to help clarifying the basic assumptions of the ToM approach, they can be summarized as follows:

1. There is a primary separation between the self and (the minds of) others;
2. The individual must bridge this separation either by some form of “Theory” or “Simulation” of the other’s mind, a process that is in some degree fallible;
3. The main bodily structures that are directly relevant for these process are innate or acquired modules engaged through an inferential or simulation- attitude;
4. Cognition develops essentially from inside to the outside world, with innate or acquired cognitive skills being eventually transferred or projected onto others for the purpose of explaining and predicting their behavior.

In contrast, phenomenologists such as Husserl, Vygotsky and Wittgenstein sustain the idea of inter-subjectivity as follow:

1. Human beings are primordially connected in their subjectivity, rather than functioning as monads who need to infer that others are also endowed with experiences and mentalities that are similar to their own;
2. The sharing of experiences is not only primarily on a cognitive level, but more basically, on the level of affect, so that perceptual processes and conative action-oriented engagements are element of main importance;
3. Such sharing and understanding is based on embodied interaction (e.g., empathic perception, imitation, gesture and practical collaboration);
4. Crucial cognitive capacities are initially social and interactional; only later are they understood as private or representational terms.

Other scholars such as Merleau-Ponty (1962), Scheler (1954) and Schutz (1966) continued this tradition by developing complementary accounts of inter-subjectivity (cf. Zahavi 2001) who share the idea that the basic forms of understanding of others are not inferential, but rather direct, stressing the importance of perception in this process. According to this view, both Theory of Mind approaches (TT/ST) are not satisfactory in providing a fuller and complete theory of how our interaction with one another works. Additionally, developmental evidence shows that young infants are capable of grasping the purposeful intentions of others through the perception of bodily movements, gestures and facial expressions. Trevarthen's notion of primary inter-subjectivity can provide a theoretical framework for understanding these capabilities and his notion of secondary inter-subjectivity shows the importance of pragmatic contexts for infants starting around one year of age. Moreover, the recent new born neuroscience of resonance systems (i.e., mirror neurons, shared representations) seems to provide further evidences to this latter hypothesis (Zlatev et al., 2008, pp.2-5).

These ideas taken together have been developed in the context of an embodied "Interaction Theory" of social cognition, even though for more sophisticated inter-subjective interactions in older children and adults some forms of ToM are still required. Furthermore, this idea is supported by appealing to the narrative competency and in particular to the Narrative Practice Hypothesis (or NPH) according to which it is through

the repeated encounters with narratives of a distinctive kind the normal route that allows children to acquire an understanding of the forms and norms with which they make sense of actions in terms of reasons. Generally speaking, Interaction Theory provides an alternative approach for understanding the basic narrative competencies that we all share, so that the critique to both TT and ST is meant to define a more positive account of the way in which every day inter-subjective abilities work.

- The Narrative Practice Hypothesis

The pervasive presence of narrative in our daily lives and the development of specific kinds of narrative competency can provide a literary alternative to Theory- theory and Simulation approaches. By assuming that the distinctive kinds of narrative which we encounter during the first years of our lives are what first allow us to develop our folk psychological competences, Daniel Hutto elaborated a new approach called The Narrative Practice Hypothesis (NPH) according to which children normally achieve a folk psychological understanding by engaging in story-telling practices. In his account, the stories about those who act for reasons (i.e., folk psychological narratives) are those which assume a great importance in children's development since stories of this special kind provide the crucial training set needed for understanding reasons (Gallagher et al., 2013, p. 215). Accordingly, children acquire their skilled competence in understanding reasons by being exposed to and by engaging with narratives when appropriately and actively supported by their care givers.

Thus, by being involved in the narrative mechanism, it is possible for them to develop an implicit understanding of how to make sense of persons as targets who act for specific reasons, an ability which they can achieve because all stories, even the simple folk psychological narratives, can successfully show the experiences of the fictional mind. As a consequence, it can be stated that there are important links between narrative abilities and our capacity to understand others. The exposure to stories constitutes a critical determiner for the development of folk-psychological abilities since controlled

studies have shown that narrative training is responsible of improving performances on false belief tasks and this is also true in the case of children affected by autism (ASD). Thus, it has been concluded that narrative is an active tool for improving children's Theory of Mind development. Similarly, it has been observed that frequent conversations about the mind can accelerate growth of a ToM (Garfield, Peterson and Perry, 2001).

Folk psychological narratives are distinguished from the others account of the mind by being about agents who act for reasons. Importantly, such specific narratives can play their peculiar role in development by being the objects of joint attention in early learning, an assumption which stands at the core of the NPH, according to which around the age of two, children already possess an early intentional understanding of persons as having internal goals and wants, namely of states of mind that are clearly different from person to person. By learning to make sense of persons in a dramatic and narrative way, young children while listening to stories or play-act (but the same is valid also for adults who are exposed to parables, plays, myths, novels) become familiarized with sets of characters and with a range of ordinary or extra-ordinary situations that help them to shape their expectations about both real and fictional situations. Therefore, an education in narratives of many sorts can stimulate the knowledge of what actions are acceptable and in what circumstances, what sort of events are important and what behaviors are socially acceptable and what are not, in order to provide an idea of what is allowed and what is forbidden in a social context.

By and large, stories, real or fictional, have the social function of depicting what others can expect from us as well as what we can expect from others in certain situations. Thus through stories it is possible to know the norms associated with social roles that pervade the everyday environment; this activity of engaging with narratives is not a passive matter but rather it presupposes a wide range of emotive and interactive abilities. To appreciate such stories children must initially be capable, at least to some degree, of an imaginative identification and of responding emotively to the situation proposed, just as they do in basic social engagements ((Zlatev et al., 2008, pp. 30-2).

Janet Astington (1990) has argued that acquiring narrative competency requires necessarily having some kind of Theory of Mind since what begins as a perceptual and

emotional resonance process in early infancy, by allowing us to infer the feelings and intentions of others from their movements, gestures, and facial expressions, lead to the development of a more mature understanding of how and why people act as they do in adulthood, founding in this way the ability of framing their actions most and foremost in a narrative way. As a consequence, our everyday abilities for inter-subjective engagements and interactions are, in the later stages of childhood, transformed by encounters with narratives. In other words, it is the exposure to the complex objects of joint attention proposed by narrative stories that is responsible for the development of the sophisticated folk psychological abilities and understandings, competencies which will remain importantly in play for the person's entire life.

- Intentional Relation Theory

According to Intentional Relations Theory (Barresi and Moore 1996), the understanding of the self-other equivalence requires parallel knowledge of mind both from a first and a third person point of view and an awareness that any mental concept must directly match and link these two ways of knowing reality. In the Theory of Mind (ToM) approach to social understanding, emphasis is placed on sophisticated abilities to understand mental states, in particular on the ability of attributing representational mental states such as beliefs to self and others. As a consequence, it is the capacity of attributing false beliefs that is taken as an hallmark of the specifically human form of mentalistic social understanding that characterizes Theory of Mind. Precisely, as already mentioned, according to the Theory- theory approach, humans have an innately (or acquired early in development) ToM mechanism that can be applied to both self and other, a system which is based purely on inferences from external behaviors, so that the self-other equivalence in this account is based on the fact that one can interpret one's own behavior in the same way that the behavior of others can be understood. In contrast, Simulation Theorists suggest that what someone else feels can be understood only by observing a particular behavior in a particular context.

A third kind of theory, which gives prominence to inter-subjectivity, invokes the notion of matching or sharing attitudes or psychological states between self and other and this view is represented by a range of different accounts. Though, while the capacity for mind-sharing is evident in the forms of inter-subjectivity, what is not completely clear is how the infant moves from sharing mental states with others to the understanding of mental phenomena as distinct and possibly different in self and others. The Intentional Relation Theory (IRT) differs from the other accounts in specifically addressing the beginning of the recognition self-other equivalence and difference as involving a developmental shift from mind-sharing to mind-understanding. The key notion in IRT seems to be that the first-person information that we have about our own IRs (e.g., the “feeling” of love for someone) is distinctly different from the third-person information that we have about the IRs of others (e.g., another’s “behavior” toward the object of love). As a consequence, in order to develop uniform concepts or representations of IRs that can be applied equally to self and other, these two types of information need to be matched in a single concept or form of knowledge that contains both (Zlatev et al., 2008, pp.42-4).

In order to better understand IRs, the organism must be able to combine first person information about IRs with third person information about IRs into integrated representations involving an agent, an intentional relation and an object that can be equally applied both to the IRs of self and the IRs of others or to the joint activity of self and other. This complex process seems to begin in the fourth year of age, when pre-school children achieve an high level of social understanding by being able to imagine both first and third-person properties of a mental state. The results of the combination of first-third person information in children developing knowledge of mental representation as such allow them to understand conceptually the mind of others. Though, according to IRT, the levels of intentional understanding at which there is a comprehension of individual minds derive from previous shared intentional activities where first- and third-person information are originally associated.

The inability to readily transform third-person perceptions into first-person experiences, as well as to make the reverse mapping, makes it difficult for autistic

individuals to make sense of mind because of the absence of a direct connection between the two necessary, inseparably tied aspects of all mental phenomena, namely an externally available bodily expressive component and an internally available feeling. As a result of this deficiency, in the process of sharing mind with others, they lose interest in other people and have difficulty in learning from them. Eventually, if they do attempt to reflect on minds and understand them both in self and others, they form two radically different accounts: on the one hand, they develop rather complex TT-like accounts of mind from a third-person view of their own and other people's behavior; and on the other hand, they overgeneralize in apparent simulation their project their own egocentric first-person perspective to the perception of others (Frith and De Vignemont, 2005). Moreover, because of a lack of mind sharing during infancy and beyond, they are faced with intractable problems in understanding mind beyond those that appear as purely third-person TT types or purely first-person ST accounts.

Regarding this issue, IRT provides three important elements: first, it postulates a distinction between first- and third-person information concerning intentional relations, as well as a requiring that both forms of information have to be combined in order to generate representations of intentional action that are shared between self and others. Second, it claims that a clear distinction has to be made between different levels of social understanding in which first and third-person information are integrated. Third, it postulates that under certain conditions, first- and third-person information about intentional relations may be processed separately so that the activities of self and other are represented independently, a condition that can be clearly seen in autism. Although ST might account for some instances of Theories of Mind generated purely from behavior, it is only in autistic individuals where it is possible to observe a clear anomalous attitude of its workings since the failure in mapping first- and third-person information is observable in this pathology from very early on in life, an inability which prevents them from the sharing of mental activity (Zlatev et al., 2008, p.61).

3.3 Further developments of the debates on Theories of Mind

In recent years, much of the discussion about the nature of social cognition has taken place within the framework of the so-called Theory of Mind debate. The neat division between TT and ST is an oversimplification not only because of the existence of several hybrid theories but also because neither of the main positions are theoretical monoliths: Theory-theorists are basically split on the issue of whether the theory in question is innate and modularized (e.g. Carruthers, Baron Cohen) or whether it is acquired in the same manner as ordinary scientific theories (e.g. Gopnik, Wellman). As for the Simulationists, some claim that the simulation in question involves the exercise of conscious imagination and deliberative inference (e.g. Goldman) while some others claim that the simulation, although explicit, is non-inferential in nature (e.g. Gordon); finally there are those who argue that the simulation rather than being explicit and conscious is implicit and sub-personal (e.g. Gallese).

Dan Zahavi asserts that one of the most important powers of the human mind is the ability of simultaneously conceive of and think about itself and other minds; because the mental states of others (and even our own minds) are completely hidden away from the senses, these states of mind can only be inferred, recalling again the idea of a veiled, internal set of mental states that has to be codified (Zahavi, 2014, p.173). Whereas the sustainers of the Theory-theory argues that our understanding of others chiefly engages detached intellectual processes, being moved by inference from one belief to the other, the Simulationists, on the other hand, argue that our understanding of others exploits instead our own motivational and emotional resources. As a consequence, in contrast to the Theory-Theorists, they deny that what lies at the root of our mentalizing abilities is a sort of mental postulations. Instead, in their view, we possess no such theory, or at least an incomplete one which is not enough to underpin all our competences with psychological notions.

According to the account on Simulation Theory as expressed by Goldman (2006), we do not need a theory in order to understand others but rather we can simply use our own minds as a model, so that the understanding of others' minds would be grounded in

the introspective access to our own mind. In this view, the capacity for self-ascription precedes the capacity for other-ascription. More specifically, Goldman argues that the understanding of the others is rooted in the ability of the “I” to project imaginatively itself into other’s situation. In doing so, imagination plays an important role since it seems to guide entirely the process of assuming the target’s mental setup.

Despite their differences, TT and ST both agree that it is impossible to experience other minded creatures empirically and this is the reason for which we need to rely on and employ either theoretical inferences or internal simulations. By agreeing about the hidden mental states of others, the two main ToM standpoints challenge the problem of facing a theory of social cognition to solve the question of how and why we start to ascribe such hidden mental entities or processes to certain recognizable targets. As it was already mentioned, phenomenologists would question this issue from the beginning of their study on the mind. In particular, they argue that the appeal to either theory or simulation is unnecessary since this attitude is motivated by a too impoverished conception of what is experientially available. Interestingly, it is occasionally assumed that a phenomenological account of inter-subjectivity is by and large opposed to the Theory-Theory of mind, whereas the relation between phenomenology and the simulation theory is much more conciliatory since they share the fundamental points mentioned above.

The current simulation proposal has been closely linked to an empirical investigation of the underlying psychological mechanisms of our folk-psychological abilities to interpret, explain and predict the behavior of other agents; the general principles of this approach have been laid down by a great number of contributors over time. In the 1980s philosophers advocated simulation or replication as an alternative to the dominant functionalist’ approach to folk psychology (Gordon, 1986; Heal, 1986; Goldman, 1989) while in the 1990s a developmental viewpoint on Simulation Theory was presented by Harris (1992) and later in the 1990s and 2000s neuroscientific findings further sustained Simulation Theory (Gallese and Goldman, 1998; Currie and Ravenscroft, 2002; Decety and Greze, 2006; Goldman and Sripada, 2005; Gallese, 2007). Moreover, many treatments of Theory of Mind incorporate two or more levels,

components or systems of mindreading so that to offer a bi-level approach (Goldman, 2006).

Generally speaking, it can be stated that investigating folk-psychological capacities is an interdisciplinary enterprise involving the study of animal behavior and intelligence, developmental child psychology, psychopathology and neurobiology, to mention some of the disciplines involved in the field. A major component of the ST-TT debate has addressed to the plausibility of the thesis that a complex skill like mindreading is driven by a theory that unfolds during a child's early years. Early defenders of the child-scientist version of TT claimed to find evidence that children revise their theory of belief between two and four years of age, reaching a mature stage only around four (Wellman, 1990; Perner, 1991; Gopnik and Meltzoff, 1997). The details of this position, though, were overturned when Onishi and Baillargeon (2005) found false-belief competence (in non-verbal tasks) at 15 months of age.

Psychologists refer to research into our folk-psychological mindreading abilities with the label of Theory of Mind since Premack and Woodruff's 1978 publication. Specifically, in the first definition, Theory of Mind was described not primarily in terms of specific underlying causal mechanisms, rather in terms of engaging in a conceptual and interpretive practice whose developmental aspects had been central to the ToM debate for decades, including the enormous cognitive advances that children seem to make between the ages of three and four in their folk-psychological understanding of other agents. In fact, within this period of time, it seems that infants come to understand that other persons can have false beliefs about the events in the world and thus they can misrepresent their subjective states of affairs (Stueber, 2006, pp.99-100).

The second set of centrally relevant phenomena in this debate are the observed deficiencies and abnormalities in social behavior and cognitive development resulting from certain psychopathologies, as those observed in autism, ever since such psychopathologies have been diagnosed as specific deficiencies in mindreading abilities.¹³ Autistic children, who show a variety of deficiencies in their propensity to

¹³ For a complete description of the research on autism, see also Baron Cohen, S. *Mind-blindness: An Essay on Autism and Theory of Mind*, USA: The MIT Press, 1995

socially relate to other people as well as to manifest gaps in their linguistic and communicative competences, do not pass the false belief task at the normal age, in contrast to children not affected by this pathology and even in contrast to children with Down syndrome; moreover, they seem unlikely to engage in spontaneous pretend play.

Baron-Cohen has proposed that our mindreading abilities are due to the interaction of four subsystems: the intentionality detector (ID), the eye-direction detector (EED), the shared attention mechanism (SAM) and what he calls the Theory of Mind module (ToMM). ID and EED are quasi-perceptual systems that respectively process information about bodily movements in terms of primitive, intentional agency and important information about the direction of another's gaze, both feeding the information into SAM, which determines the exact relationship between the observed agent and the self. In a second moment, this information is conveyed to the ToMM and interpreted in terms of a full range of mental concepts characterizing our ordinary folk psychological practices. In this way, Baron-Cohen diagnoses autism as due to a shortcoming in SAM and ToMM, since children with autism have difficulty in passing the false-belief task and presents deficiencies in joint-attention behavior (Stueber, 2006, pp.103-5).

Since the original Theory of Mind debate focused on trying to understand our mature folk psychological abilities (i.e. the capacity to acknowledge that other people can have different subjective perspectives on the world and that these perspectives might not represent the world as it really is), it is not surprising that researchers were particularly impressed by four-year-old children's ability to pass the false-belief task. Interestingly, the capacity for grasping the mind state of another person seems to be present much earlier in ontogeny, as the phenomenon of joint attention clearly demonstrates. Thus it would be better to distinguish between two stages in the process of human capacity for understanding social cognition, namely an early stage in which the child is able to understand persons as intentional agents and a later stage, where the child is able to understand persons as mental agents having proper beliefs and thoughts (Seemann, 2012, p.268).

Analogously, Tomasello (2010) seems to align himself with Simulation Theory since he perceives the ability to view others as intentional agents, an ability which

normally develops during an infant's first year, as being causally dependent on the infant's increased faculty to interact with the world in a structured and persistent manner as well as on experiencing himself or herself and his or her own movement as being goal directed. In interacting with objects belonging to the world, the infant firstly perceives his or her place in the world as an intentional agent and only subsequently understands others in the same manner. Moreover, Peter Hobson by stating that to identify with someone else means to relate to the actions and attitudes of someone else from his/her own perspective (Hobson, 2007), emphasizes that the phenomenon of identification proceeds by assimilating the world-orienting perspective of another within one's own psychological outlook, allying himself in this way with the Simulation point of view. Additionally, mirror neuron research supports Hobson's focus on the primary role of emotions in infant's engagement with others and it also supports Tomasello's claim that the early understanding of another person as an intentional agent contains an egocentric element, strictly depending on the infant's ability to interact with the world (Seemann, 2012, p. 274).

Nevertheless, the interpretation of the functions of mirror neurons is rather problematic. Alvin Goldman has conceived mirror neuron systems in humans as an evidence that supports the Simulation Theory while Shaun Gallagher views mirror neuron activity as a perceptual process that allow us to understand the meaning of observed events and facial expressions. By taking into account these different positions, it is clear that the existence of the mirror neuron system in the brain is not a sufficient evidence for giving a satisfactory and mature interpretation of human folk psychological abilities. Specifically, Alvin Goldman acknowledges that the mere existence of the neurobiological resonance on part of the observer does not automatically imply that the observer systematically reads the mind of the other person, so that mirror neuron activity cannot automatically be conceived as the main responsible for the mind reading ability.

Similarly, Shaun Gallagher proposes that the understanding of the mirror neuron process as part of a system directing perception has to be viewed as mechanism strictly related to our inter-subjective interactions within each other. In this view, mirror neuron activity enables us to directly perceive the veiled meaning in the action or gesture of

others, namely their joy and anger or their intentions as expressed in the face or in the posture of the others.¹⁴ Moreover, Rizzolatti and Sinigaglia added that “our perceptions of the motor acts and emotive reactions of others appear to be united by a mirror mechanism that permits our brain to immediately understand what we are seeing, feeling, or imagining others to be doing, as it triggers the same neural structures that are responsible for our own actions and emotions”. Their main point of discussion is that the mirror neuron system provides an “understanding of the meaning of motor events in that it immediately perceives the meaning and interprets them in terms of an intentional act performed singly and when they are part of motor chains” (Rizzolatti and Sinigaglia, 2008, pp. 97- 8: 113).

According to Karsten Stueber (2006), resonance mechanisms realized by the mirror neuron systems as mechanisms of basic empathy are a form of inner imitation that allow us to grasp other person’s intentions because others are anatomically analogous to ourselves. In her account, mirror neurons create a perceptual similarity and spaces of familiarity in the direct encounter with other persons but are not enough to predict other’s thought, desires and behaviors. Thus her explanation of basic empathy supports the fact that our grasping of other persons contains an irreducible egocentric moment of recognition that is different from the accounts of empathy described by the phenomenological traditions, such as those articulated by Stein and Husserl (1989), according to which the “I” directly experiences foreign subjectivity without using his/her own experiences as a model for understanding the subjectivity of the other.

Basic empathy and mirror neuron mechanisms certainly are not the unique necessary tools for the proper development of our mature mindreading abilities since other attitudes such as the capacity to follow the gaze of another person, detecting the direction of another’s eyes and the ability of recognizing faces also play an important role in the development of this fundamental tool. Again, it is not to be expected that one will find a module for joint attention in the brain but rather, the phenomenon of joint attention

¹⁴ For further explanations see also Gallagher, S. *Logical and Phenomenological Arguments Against Simulation Theory* in Hutto, D; Ratcliffe, M. *Folk Psychology Re-Assessed*, p.71, Springer Press, 2007

is best understood as emerging from a variety of more primitive capacities together with motivations for cooperation and sharing emotions. For Gallagher, more specifically, the normal understanding of a person's action within the social realm is accomplished by our ability to posit the person's action into a larger context in light of shared cultural background assumptions. Accordingly, the function of folk psychology is that of allowing us to maintain a meaningful relationship between individuals who reciprocally understand each other's thought, beliefs and desires in particular environmental conditions.

The potential capacities of basic and re-enactive empathy contribute greatly to our integration into the social realm so that an egocentric, irreducible element is always present in this type of communal relation. In John Campbell's point of view, joint attention seemingly involves social cognition of the sort that is called mind reading, for which the "I" not only have to know that the other person is attending, but also that he/she is capable of having such a mental state. These Theory of Mind versions of joint attention seem to include TT and some ST accounts: for TT, a concept of attention is needed and also an explicit or tacit knowledge of a theory that would allow the subject to understand that the target is attending to something, while ST requires to simulate the possession of the target's mental state; for both TT and ST, decoupling requires action because they have traditionally been described as a third-person observational accounts rather than second-person ones (Seemann, 2012, p.296).

Daniel Hutto proposes another alternative to ToM, which is called the Mimetic Ability Hypothesis (MAH). MAH claims that an increased capacity for re-creative imagination best explains the kind of sophisticated inter-subjective engagements of which hominids would have been capable of as well as sustaining that these abilities constitute an important basis for the development of a complex language system. His proposal places the idea of an evolutionary value of ToM devices since it has been argued that ToM abilities would have been needed in order to share a plan or a goal for developing and implementing the sophisticated hunting tactics or even for erecting various constructions (Baron Cohen, 1999). Yet this thought must be analyzed in light of the fact that even wild chimpanzees and other group of animals are quite capable of coordinating

their hunting efforts, despite a manifest lack of mature ToM abilities (Boesch and Boesch-Achermann, 2000; Brinck and Gärdenfors, 2003), considerations that lead to take an extreme caution when drawing inferences about the degree of mentalizing capacities that our ancestors might have possessed.

According to Robin Dunbar (2003) Theory of Mind is probably essential for language not so much because it is involved in the production of speech per se but because it provides a mechanism that both enables speakers to guarantee the spreading of their message and allows hearers to correctly understand the speaker's message. Interestingly, he has independently established that there is a direct correlation between neocortical volume in primates and the maximal size of their workable social groups (Dunbar 1992, 1993). By organizing ideas from this data and estimating the brain size of *Homo erectus*, he suggests that these hominids may have been operating in social groups with as many as 150 members, thus a device for normalizing social relations should have been indispensable. In this view, language has a role of mere facilitator, lacking any intrinsic representative power of its own; by accepting this picture of the function of language, meta-representational devices would have been needed for allowing public exchanges. Building directly on Dunbar's work, Dautenhahn (2002) has proposed that the capacity to produce and comprehend narratives, under the form of conversational gossip, may have evolved in order to resolve the dilemma of living in larger social groups so that the evolutionary origin of communicating through stories is thought to have evolved through the increasing social dynamics among our human ancestors and, in particular, the necessity to communicate about third-party goals and plans.

There is a clear and interesting connection between the kind of imaginative re-enactments that a mimetic culture would have promoted and the basis for the kind of narrative competency young children first exhibit: social dramas are the core of many narratives and these narratives are the vehicle that provide the subject matter for narrating the actions of our ancestors even though they, lacking the appropriate linguistic abilities, were in no position to tell their stories in an organized way, so that there could have not been narratives without an appropriate linguistic capacity for shaping and narrating their experiences.

By assuming that *homo erectus* lacked a coherent system of language, the resources for conducting discursive conversation and telling stories would have been missing (Dunn, 1991). In place of such discourse, a developing capacity for mimesis may have arose, such as exemplified by the meta-representational mechanisms, which played an instrumental part in the development of the very first public language. Notably, a capacity of this kind is well shown by human infants around the first year of their lives, so that they can be taken as an example of the most sophisticated form of non-verbal inter-subjective engagement. This first form of early interaction involves participants not only attending to the same objects at the same time but also it requires them to mutually recognize their attending as a common, shared activity. As a consequence, inter-subjective interactions of this kind should be regarded as empirical proofs for the existence of a Theory of Mind neatly different from the more speculative theories of mirror-like reciprocal imitation (Zlatev et al., 2008, pp. 260-1).

Simulation is usually presented as a kind of imaginative attempt to adopt another's perspective on events but there are problems even for the defenders of this theory, such as Gordon (1995, 1996), who only invokes a first-personal "transformation" as opposed to a "projective" version of the theory. For, unless such an act of simulation is assumed to take place within a context in which the simulators are somehow already independently aware that the other in question has a different point of view, the empirical perspective would be an utterly inadequate means of approaching to the perspective of others (see also Gallagher, 2007). To sum up, whereas in joint attention what is required is both identification and recognition of difference (Hobson, 2007) simulation per se involves only the experience of acting in and attending to a shared world, without taking into account others' point of view in the process of response to public objects.

As mentioned in various sections, the mechanisms that underpin the mutual connectedness that enable identification and common focus are likely to have their ancient roots in mirror neuron systems. Interestingly, key elements of the human mirror neuron system for grasping are found in Broca's area, a part of the brain that relates to language production and comprehension. This founding has inspired researchers to speculate that there may be interesting, possibly quite tight, connections between this

kind of intentional attunement and the basic lexicon which underlies the various abilities for grasping (Rizolatti and Arbib, 1998; Billard and Arbib, 2002; Arbib, 2005). This hypothesis makes plausible that a resonance systems of this kind may have played a crucial role in enabling humans to enjoy a shared world as well as developing a common language for describing it, ideas that lead directly to the already mentioned hypothesis of a possible connection between joint attention and word learning. The activity of jointly attend requires more than mere identification since it is indispensable to be able to recognize the other as other and at the same time to be able to make inferences from inside to outside, based on assumptions of similarity. Described as such, joint attention might be thought to integrate mindreading abilities with the use of propositional attitudes, emphasizing the importance of linguistic abilities.¹⁵

In particular, according to MAH theory, seeing another's seeing does not necessarily requires the representation of the other's cognitive abilities but just imagining their perceptual ones; as a consequence, even the most sophisticated form of non-verbal inter-subjective engagement does not involve the manipulation or attribution of propositional attitudes so that there is no need to appeal to a sort of re-creative imagination to explain pre-linguistic joint attention. Non-verbal acts of joint attention are in fact only based on responses to intentional attitudes, so that infantile capacities to identify and respond to intentions-in-acting are quite distinct from the understanding of intentions that rely on having mastered the concepts of desire and belief. Moreover, the existence of linguistic abilities in individuals with autism can be taken as a proof of the non- reliability of the claim that it is necessary to have a ToM in order to learn words. Presumably, autistics are able to achieve this social capacities because, even though they cannot jointly attend with others, they are supported by expert language users who are able to do this.

As previously mentioned, the original Theory of Mind framework was a development of philosophical cognitivism and especially of the concept of folk psychology (see, e.g., Dennett, 1971 and 1987; Bennet, 1976; Stich, 1983; Fodor, 1987;

¹⁵ For further explanations, see Zlatev, J. et al. *The shared Mind: Perspective on Intersubjectivity*, pp. 275-9, USA: John Benjamins Publishing Company, 2008

Goldman, 2006). Yet the treatment of children and adults as if theoreticians is not actually an empirical discovery but rather a background assumption inherited from Chomsky's theory of language, namely the conviction that the study of intentional social activities seems no longer to deal with social psychology but rather with cognitive psychology. This essentially cognitive approach has been widely adopted by developmental psychologists aimed at explaining the development of social cognition (Bretherton, McNew and Beeghly-Smith, 1981; Perner and Wimmer, 1985) and has provided the basis for influential accounts of psychiatric disorders such as autism (e.g., Baron-Cohen, Leslie and Frith, 1985) and schizophrenia (e.g., Frith and Corcoran, 1996).

The original version of Theory of Mind has, over time, been joined by a number of alternative theories of social understanding that have also been included under that original heading of ToM, so that the kind of theory proposed by Premack has now to be referred to by the umbrella term Theory- Theory of Minds (TToM). One of the earliest alternatives to TToM postulated a biological module, a Theory of Mind Mechanism, or ToMM, according to which it is theorizing that in effect allows us to make sense of one another due to a biological inherited mechanism so that ToM, in this version, is intelligence by Nature. As Tooby and Cosmides (1995) claim, we are all mind-readers since infancy and we have developed this ability by building interpretations of the mental events of the others and by sharing their feelings. Moreover, humans appears to have evolved this cognitive tool because, as members of an intensively social, cooperative and competitive specie, their lives strongly depended on their abilities to infer what was going on in another's minds. As a consequence, the problem of other minds is not only a natural one but rather the direct consequence of the necessity of engaging with other people for establishing a collective society. The issue about the relation with other minds is not something universal, rather it is created by thinking about people as being made of minds and bodies, a dualism reproduced in psychology as the contrast cognition-behavior.

As already mentioned, it is has become a customary to blame Descartes for starting the problem mind- body in philosophy, through the postulation of dualism mind-body and the use of the systematic doubt. According to Gilbert Ryle's account about what he defined the Cartesian Myth:

Every human being [...] has both a body and a mind. [...] Human bodies are in space and are subject to the mechanical laws which govern all other bodies in space. Bodily processes and states can be inspected by external observers. So a man's bodily life is as much a public affair as are the lives of animals and reptiles and even as the careers of trees, crystals and planets. But minds are not in space, nor are their operations subject to mechanical laws. The workings of one mind are not witnessable by other observers; its career is private. Only I can take direct cognisance of the states and processes of my own mind. A person therefore lives through two collateral histories, one consisting of what happens in and to his body, the other consisting of what happens in and to his mind. The first is public, the second private (Ryle, 1949, p.13)

Even though this view explains minds as complicated mechanisms and claims to be materialist rather than dualist, it inherits the problem of other minds in distinguishing mind from behavior and in assuming that mental states are essentially private, whereas behaviors are not. In fact, both the categorical behavior-mind distinction and the privacy of the mind are formulated in the ToM debate in a similar way Ryle found them in Descartes. Accordingly, it can be stated that behaviorists, in their determination to be scientific, failed to notice that they adopted a definition of behavior that is basically Cartesian and were thus themselves responsible for the dualistic meaning of their own theories.

The cognitive revolution, which is aimed at bringing back mind from psychology to the cognitive sciences (Bruner, 1990), actually maintained the problem of other minds since cognitivism, for instance, reinstated the legitimacy of scientific inference to the unobservable, but retained the conception of behavior as logically separate from mind, regarding it as evidence of an essentially unobservable mind so that behaviorism and cognitivism, despite their declared materialism, both operated in what is defined as a Cartesian framework for thinking about people insofar they remain epistemically divided.

Chomsky is famous for having delivered modern psychology from behaviorism and his influence in psychology was deeply considerable. In marginalizing behavior, he reinforced the Cartesianism implicit in scientific psychology by focusing on the mind and even though he did not precisely write about it, he has been credited to have provided a role for the Theory of Mind (e.g., Smith, 1999: 25–6). As Chomsky puts it:

One might attempt to characterize the knowledge of language perhaps knowledge more generally, as a capacity or ability to do something, as a system of dispositions of some kind, in which case it is perhaps not unreasonable to think of behavior as providing a criterion for the possession of knowledge. In contrast, if such knowledge is characterized in terms of mental state and structure, then behavior simply provides evidence for possession of knowledge¹⁶.

Thus he significantly contributed to the acceptance within psychology of the notion of language as a system of grammatical rules and to the extension of this way of thinking to other psychological faculties, such as to the various disciplines involved in the Theory of Mind study.

Research on Theory of Mind has proliferated since the 1980s and there are now well-known different versions of these theories. Indeed, the name for this field of research derives from a specific theory which has now to be referred to as the already mentioned Theory-Theory of Mind. Despite the apparent diversity of the positions in the field of the Theory of Mind debate, these different assumptions share, at least at varying degrees, the following specific ideas:

1. Somehow we all have to engage in an inferential jump beyond what we can actually observe about other people, namely their behavior, and their intentions, emotions and thoughts. Thus it is only through an intellectual or quasi-intellectual process that we can truly know other people as beings with minds;
2. Any development towards a proper understanding of other people, and engagement with them, consists of a single transition;
3. Very young children and most other animals are incapable of properly relating to others as psychological beings or really communicating with them; despite their lack of the relevant Theory of Mind capacity, they can infer what the people around them are thinking if guided by their caretakers;
4. Autism and schizophrenia can be understood as evidences of the complete or partial absence of the ToM mechanism that allows the rest of people to make sense of others' thoughts and beliefs. Due to this assumption, various experiments

¹⁶ See Chomsky's account on language as reported in Leudar, I; Costall, A. *Against Theory of Mind*, pp.25-6, New York: Palgrave Macmillan, 2009

have been conducted within the field of impaired ToM abilities (e.g., on false belief);

5. There is supposed to be a very big problem inevitably facing people when they are making sense of one another, since there is a profound logical gap between understanding behavior and truly understanding minds. This problem, in actual fact, is the traditional problem of other minds, sustained by a radical skepticism regarding what we can know about other minds and, in its most extreme form, regarding the very existence of other minds (Kirschner, 2003).

Since TT postulates that the mentalist vocabulary has an organization and is coherently structured in the way that scientific terminology is, that is, combined in a propositional scheme, the supposed ToM users are not merely theoreticians hypothesizing unobservable phenomena as TT, unlike ST, takes for granted that adult human beings understand other humans on the basis of a theory regulating the ascription of thoughts, beliefs and intentions: an individual has a ToM if he imputes mental states to himself and others. A system of inferences of this kind is properly viewed as a theory because such states are not directly observable and therefore it can be used to make pure speculations about the behavior of others.¹⁷

For both TT and ST, it can be found an adherence to the Cartesian conception of the mind as belonging to an hidden, unobservable domain as well as it can also be noted a certain tendency to construe the demarcation line between which can or cannot be observed about human conduct along empiricist-behaviorist lines. Thus the use of expressions such as intention, thought and belief are treated by the ToM sustainers within the hypothetical and speculative field as they are focused on the observable and the unobservable, an idea that was already postulated by the behaviorists. Moreover, TT argues that we must be similar to professional psychologists who need to appeal to mind as a theoretical construct of something both unobservable and hypothetical if we desire to understand how it is possible to predict and explain human conduct in everyday affairs as it usually happens in the scientific practices.

¹⁷ For the complete theory see Premack and Woodruff, Does the Chimpanzee Have a Theory of Mind? in *Behavioral and Brain Sciences*, 4: 515-6, 1978

The supporters of the TT claim that what Dennett and others have termed a folk psychology is nothing but the phylogenetic and ontogenetic precursor of scientific psychology and thus it consists of the set of interdependent mental concepts/predicates in our language so that it is possible to call our folk psychology a theory because it facilitates predictions. One of the main standpoint of TT is the idea of partitioning a section of our language into a mentalistic domain, an assumption that underpins a further development, namely the postulation of cognitive modules; modularity theory, in cognitivism, asserts that a human capacity for speech is encapsulated and distinguishable from the capacity of perception and from other forms of conduct so that TT seems to be based upon an even more radical version of the modularity thesis (i.e. our capacity of using a portion of the language as distinguished and studied independently from the rest of the linguistic abilities).

One of the most important ideas behind TT is that humans, as very small children, have a generic problem: they do not understand other people. At some point in childhood this problem is solved by finding a solution that, according to Theory-Theory, is a matter of establishing a system able to explain how other people's mental states cause their behavior. Simulation theory, on the other hand, takes into account interpersonal comprehension as involved in the process of creating simulations of other people's mental states, a processes which presuppose to introspectively understand how those mental states are configured in the reasoning processes that would eventually produce behavioral attitudes toward others (Leudar and Costall, 2009, pp. 118-20).

As far as the linguistic aspect of ToM is concerned, the capacity for understanding other human beings is supposed to be strictly connected with the acquisition of language. As for the Simulation Theory, the linguistic aspect was elaborated to show that language learning is not a distinct and isolated activity, but rather is something that is done through social practices, so that language constituents are an integral part in the learning process. Again, the capacity for understanding others seems to derive from a sort of distributed learning and it is expanded, elaborated and diversified by the engagement in and mastery of the numerous and diverse practices of life, rather than being its generic and inherited precondition.

Discursive Psychology is a useful discipline since it provides an attempt to offer an alternative to mentalism within the psychological studies. Generally speaking, the doctrinal content is largely derived from other theories, supposedly adapting those theories to psychology's specific problems as a way for bringing psychology into closer contact with two general movements that have divided the humanities and social sciences in the past half-century, namely the linguistic turn and the turn to the social.

As the name implies, Discursive Psychology (DP) proposes that psychology can be, at least partly, converted into the analysis of discourse and that the nature of the mind has to be understood through the examination of a discourse that involves mental matters. By considering the psychological aspect in the context of discursive transactions and in the relations between people, this view avoids the excessive individualism of which mentalist psychology is often accused of and directly connects it with the turn to the social. DP's own exercise, therefore, is meant to be at the same time a compromise and synthesis between two supposed different views since it seeks to resolve the issues regarding the nature of language both by examining the language in practice and by casting doubts on the supposition that language use per se is entirely an independent mentalist concept.

It is worth to mention that nowadays a number of dialogical approaches in the field of language studies and in human and social sciences have been developed and each approach assumes different denominations, such as "the dialogical self", "the dialogical turn", "dialogism(s)" and so on. These different positions have a long list of predecessors and their beginning can be recognized in Vico's *New Science* (1744) in the seventeenth century, connecting Hegel, to the neo-Kantians and to the Bakhtin's Circle, creating a bond that connects them to the present days. All these scholars have in common the idea that the mind, knowledge and language are products of social interactions rather than of the individual's cognition. These perspectives, however, do not paint a definite and clear picture of the human mind but instead inspire a series of alternative ways of thinking about it (Leudar and Costall, 2009, 217-9).

Folk Psychology explanations, conversely, proceed by focusing themselves upon a common stock of familiar mental terms deployed in a systematic way. In line with this, mentalistic vocabulary is generally regarded as being a kind of well-established and

clearly demarcated rule-governed domain at the heart of which it is possible to find intentional/ attitudinal psychology and belief/ desire psychology. Simulation theorists take for granted that humans regularly engage their FP, but they also claim that FP do not involve a set of principles or a structured theory but rather that the minds are already settled with the relevant mental states which can be systematically manipulated in precisely the same ways the Theory-Theorist's principles aim to describe their functioning, so that it is possible to construe the relevant mental ascriptions of other minds by using our own minds as direct models (Hutto, 2007).

Although this activity is a structured one, involving the use of mental concepts and beliefs, no theoretical principle would actively be consulted or used in it. As mentioned in other sections, the kind of TT/ST union just mentioned is possible only because TT and ST theorists agree on several big issues concerning the mind and its working: in fact, proponents of both theories simply assume that the primary function of FP is to predict and explain actions in third person contexts through the attribution of causally efficient inner mechanisms of decoding mental states; additionally, both accounts recognize FP as having two undeniable features which are coherence and mentalism (Wellman, Cross and Watson, 2001), concepts that both TT and ST identify as fundamentally relevant. As a consequence, the most important common characteristic shared by the various position in the Theory of Mind debate is that they posit the existence of complex, inherited mechanisms that manipulate structured mental representations in one way or another.

Chapter 4: The Communicative Strategies of Joint Attention and the Phenomenon of ‘Mind-blindness’

4.1 Joint attention

Joint attention can be seen as a concept that occupies a central position in any account of the mind that acknowledges the importance of its social dimension. It can be said that this idea is located at the intersection of a complex set of capacities that serve our cognitive, emotional and action-oriented relations toward others since it involves both social cognition and our ability to understand others, what they intend and what their actions mean. As a consequence, joint engagement is at least helpful, and in some ways necessary, for infants at fourteen months to register other creatures as different from they own as well as becoming familiar with the environment around them.

Developmentally speaking, this capacity could be considered as a bridge between primary inter-subjectivity, defined as the set of sensorimotor abilities necessary for understanding the meaning of another person’s movements, gestures, facial expressions, eye directions and intentional actions and secondary inter-subjectivity, which is the employment of the primary abilities to join complex, intentional situations so that joint attention, by involving our abilities to understand the world through our interactions with the others, can be defined as an interactive activity by nature. Its functional role is useful to understand other minds and thus a shared attention module can be perceived as a precursor of a Theory of Mind module, as suggested by Simon Baron Cohen (1995). In the recent decades, numerous researchers have theorized about humans having an unique inclination for joint attention. For Baron Cohen, this phenomenon involves a form of meta-representation and the hypothesis of a human cognitive specialization for sharing attention seems to be extremely popular in contemporary psychology, philosophy and cognitive sciences. In this account, experience is gained within the feelings of our individual active bodies, in awareness of the world in which we move and in engagement with one another’s different and potentially collaborative intentions, interests and

emotional states, regarding both ourselves and the objects around us. In this way, the world becomes necessarily a social experience.

According to joint attention theory, before language is learned to represent anything that is customarily distinguished as having a symbolic identity, a common sense reality or “proto-habitus” comes to life, which derives from the convergence of shared interests and highly vocal affective appraisal of operational or practical intentional society, strictly dependent on the uniquely human abilities for interpersonal life and sympathetic emotions. For the infant, joint visual attention with the adult initially takes the form of direct eye contact or mutual gaze before developing into a more complex formation which alternates between the activity of attending to each other’s eyes and jointly intending an object present in their visual field.

Juan-Carlos Gómez’s account, as expressed in Eilan’s book (2005), provides an interesting alternative view to the Theory of Mind approach to social cognition. In his book, he theorizes eye contact as a perceptual rather than an intellectual process, so that visual attention is not an invisible mental process to be inferred but rather is directly perceived in the other person’s overt behavior via external signs of attention such as the direction of the gaze and the facial expressions. In the sighted population, eye contact is useful for reading or seeing the minds of others in a process of direct mutual interaction; therefore, it can be considered as a direct evidence of other’s mental states, obtained through a reciprocal gaze sharing. As a perceptual rather than purely cognitive process, mind-sight is firmly rooted in embodied human capacities such as gaze tracking and other’s facial expressions.

As a consequence, as far as the sighted population is concerned, the experience of mutual gaze serves the formative function of establishing social bonds by directing the infant’s and adult’s intentional states onto each other, making them aware of their reciprocal existence. Thus the importance of joint attention studies in relation to Theory of Mind seems to concern primarily the bodily sensibility and the hypothesis that human bodies engage in direct contact with others via mutual gaze, actively shaping the kind of relations that occur between minds, so that not only it is possible to see the other’s mind

in the other's body, but also the body and its sensibility actively inform the specific ways in which minds intercommunicate.

4.2 Mind-blindness

The concept of mind-sight described above could be used to explore the social difficulties within the autistic population from a different and possibly revealing perspective. By investigating the broadly construed perceptual abilities of the autistics, which include not only perceiving the physical environment via the visual field but also registering and responding to the manifestations of other persons' feelings, it could be argued that the challenge they face lies primarily at a cognitively lower level than that commonly assumed in the Theory of Mind model. In the particular case of autism, what is critical is the inability to perceive other people as having feelings as well as to respond to those feelings with feelings of one's own, identifying in this way with attitudes of others directed at events and objects in the world. Researches on social cognition and on the problem of understanding other minds in autism have, over the last 30 years, been largely dominated by the idea of a lacking of Theory of Mind, concerning the patients' inability to infer the internal mental states of other individuals.

According to Peter Hobson (1993), the difficulties concerning autistic people have to do with seeing, but in a sense that extends beyond vision. In order to clarify this concept, it could be argued that human seeing typically and normally extends beyond vision so that to include the affective component of the world; in this way we perceive people as well as objects and events in an affective way. As far as the researches on autism are concerned, the main aim is that of investigating this pathology at the foundational level, namely that of perception and affectivity, rather than at a cognitively higher level of theory, as it has become customary in the current debates on social cognition. The key question to ask seems to be why seeing has become dissociated from feeling or affectively impoverished in the autistic targets in a way that restrains the typical ability of mind-sight. As previously noted, the nature of interpersonal engagement

such as interpersonally coordinated affective states seems to have a pivotal developmental role in establishing the phenomenon of jointness. To better understand this concept, recent studies in developmental psychopathology and specifically clinical reports and scientific investigations of early childhood manifestations of autism, are relevant since clinical and experimental studies of children and adolescents affected by the autistic syndrome disorder can be helpful for revealing aspects of jointness in social engagement and orientation toward the world.

The syndrome of autism is an expression of limitations in the inter-subjective engagement in relation to a shared world, which manifest itself as a lack of joint engagement with other people's activities as well as within activities of their own. Among the common features of this syndrome, it can be mentioned the children's difficulty in using language to transmit meaning to others, their tendency to produce sentences that are disconnected, their often inflexible (context-insensitive) use of words, their abnormal use of the personal pronouns "I" and "you", their overriding desire to maintain sameness in their routines and surroundings and, more generally, the lack of variety in their spontaneous activities. Therefore, whatever feelings accompanied the children's achievements, these appeared to lack an orientation toward other people's appreciation and a seemingly lack of the ability and propensity to share experiences with them.

In the domain of joint attention, children with autism show a reduced tendency to use eye contact and deictic gestures (e.g., pointing or showing) to coordinate attention and share experiences with social partners. Since sharing (or jointness) requires connectedness and differentiation between two people, it is a feature that presupposes that one person apprehends other person's centered and bodily stance of the communicative in such a way that each person involved in the communicative web can make that stance his or her own for informative purposes. Interestingly, this particular feature of social engagement seems to be the core of what has been recognized as the hallmark for the activity of identifying with the behavior or attitude of someone else, an attitude which is dramatically lacking in the autistic targets.

4.3 Empirical studies on 'Mind-blindness'

According to recent studies (Hobson and Meyer, 2005), while making reference to their own bodies for trying to convey something of their communicative intentions, children without autism normally tend to foreshadow other person's intentions. Conversely, children with autism rarely communicated in this way, nor did they identify with the other person's intentions using bodily gestures toward them. Since sharing looks a vis-à-vis constitute an index of a deep interpersonal engagement, a series of studies have been conducted by Hobson (2005) in which children with and without autism were matched in order to study their propensity to imitate self-other orientation in four different actions involving different objects. The results showed that although all of the children copied the actions, those with autism were significantly less likely to imitate the self-other orientation of the actions.

Among a long series of tests made by Hobson and Lee (2007), it is worth to be mention a study involving a setting in which participants had the task of observing an investigator who demonstrates an action and then communicates this action to another tester who only subsequently entered the room, with the aim of completing the same action. There were six actions demonstrated and for each the demonstrator's instruction was "Get Pete to do this." Three actions involved goal-directed use of objects (e.g., using a mechanical arm to place a cloth frog into a waste bin), two were non-goal directed involving the body (e.g., raising hands above the head), and three included a form of expressive style (e.g., placing hands on hips in a proud, assertive stance). The ratings of the videotaped interactions were designed to capture the diverse expressions of the hypothesized underlying process of identification with the attitudes of the communicative partners. As predicted, the results were that participants with autism contrasted with matched participants not affected by the pathology in showing a minor degree of emotional engagement with the testers, forms of joint attention that implicates a sharing of experiences, communication of styles of action and role shifting from the learner to the teacher. When these measures were combined in a composite index for examining the data, the two groups were almost completely distinct (Leudar and Costall, 2009, p.122).

The results provided further proofs for sustaining three a priori predictions, which are the follows:

1. Children with autism contrasted with control participants seem to spend more time looking at the objects acted upon and less time looking at the tester;
2. Participants with autism showed fewer sharing looks toward the tester and, although they also showed fewer checking and orientating looks, they were specifically less likely to show any sharing looks.

Within each group, individual differences in sharing looks were associated with an activity of imitation of self-other orientation. Although only five of sixteen children with autism ever showed a sharing look, these same five were more likely to show role reversal in their imitation than manifesting a real interest for the other. This mode of social perception that is believed to be implicated in the process of identifying with someone else, is especially impaired among many individuals with autism. Not only are such individuals less likely to look to another person when copying their actions on objects but also they are less disposed to engage in sharing looks and less likely to imitate the self -other orientated aspects of the actions demonstrated.

In another study (Hobson, García- Pérez and Lee, 2010) semi structured tests were employed to determine whether children with autism produce and comprehend person-centered (deictic) expressions. Notably, it was observed that a great amount of children with autism sometimes referred to a location that was distant from themselves with the terms “this” or “here” (instead of “that” or “there”). These findings reveal something further about the communicative jointness and the structure of joint attention: children with autism use the deictics in an egocentric way, an a-typicality that appears to reflect children’s limited organization of interpersonal experience and a lacking of the reciprocal role taking abilities in relation to a world jointly shared with others.

Currently, there are two opposing theories claiming that people with autism have difficulty understanding goals and intentions of others. These are the “mentalizing theory” and the “broken mirror theory.” In the mentalizing theory, it is proposed that only the mentalizing network is abnormal, while at least the basic processing in the mirror system is normal and the same is valid also for the brain activity in the mentalizing

regions (Castelli, Frith, Happé and Frith, 2002). Thus according to this view, the difficulties displayed by the autistics are symptoms of an inability to represent other people's mental states.

In contrast, the broken mirror theory proposes that a core deficit in mirroring leads to difficulties with mentalizing; as a consequence, the developmental failure of the mirror system is perceived as the primary social difficulty in autism as well as the cause of its poor overall activity of mentalizing. Following this theory, deficits in understanding the kinematic aspects of an action and goals related to that action would advance further difficulties in understanding emotions and mental states (Baron Cohen et al. 2013, p.386). Additionally Kanner (1943), by considering that in the group of 11 children he examined during an experiment, all showed an “inborn autistic disturbances of affective contact” underlined how people “so long as they left the child alone, figured in about the same manner as did the desk, the bookshelf, or the filing cabinet” (Baron Cohen et al., 2013, p. 246), a case that clearly illustrate children's lack qualities of self-awareness as well as their impairment in connecting with others.

As it was noted in various sections, Theory of Mind reflects an understanding that people have mental states (desires, beliefs, intentions) that are linked to feelings and behaviors, so that, typically, children of 3 or 4 years of age start to show some of these abilities by recognizing that others have beliefs about things in the world, which are sometimes false or different from their own. The measurement of false belief (or first-order ToM) is generally regarded as key marker in children's understanding of mental states and its emergence at early age in development has been characterized as a crucial shift in the improvement of their thinking capacities. A large body of research has found that children and adolescents with autism spectrum conditions (ASD) experience difficulties understanding mental states or show a delay in the ToM development (Baron-Cohen, 2000). Having linked the positive associations between ToM and social behavior in typically developing children, has given prominence to investigate ToM as one of the main mechanism that potentially underpins social behavior in autism. Moreover, several studies have developed methods to teach ToM to the supposed mind-blinded population, with the aim of establishing whether it could be possible to promote an improvement in

the process of ToM understanding through intensive teaching by using structured progressive techniques (Swettenham, 1996).

Human communicators usually create the prerequisites for a successful joint communication by adjusting for their purposes the linguistic and communicative features as needed, so that human communicative acts are grounded in jointness and are by nature a set of shared conventions. Other primates do not structure their communication in the same way with joint intentions, an activity that mutually assume cooperative motives and communicative conventions; rather, they simply attempt directly to predict or manipulate the individual goals, perceptions and actions.

The most systematic comparative study aimed at proving the reliability of this hypothesis is that of Carpenter, Tomasello and Savage-Rumbaugh (1995; see also Bard and Vauclair 1984, for some similar observations). They observed human 18-month-olds and a chimpanzee in interaction with an adult human female target and some objects, with a focus on their objectively looking patterns. In this situation, the two species interacted with objects and notably the chimpanzee simultaneously monitored the adult human's behavior with a certain frequency. The result was that the human infant spent far more time than the ape looking back and forth from object to adult and, interestingly, their looks were almost twice as long as those of the apes, sometimes being accompanied by smiles, whereas apes do not smile at all. These differences helped in drawing the conclusion that the ape's looks to the adult were "checking" looks (to see what the adult was doing or was likely to do next), whereas the infant's looks to the adult were "sharing" looks (to share interest). In this way it was stated that the apes knew that the other had goals and perceptions, but they had not the ability or desire to share them, so that it was concluded that apes interacted with others and with the objects around them, but they did not engage in shared activities sustained by common goals and experiences.

Tomasello and Carpenter (2005) found something very similar when they had an attempt in encouraging three human-raised chimpanzees to share attention with a human in the context of a play with objects. They found that the chimpanzees sometimes looked at the interacting human to check what she was doing, but they did not look to her with the aim of sharing interest in what she was doing nor did they attempt to initiate an

activity of joint attention by communicating gesturally with her. Thus, although great apes share with humans the ability to learn instrumental actions for sharing intentions with others, the social function of imitation and the resulting evidences from various experiments concur in stating that the ability of conforming to a set of norm groups seem to be an uniquely human capacity (Tomasello, 2010, p.196).

Since identification may find expression in imitative acts, this process is one that has cognitive, conative and affective aspects: through others and in virtue of identification, a person is moved in orientation, which is a motivational and emotional as well as cognitive process. It may thus be that humans' evolved abilities to represent the world iconically in gestures for purposes of interpersonal communication manifests itself in ontogeny in several different ways as a result of the emergence of a vocal language. To identify with someone else is to feel something of the psychological stance expressed by the other and, potentially, to exploit this ability in order to relate to the others as creatures oriented toward the world; notably, this ability seems to develop over the first year of a typically developing infant's life. Unfortunately, children with autism tend not to be moved by external emotions and events; as a consequence, without the powerful advantage of the engagement with other people perceived as other minds since birth, their cognitive and social development is severely compromised.

In the "Hello- goodbye" study conducted by Hobson and Lee (1998), children with and without autism were introduced to an unfamiliar target and, in order to capture the spontaneous greetings and farewells of these subjects, a professor videotaped participants as they entered and departed from a familiar but empty classroom in which there was a stranger. The results were as follows:

Compared with participants without autism, there were about half as many of those with autism who gave spontaneous expressions of greeting in the 'Hello' episode, and a substantial proportion failed to respond even after prompting. All the young people without autism made eye contact, but a third of those with autism failed to do so; no fewer than 17 out of 24 of the former group smiled, but only six out of 24 of those with autism. In the 'Goodbye' episode, half the individuals without autism but only three of those with autism made eye contact and said a goodbye. And not only were there few participants with autism who waved in response to PH's final prompt, but also their waves were strangely uncoordinated and limp (Zlatev, 2009, p. 87).

These findings further sustain the hypothesis that children with autism are limited in the degree to which they do not identify with another person during a conversation.

An additional study was made by Hobson and Hobson (2007) where a set of children were invited to watch videotapes of participants interacting with an adult and, successively, to judge the quality of each look they made to the adult's face. Interestingly, it was found that an imitation of self/other-orientation occurred when the child adopted the examiner's demonstrated self/other anchored orientation, thereby reversing the positioning of the object and directedness of the action. The methodology was as follows:

As a first step, two raters were found to agree in the amounts of time for which children directed their gaze to the tester. The next stage involved raters judging each look with respect to the quality and/or function of the look according to the following mutually exclusive and exhaustive scheme. 'Sharing looks' were defined as those looks directed to the tester that could be seen to express a participant sharing experience through interpersonal contact with the tester. They involved a deep gaze which conveyed personal involvement with reciprocity, depth and affective contact, in contrast to checking looks that involved superficial glances at the tester and were more superficial and lacking in mutuality. 'Checking looks' were defined as those looks towards the tester that were used in order to assess or check out either the situation, the tester's response, or to determine what might happen next. Each of the three forms of looking were less prevalent among participants with autism, and most of the participants in each group showed some 'checking' and 'orientating' looks. However, two-thirds of participants with autism never showed a 'sharing' look, whereas this was the case for one-third of the comparison group. Concerning our critical prediction that sharing looks, and only sharing looks, would relate to imitation of self/other-orientation, it turned out that indeed, participants in each group who showed sharing looks tended to be those who imitated the demonstrator's self/other-orientation, whereas all those with the lowest scores of self/other-orientation also showed a complete absence of sharing looks (Zlatev, 2008, pp.72-4).

Thus the conclusion of Hobson and Hobson was that sharing involves a structure of interpersonal engagement (involving identification) that is easily overlooked until it becomes manifest through imitative self/other-orientation. Notably, in the Hello-Goodbye study, individuals with autism are less likely to orientate to and affectively engage with a stranger or to depart with typical gestures of farewell, namely, they lack an inter-subjective attitude in relation to the bodily and expressive dispositions that are thought to reside at the core of an irreducibly interpersonal experience.

By and large, it would seem worthwhile to think about what sharing might involve. As Trevarthen (1979) pointed out, a necessary division between primary and secondary

inter-subjectivity is needed as primary inter-subjectivity concerns the transactions that go on between two people, paradigmatically in face-to-face engagements, while secondary inter-subjectivity concerns those transactions implicating the shared experience of some real or imagined object or external event to the two or more people involved in the situation. As a case described by Hobson (2005) had shown, a typical developing two-months-old videotaped during a still-face procedure as part of a study of mother-infant relationship, reveals that when the mother assumed a still-face and an unreactive posture the infant responded by becoming uneasy, restless and jerky in his movements; this result can be considered as a first example of how an newborn participates and experiences the others since his first days.

The key point of these studies is that movements in subjective orientation are essential for what is so special about human life, namely the transitions in thoughts as well as in feelings since infancy through all life: to be moved by others requires necessarily an emotional engagement, a process that becomes possible due to special forms of interpersonal engagement. Initially, there is a dyadic meeting (as described in the study on infant joint attention mentioned above) and subsequently there is also an engagement with others' thoughts and behaviors through which the target can be moved in attitudes towards the world (as in social referencing). Out of this evidence, emerges the human propensity to move within their own minds, from one psychological orientation to another, an ability that can be considered the core of the human sharing capability (Zlatev, et al., 2008, pp. 94-6). In children affected by autism, this complex activity is impaired at various the degree, so that the empirical studies on autism is useful insofar it provides critical insights into the development of social understanding and empathy.

Autism is defined as a lifelong, complex developmental disability that typically appears during the first three years of life and affects the way a person communicates and relates to people for his/her entire life. Since 1943 when Kanner described for the first time early infantile autism and Asperger (1944) identified the first forms of autistic psychopathy, a lot has been learnt about these pathologies but still not enough to truly understand their bewildering condition. Autistic spectrum disorders include a wide range of deficits in social cognitive skills and people are affected by it to different degrees so

that some individuals manifest severely compromised attitudes while others present just a weaker impairment. Since the main point of the inability of mind-reading is described in terms of a lacking of meta-representation abilities, autism is helpful to understand the disruption of the relation self-other, known as “mind-blindness”.

As far as the methodologies regarding the study of the phenomenon of autism in relation to Theory of Mind and joint attention are concerned, one of the most widely respected approach to formal research diagnosis is the Autism Diagnostic Observation Schedule, a semi-structured series of planned social presses to prompt requests, and engagement in joint attention and communication with the tester which has turned out to be extremely effective for studying certain types of interpersonal exchanges that are often limited in autism. Moreover, since it has been noticed that autistic people find it hard to perceive the mental states of others, Simon Baron Cohen (1995) suggested that the central feature of autism can be described as an inability to infer another person’s mental states and, in extreme cases, autistic children may have no concept of mind at all, impairment that has been described in terms of a lacking Theory of Mind (Zlatev et al., 2008, pp.84-5).

In one widely used ToM task which is known as the Sally–Ann test, the child sees a doll named Sally watching another doll named Ann placing a marble in her basket. When Ann leaves, Sally moves the marble from the basket into her box; the child is asked where Ann will look for her marble when she returns, namely whether in the basket or the box. Normal children (and most children with Down syndrome) will correctly infer that Ann will look in the basket because she does not know that the marble has been moved while the autistic ones repeatedly fail the test. The reason for this result seems to be that for non-autistic people the answer is spontaneous, while individuals with autism elaborate a solution logically, so that such mind-awareness of autistic people shows that they engage deeply in ToM tasks because they find these tests more difficult than the non-autistic children. In conclusion, it is useful to underline that inter-subjectivity has a self-other structure which is highly important in the process of identifying with others. Studying children with autism seems relevant for further reinforce this theory since one of the most promising starting points for understanding, diagnosing and treating autism is

to consider how inter-subjective engagement shapes the development of social relations. Regarding that, recent advances that affective neuroscience has obtained in the last decades supports the dynamic developmental approach to ToM described above so that the roots of the so-called “mind-blindness” seen in autism can be traced back to biological deficits that appear much earlier than what is generally presumed in the modularity view of ToM.

As far as the development of ToM abilities is concerned, the important point seems to be that a child’s unconscious processing of the non-verbal affect signals of those around him/her is forged in the earliest years of development, when the brain is going through its massive growth spurt and subsequent pruning. At this stage, the connections formed between the amygdala and the paralimbic and prefrontal neural systems influence how an individual responds to certain types of social stimuli, not only in childhood but for their entire life. As a consequence, failures in mind-reading seen at four years seem not to be the product of a faulty ToM, or a lower-level system related to it; rather they are the by-product of compromises in the social brain network.

According to the original Theory-Theory account of social cognition, children normally develop during the first years of their lives a succession of Theory of Mind abilities that, just like scientific theories, postulate abstract, coherent mental entities and laws which enable them to provide predictions, interpretations and explanations to the events that happen in their lives, a mechanism which successively enable them to successfully interact with the people inhabiting their environment. Unfortunately, individuals affected by autism or Asperger’s syndrome are unable to do this, as clearly demonstrated by the series of studies mentioned above.

Conclusion: Do We Really Need a Theory at all?

In this dissertation, I have made an attempt to analyze the various positions on the Theory of Mind debate by taking into account both the classical standpoints of Theory- Theory and Simulation Theory and the recent developed alternative accounts to these well-established approaches, trying to underline both the positive and negative aspects of the various hypothesis about how we, as human beings, make sense of other's intentions, goals and beliefs from childhood to adulthood while engaging in a social, collective world. By doing so, I have considered both philosophical and developmental alternative critiques in the Theory of Mind debate, with the aim of emphasizing the importance of an integrated, mutual collaboration among the various positions in the ToM discussion in order to arrive at a fuller understanding of how the human brain works while engaged in both everyday interactions and in the literary experience.

My attempt in providing a description of the various parts of the brain which are supposed to “fire” during the narrative encounter is due to unify the neuroaesthetic of literature and the humanities, bridging in this way an historical gap between two apparently divergent disciplines in order to arriving at a more productive dialogue in which every branch can reciprocally enrich the other. In particular, the experience of reading is a place where the humanities and neuroscience can productively meet since how readers read and how literary works engage and manipulate these processes in various ways are matters about which neuroscientists and literary critics can say useful things to one another, even though their conversations will not solve entirely the problem of other minds and thus give a satisfactory answer to the question of which neurological functions are responsible for the mind processes. The clinical experiments of neuroscience may be useful to better understand literary experience since neurobiological studies of how the brain stimulates body states can help clarifying how aesthetic emotions affect the perception of the world. Neuroscientists exploring art and literature generally agree that there is no “art neuron” in the brain, so that the aesthetic experience is a complex process, distributed more widely over the brain than other manifestations such as color perception or facial recognition. Therefore, the results of the studies

mentioned in my work are not aimed at providing a better understanding of the absolute uniqueness of the perception of art but, rather, at showing the complex continuities and discontinuities between aesthetic experiences and the everyday life of our species.

Concerning the neuroscience of reading, I hope to have accurately shown how it draws important evidence from the studies of cognitive psychology as well as from aesthetics studies. Similarly, contemporary mirror-neuron research relies heavily on psychological observations of infant imitation in order to explain the mechanisms of inter-subjectivity, so that a productive dialogue between literary studies and neuroscience can emerge, a collaboration that will be facilitated if the mind-brain dispute can be bypassed by finding a way to correlate aesthetic experiences and neurological mechanisms. As far as the emotional response to literature is concerned, I have urged through my thesis that the importance of feelings in our engagement with both real and fictional characters has been historically underestimated while they are in actual facts extremely important as stories are about people in social contexts, having the great potential of showing us the dynamics of feelings at work without committing us directly to the situation described.

Furthermore, I have provided an account of how the phenomenon of empathy works and, insofar literature can foster or inhibit our willingness and ability to simulate other people's thoughts and feelings, an emotional reaction toward literary texts has always been the natural outcome of the reader's engagement with the work of art. Regarding that, I have provided some instances of rhetorical techniques employed by some writers in order to stimulate the reader's empathetic response and the effects these features usually have on the audience. For this reason, some hints at the issue about whether reading can influence pragmatically and morally their readers' lives have been provided, even though it still remains an open question. What has been proved instead, through the empirical studies on the brain of which I have provided some examples, is the existence of a physical, neural response in specific zones of the brain (especially AI) while a person is exposed to certain types of stimuli such as stories, so that the basis for a

neural, emotional empathy that has been theorized since ancient times nowadays seems to be scientifically proved.

In the last part, I have taken an insight into the phenomenon of “mind-blindness” with some instances of empirical studies on how Theory of Mind works and how it is impaired in individual affected by ASD, in order to underline the importance of having a Theory of Mind mechanism to fully and successfully engage with other minds in everyday circumstances. The term theory derives from the Greek word θεωρία which, according to the German philosopher Martin Heidegger (1889-1976), implies not only a certain degree of contemplation but also indicates the struggle to bring that contemplation into one’s activities in the world as the highest form of *energia*, an activity that represent human’s commitment in the world for asserting their existence over the nothingness. Accordingly, theories are by nature abstract and not content specific, a set of principles with the role of mediating between what the eyes see and what the mind perceives (thus allowing the contemplation and the comprehension of the aesthetic) which speculate about a theoretical issue in order to better define it, even though the thing per se remains ungraspable, escaping every kind of definition.

Specifically, Theory of Mind can be defined as an umbrella term which present a series of hypothesis that are possibly true but not necessarily or proven to be true just to some degree since there is no absolute method which can prove the physical existence of a ToM mechanism in our brain. Nevertheless, in the last decades the development of the empirical science with the introduction of newly developmental techniques such as brain imagines (fMRI, PET, EEG), helped gaining validity to the speculations made since the beginning of the 1980’s about the existence of a “resonance mechanism” in the brain responsible for sharing the feelings of others as well as for understanding and anticipating their intention with all the benefits that derives from this faculty. As I have tried to show by reporting some case studies conducted by cognitivists and neuroscientists on both children and adult’s ToM by submitting them to empirical tests, there are concrete evidences of the existence of a Theory of Mind in the human brain which is the tool supposed to be the real responsible for our normal interaction with the environment and

the society in which we live. Accordingly, the description of recent experiments pivoted by Baron Cohen, Hobson and Lee (1998) on individuals affected by ASD and thus carrying the so-called “mind-blindness” syndrome, were provided in order to better understand both the importance of having a Theory of Mind and to show the many ways in which this theory allows us as human beings to live a normal and serene life as well as enables us to engage and empathize with fictional characters’ experiences and lives.

Bibliography

Antonietti, Alessandro; Liverta-Sempio, Olga; Marchetti, Antonella. *Theory of Mind and Language in Developmental Contexts*, Canada: Springer Press. 2006.

Apperly, Ian, *Mindreaders. The Cognitive Basis of Theory of Mind*, New York: Psychology Press. 2010.

Aristotle, *Poetics*, USA: Penguin Classics. 1997.

Armstrong, Paul B. *How Literature Plays with the Brain: The Neuroscience of Reading and Art*, USA: The Johns Hopkins University Press. 2013.

Astington, Janet Wilde; Jodie A. Baird, *Why Language Matters for Theory of Mind*, New York: Oxford University Press. 2005.

Bakhtin, Mikhail, *Problems of Dostoevsky's Poetics*, USA: University of Minnesota Press. 1984.

Bakhtin, Mikhail, *The Dialogic Imagination*, USA: University of Texas Press, 1981

Baron Cohen, Simon, *Mindblindness: An Essay on Autism and Theory of Mind*, USA: The MIT Press. 1995.

Baron-Cohen, Simon; Lombardo, Michael; Tager-Flusberg, Helen, *Understanding Other Minds. Perspectives from Developmental Social neuroscience*, USA: Oxford University Press. 2013.

Bernaerts, Lars, *Stories and Minds: Cognitive Approaches to Literary Narrative*, USA: University of Nebraska Press. 2013.

Bogdashina, Olga, *Theory of Mind and the Triad of Perspectives on Autism and Asperger Syndrome, A View from the Bridge*, London: Jessica Kingsley Publishers. 2005.

Carruthers, Peter, *Language, Thought and Consciousness. An Essay in Philosophical Psychology*, Cambridge: Cambridge University Press. 1998.

Cohn, Dorrit, *Transparent Minds*, New Jersey, Princeton University Press. 1984.

Coleridge, Samuel T. *The Major Works*, New York: Oxford University Press. 1985.

Damasio, Antonio R. *The Feeling of What Happens. Body, Emotion and the Making of Consciousness*, USA: Vintage. 2000.

- Danta C, Helen Groth, *Mindful Aesthetics: Literature and the Science of Mind*, London: Bloomsbury Academic Press. 2014.
- Descartes, René, *Discourse on the Method and Meditations on First Philosophy*, USA: Hackett Publishing Company. 1999.
- Dostoevsky, Fyodor, *Crime and Punishment*, USA: Vintage Classics. 2008.
- Dostoevsky, Fyodor, *Notes From Underground*, USA: Vintage Classics. 1994.
- Elderman, Gerald, *The Remembered Present: A Biological Theory of Consciousness*, USA: Basic Books. 1990.
- Eilan, Naomi; Hoerl, Christopher; McCormack, Teresa; Roessler, Johannes, *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology*, USA: Oxford University Press. (2005).
- Elkman, Paul, *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, New York: Holt Paperbacks. 2004.
- Frackowiak, Richard S.J; Friston, Karl J; Frith, Christopher D; Dolan, Raymond J; Cathy J. Price, Zeki, Semir; Ashburner John T; Penny, William D. *Human Brain Function*, USA: Academic Press. 2003.
- Gallagher, Shaun; Dan Zahavi, *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*, Abingdon: Routledge. 2008.
- Gallagher, Shaun; Schmicking, Daniel, *Handbook of Phenomenology and Cognitive Science*, USA: Springer Press. 2010.
- Goldman, Alvin, *Simulating Minds. The Philosophy, Psychology, and Neuroscience of Mindreading*, New York: Oxford University Press. 2006.
- Guttenplan, Samuel, *A Companion to the Philosophy of Mind*, Oxford: Blackwell Publisher. 1994.
- Hemingway, Ernest, *Hills Like White Elephants in The Complete Shorts Stories of Ernest Hemingway*, pp. 211-14, New York: Scribner's. 1987.
- Herman David, *Storytelling and the Science of Mind*, USA: The MIT Press. 2013.
- Herman, David, *The Cambridge companion to narrative*, Cambridge: Cambridge University Press. 2007.

Hirschfeld, Lawrence A.; Gelman, Suzanne A. Mapping the Mind. Domain Specificity in Cognition and Culture, USA: Cambridge University Press. 1994.

Hobson, Peter, Autism and the development of mind, UK: Lawrence Erlbaum Associates Publishers. 1993.

Hoffman, Martin L. Empathy and Moral Development. Implications for Caring and Justice, USA: Cambridge University Press. 2010.

Hogan Colm, Patrick, Affective Narratology, USA: University of Nebraska Press. 2011.

Hogan Colm, Patrick, The Cambridge Encyclopedia of Language Science, New York: Cambridge University Press. 2011.

Hogan Colm, Patrick, The Mind and its Stories: Narrative Universals and Human Emotion, New York: Cambridge University Press. 2003.

Holland, Norman, Literature and the brain, Florida: The Psy Art Foundation. 2009.

Hutto, Daniel, Folk Psychological Narratives. The Sociocultural Basis of Understanding Reasons, USA: Bradford Books. 2008.

Hutto, Daniel; Ratcliffe, Matthew, Folk Psychology Re-Assessed, The Netherlands: Springer-Verlag. 2007.

Jaén, Isabel; Simon, Julien Jaques, Cognitive Literary Studies: Current Themes and New Directions, Austin: University of Texas Press. 2013.

Jakobson, Roman, Language in Literature, USA: Belknap Press. 1990.

Joyce, James, A Portrait of the Artist as a Young Man, New York: Oxford University Press. 2000.

Keen, Suzanne, Empathy and the Novel, New York: Oxford University Press. 2007.

Lakoff, George, The Political Mind: A Cognitive Scientist's Guide to Your Brain and Its Politics, USA: Penguin Books. 2009.

Lemon, Lee T; Reis, Marion J. Art as a Technique, ed. Russian Formalist Criticism: Four Essays, USA: University of Nebraska Press. 1917.

Leudar, Ivan; Costall, Alan, Against Theory of Mind, New York: Palgrave Macmillan. 2009.

Nichols, Shaun; Stephen P. Stich, *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, New York: Oxford University Press. 2003.

Oatley, Keith, *Best Laid Schemes: The Psychology of Emotions*, USA: Cambridge University Press. 1992.

Palmer, Alan, *Fictional Minds*, USA: University of Nebraska Press. 2008.

Premack, David and Guy Woodruff, Does the Chimpanzee Have a Theory of Mind? in *Behavioral and Brain Sciences*, 4: 515-526, 1978.

Raymond W. Gibbs Jr. *The Cambridge Handbook of Metaphor and Cognitive Sciences*, New York: Cambridge University Press. 2008.

Ratcliffe, Matthew, *Rethinking Commonsense Psychology: A Critique of Folk Psychology, Theory of Mind and Simulation*, New York: Palgrave Macmillan. 2007.

Richardson, Alan, *British Romanticism and the science of mind*, Cambridge: Cambridge University Press. 2005.

Richardson, Samuel, *Clarissa: Or the History of a Young Lady*, London: Penguin Books. 1985.

Rizzolatti, Giacomo; Sinigaglia, Corrado, *Mirrors in the Brain: How Our Minds Share Actions, Emotions, and Experience*, USA: Oxford University Press. 2008.

Ryle, Gilbert, *The Concept of Mind*, USA: University of Chicago Press. 2002.

Sartre, Jean Paul, *Being and Nothingness*, USA: Washington Square Press. 1993.

Schmicking, Daniel; Shaun Gallagher, *Handbook of Phenomenology and Cognitive Science*, New York: Springer Press. 2010.

Seemann, Axel, *Joint Attention. New Developments in Psychology, Philosophy of Mind, and Social Neuroscience*, USA: The MIT Press. 2012.

Stueber, Karsten, *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*, USA: The MIT Press. 2012.

Tomasello, Michael, *Origins of Human Communication*, USA: The MIT Press. 2010.

Trevarthen, Colin, Action and Emotion in Development of Cultural Intelligence: Why Infants have Feelings Like ours in *Emotional Development: Recent Research Advances*, ed. Jacqueline Nadel, Darwin Muir, p.61-4. 2005.

Trevarthen, Colin, *Communication and Cooperation in Early Infancy. A Description of Primary Intersubjectivity in Before Speech: The Beginning of Human Communication*, UK: Cambridge University Press. 1979.

Turner, Mark, *The Literary Mind: The Origin of Thought and Language*, USA: Oxford University Press. 1998.

Vermeule, Blakey, *Why Do We Care about Literary Characters?*, USA: The John Hopkins University Press. 2010.

Wellman, Henry M. *Making Minds. How Theory of Mind Develops*, New York: Oxford University Press. 2014.

Woolf, Virginia, *The Complete Works of Virginia Woolf*, London: Centaur Classics. 2014.

Wordsworth, William, 1799 *The Prelude*, London: Wordsworth Poetry Library. 1998. Vol.I-II.

Zahavi, Dan, *Self and Other: Exploring Subjectivity, Empathy, and Shame*, New York: Oxford University Press. 2014.

Zeki, Semir, *Splendori e Miserie del Cervello*, Torino: Codice Editor. 2010.

Zlatev, Jordan; Racine, Timothy P; Sinha, Chris; Itkonen, Elsa, *The Shared Mind. Perspectives on Intersubjectivity. Converging Evidence in Language and Communication*, USA: John Benjamin Publishers. 2008.

Zunshine, Lisa, *Introduction to Cognitive Cultural Studies*, Maryland: The John Hopkins University Press. 2010.

Zunshine, Lisa, *Why We Read Fiction: Theory of Mind and the Novel*, Ohio: Ohio State University Press. 2012.

