**UNIVERSITÀ DEGLI STUDI DI PADOVA**

**Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Dipartimento di Psicologia Generale**

**Corso di laurea Magistrale in Psicologia Clinica dello Sviluppo**

**Tesi di laurea Magistrale**

# Empathic processes and conscious perception: intersubjectivity at the basis of consciousness

***Relatore***
**Prof.** Paola Sessa

***Correlatore esterno***
**Dott./Prof.** Giorgia Silani

***Correlatore***
**Dott.** Filippo Gambarota

***Laureando/a: Tania Di Tillio***
***Matricola: 2016868***

Anno Accademico 2021/2022

Table of contents

# INTRODUCTION

Emotions and consciousness are among the most debated issues in psychology and neuroscience, since they represent experiences deeply rooted in our human nature. In fact, it is complex to disentangle and analyze these concepts, as they are commonly understood at an implicit level. In this study, the aim is to explore the relationship between empathy for pain and consciousness and, precisely, if and how an emotional context can shape our awareness of the others' emotions and our empathetic processes. The contextual cues that have been considered for this purpose are related to language and they are represented by written sentences. In addition, the emotional context analyzed in this work is related to painful situations. Finally, psychophysiological signals have been recorded to investigate if the effect that has been hypothesized has specific bodily correlates in terms of arousal.

The underlying purpose of this research is to explore the possible bases of conscious perception when this latter is framed by an emotional context. The relevance of this study can be found in the investigation of the factors that enable us to be aware of emotional cues, especially when we are in dangerous, emergency or aid situations. In fact, it can be hypothesized that being conscious of the emotion of another person when he or she is in pain could facilitate prosocial behaviors related to empathetic processes, thus promoting an active helping intervention.

The theoretical basis of this study is multifaceted, encompassing different topics whose relationship is intended to be analyzed. First, the difference between conscious, preconscious, and subliminal processing will be outlined, followed by a review of consciousness theories in neuroscience. Then, predictions and the predictive coding framework will be discussed, with a specific focus on the relationship between top-down expectations and conscious perception of the stimuli. Finally, the area of emotions will be presented, taking into consideration the theory of constructed emotions of Feldman-Barrett (2017) and exploring the most relevant empathy conceptualizations. These theories and areas of research are framed by an embodiment perspective, since there is a consistent literature that agrees on the statement that it is not really possible to separate mind and body and they continuously interact with each other (Rosch, Varela, & Thompson, 1991).

In the last part of the dissertation, limits and implications of the study will be outlined. In particular, a developmental and psychodynamic point of view will be applied to the more experimental part, in order to provide a wider view on the results and the theoretical background presented at the beginning of the thesis. In fact, it is known that experimental paradigms are conceived to test a hypothesis, but they are not perfect reproductions of internal processes, both psychological and neuronal. Based on this assumption, the limits of the research will be analyzed, and future directions will be outlined.

# 1. CONSCIOUSNESS AND PREDICTIVE CODING FRAMEWORK

## 1.1. CONSCIOUS, PRECONSCIOUS, AND SUBLIMINAL PROCESSING

Dehaene, Changeux, Naccache, Sackur and Sergent (2006) have investigated the brain bases of conscious and non-conscious perception in the light of global neuronal workspace hypothesis. This conceptualization (Dehaene, Kerszberg, & Changeux, 1998) discriminates two main computational spaces: a unique global workspace comprising distributed and strongly interconnected neurons with long-range axons, and a set of specialized and modular perceptual, motor, memory, evaluative, and attentional processors. Workspace neurons are activated in effortful tasks for which the specialized processors are not sufficient. They selectively switch on or off, through descending connections, specific processor neurons input. All along task performance, workspace neurons become spontaneously coactivated, creating discrete though variable spatio-temporal patterns that can be modulated by vigilance signals and selected by reward signals.

More recently, Dehaene et al. (2006) have suggested that conscious perception founds its associated neural mechanism into spikes of parieto-frontal activity leading to top-down amplification. Different authors have tried to explain the brain bases of conscious processing, but a coherent picture from these studies is hardly emerging (Zeki, 2003; Dehaene & Changeux, 2005). Dehaene et al. (2006) have suggested that these apparent contradictions can be solved through a theory of the physiological conditions of consciousness. The authors constructed a taxonomy of brain activity states associated with conscious and non-conscious processing, and within the latter one they identified a transient "preconscious" state of activity, indicating that information is only potentially accessible. Consciousness has several meanings and, in particular, the expression "states of vigilance" indicates the non-transitive nuance, i.e. a continuum of states which encompasses wakefulness, sleep, coma, anesthesia etc. In a nutshell, vigilance represents a graded variable, and a minimum level is necessary to place thalamo-cortical systems into a receptive state. The second meaning of consciousness is the transitive one, and it refers to the experience of consciously perceiving a (visual) stimulus or to the access to conscious report. Hence, it could be stated that the transitive meaning of consciousness indicates the state, while the intransitive something more dynamic. The authors suggest that early sensory activation is necessary but not sufficient for conscious access, because the activity in the extrastriate visual areas is often observed when participants deny having seen the stimulus (Moutoussis & Zeki, 2002). Besides vigilance and bottom-up activation, Dehaene et al. (2006) suggest a third factor underlying conscious access: the spread of brain activation to higher association cortices interconnected by long-distance connections and creating a reverberating neuronal population with distant perceptual areas. This brain state results in two main changes: since the activation reverberates, the information is retained for a long period of time and can be propagated to many others brain systems. Furthermore, considerable evidence indicates that without attention, conscious perception cannot occur (Mack & Rock, 1998). Both bottom-up stimulus strength (e.g. emotional stimuli) and top-down attention amplification are jointly needed for conscious perception, but they might not always be sufficient for a stimulus to cross the threshold for conscious perception. In conclusion, this latter has to be measured through subjective report (Dehaene et al., 2006).

The above-described distinctions lead the authors to suggest a tripartite conceptualization of two non-conscious processes and a conscious one. Firstly, subliminal processing ("below the threshold") indicates a condition of information inaccessibility where bottom-up activation is not sufficient to trigger a large-scale reverberating state in a global network with long range connections. In this case, the activation spreads but remains weak and quickly dissipating. Secondly, preconscious processing has been conceptualized by Freud (1940) as a process that involves information that are "unconscious but capable of entering the consciousness", i.e. potential conscious information. There is enough activation for conscious access, but it is temporarily buffered in a non-conscious store because of a lack of top-down attention, so that even strong stimuli can remain in this state. If the central workspace is freed, a preconscious stimulus might achieve conscious access, whereas this is impeded if the preconscious buffer is erased before orienting top-down attention. In conclusion, during this kind of processing activation can be strong and spreadable, but it requires top-down attention to reach parieto-frontal areas through long-distance connections. Finally, conscious processing represents a process where activation invades parieto-frontal areas, can be maintained, and guides intentional actions, like verbal reports.

To summarize, a stimulus is consciously perceived when it activates in a synchronized, long-lasting manner, a set of "central workspace" neurons, particularly present in parietal, prefrontal and cingulate cortices, and whose long-distance connections enable propagation to many distant areas. On the other hand, a stimulus could fail to become conscious for two reasons: it might not have enough bottom-up strength (subliminal stimulus), or a temporary withdrawal of attention (preconscious stimulus). The authors' proposal could integrate other two theories of consciousness. The first one is Lamme's hypothesis (2003) of a progressive construction of interactions, first locally (visual system) and second more globally (parieto-frontal regions). The second one is Zeki's proposal (2003) of an asynchronous build-up of perception in different sites before leading to a macro-consciousness. The only difference is the conception of pre-conscious processing, that they have attributed to micro-consciousness or phenomenal consciousness. For the present authors, the only reason for attributing phenomenal consciousness to preconscious processing is the insight that perception of reality involves experiences that we are not always able to report fully (Dehaene et al., 2006).

## 1.2 THEORIES OF CONSCIOUSNESS: A PERSPECTIVE OF INTEGRATION

Northoff and Lamme (2020) have reviewed and discussed different theories regarding the neural basis of consciousness. They suggest that they might be related to distinct aspects of neural activity and consciousness, which can in some way be integrated. Some of the most outstanding and broadly discussed neuroscientific theories of consciousness encompass: Recurrent Processing Theory (RPT) (Lamme, 2010), Synchrony Theory (ST) (Engel and Singer, 2001), Integrated Information Theory (IIT) (Tononi, Boly, Massimini, Koch, 2016), Global Neuronal Workspace Theory (GNWT) (Dehaene, Changeux, & Naccache, 2011), Temporo-spatial Theory of Consciousness (TTC) (Northoff and Huang, 2017), Predictive Coding Theory (PCT) (Hohwy, 2013), Higher-Order Thought theory (HOT) (Brown, Lau, & LeDoux, 2019), Operational Space-time theory (OST) (Fingelkurts & Neves, 2010), Entropy theory of consciousness (Carhart-Harris, 2018), and Embodied Theory (ET) (Park & Tallon-Baudry, 2014).

The differences between these theories can be categorized with respect to various characteristics (Northoff & Lamme, 2020). For instance, each of them aims at explaining a different target of consciousness. For example, phenomenal consciousness is the explanandum in RPT, IIT, ST, and TTC, whereas GNWT and HOT investigate more cognitive aspects of consciousness (e.g. awareness or access to consciousness) associated to functions like top-down attention. Other theories address more general processes of perception (content of consciousness in perception and cognition), such as PCT, or the relationship between perceptual states and action, body, emotions, or the self. Another characteristic which differentiates these theories regards the technical aspects concerning the studies. These are related to the targeted neural measure (stimulus/task evoked activity or resting state activity), the subjects included in the study (healthy or pathological groups), and the experimental paradigm used to test the theory. With respect to the latter, report paradigms are required to investigate cognitive consciousness aspects (e.g. access or awareness), whereas phenomenal features can be studied through non-report paradigms. In addition, perceptual aspects can be tested through masking or rivalry, while visual attention and memory are preferentially investigated through attentional blink or change blindness (cognitive designs). Finally, other authors indicate that emotion and affect represent the best “ground” to study consciousness (Damasio, 2010; Solms, 2019).

The last diversity tackled by Northoff and Lamme (2020) concerns the different aspects of neural activity addressed by each of the theories mentioned above: stimulus-related activity, pre-stimulus activity and spontaneous activity. Various conceptualizations identify different neural correlates of consciousness and, consequently, they rely on different measures targeting distinct aspects of it. The first aspect of neural activity taken into consideration is stimulus-related activation, which is the activity evoked after the presentation of the stimulus and represents the conscious percept of it (Neural Correlate of Consciousness, NCC). This activation can be viewed in spatial and temporal terms: respectively, elicited in specific regions of the brain and occurring in early or subsequent time intervals. Regarding the first dimension, different theories assume different areas of the brain to be important for consciousness, which depends on the explanandum. Some theories (like IIT and RPT; Lamme, 2006; Tononi et al., 2016) hypothesize that posterior regions are important for consciousness, which in this case is conceptualized as phenomenal/experiential, or "integrated information" (i.e. binding of features of the object and perceptual organization). So, from this perspective, the key feature of conscious percepts is their unity or "wholeness", and the amount of integrated information is referred to as "Phi". The so-called "posterior hot zone" (i.e. the combination of visual, other sensory and parietal cortices) is considered prone to have high Phi and consequently should be sufficient to construct a conscious sensory experience (Boly et al., 2017). Other theories (like GNWT and HOT; Dehaene et al., 2011; Brown et al., 2019) postulate that frontal regions are necessary for consciousness. The GNWT targets specifically the dorsolateral prefrontal cortex (DLPFC), that is considered the basis of the conscious "access" to the contents. It is a core region of the Global Neuronal Workspace, which is the area where sensory information is made globally "accessible" to other cognitive functions (Northoff & Lamme, 2020). In this process a key role is assigned to top-down attention. In HOT, sensory information is considered a "first order" representation, not sufficient for conscious experience. Then, a "second order" re-representation is necessary for the transition of this information to consciousness. In addition, thalamo-cortical connections are assumed to be important in the integration of two aspects of consciousness, state and content (Sanchez-Vives, Barbero-Castillo, Perez-Zabalza, Reig, 2020). Eventually, consciousness featured by both content and state/level is constructed from subcortical-cortical interaction, but the exact process underneath is still not known. On the other hand, when investigating the temporal dynamics of stimulus-related activity, it is important to define two kind of visual input processing (Northoff & Lamme, 2020): the feedforward and the recurrent. The feedforward sweep of information processing represents an early stimulus-related activity, as measured by the N100, and is not related to consciousness. It encompasses the extraction of various features of the stimuli by all parts of the visual brain. Afterwards, recurrent processing occurs, with feedback connections between higher and lower-level areas of the brain (Lamme, Super, & Spekreijse, 1998). The notion of recurrent interactions is complex, involving cortical and subcortical information integration. Giving these conceptualizations, it can be stated that the explanandum is the variable that has to be considered when investigating the timing of conscious perception after stimulus presentation. For instance, RPT theory (which refers to phenomenal consciousness) hypothesizes that conscious perception starts after 100-200 ms, which corresponds to the moment where recurrent processing sets in (Northoff & Lamme, 2020). Whereas GNWT, that considers conscious access, points to the later activity (P300), which refers to the moment when sensory information becomes available for different parts of the brain (Dehaene & Changeux, 2011). There are other characteristics of stimulus-related activity that may be relevant to consciousness (Northoff & Lamme, 2020), for example the degree of activity synchronicity between different neuronal populations. Synchronous activity between neurons in a population refers to "perceptual binding", which may not necessarily represent consciousness (Hermes, Miller, Wandell, Winawer, 2015). Nonetheless, especially for high frequency synchrony (i.e. gamma band oscillations), some theories link this process to consciousness (Fingelkurts & Neves, 2010), taking into consideration the specific explanandum (i.e. perceptual organization). Another aspect of stimulus induced activity is the complexity of the signal evoked, which measure is the "perturbational complexity index" (PCI) and reflects our brain capacity to integrate information (D'Andola, Rebollo, Casali, Weinert, Pigorini, Villa, Massimini, Sanchez-Vives, 2018). A further feature of stimulus-evoked activity is the "trial to trial variability" (TTV). Reduced TTV after stimulus onset (i.e. TTV quenching; Churchland et al., 2010) indicates that the variability in the amplitude of responses to the same stimulus is also reduced, thus enhancing the stability of the signal. TTV quenching has been related to consciousness because it has been hypothesized that this represents a suppression of the brain's intrinsic noise by the external stimulus. The possible mechanism underneath may be that a conscious stimulus evokes more synchronous/recurrent activity, thus stabilizing the signal (Northoff & Lamme, 2020).

The second aspect of neural activation tackled by the authors (2020) is pre-stimulus activity. Pre-stimulus activity provides a dual role: mediating content of consciousness and its associated level of arousal. In addition, the degree of change during post-stimulus activity with respect to the ongoing dynamics of the pre-stimulus activity is central for the conscious processing of the external stimuli.

Therefore, there has to be a specific interaction between pre- and post-stimulus activity in order to associate consciousness to contents: that is, strong TTV quenching (i.e. stimulus-related suppression of the ongoing pre-stimulus variability; Churchland et al., 2010). Starting from these findings about the relevance of pre-stimulus activity for consciousness, Northoff (2013) has proposed the Temporo-Spatial Theory of Consciousness (TTC). It concentrates on how the impact of the external stimulus on the brain depends on the brain's pre-stimulus activity: an external stimulus has to interact with ongoing activity such that the two become integrated into the current stream of consciousness. Thus, this theory focuses on the importance of pre-stimulus activity's spatiotemporal dynamics: the pre-stimulus activity can expand the stimulus's actual points in time and space beyond themselves (i.e. "going beyond", Buszaki, 2006). The last characteristic is spontaneous activity, which has two dimensions: a spatial and a temporal one (Northoff & Lamme, 2020). This kind of activation is reflected in various networks organized in a small-world way, and it shows an oscillatory pattern (measured in frequencies). Taking these two features together, it can be observed that the spontaneous activity's structure is not static but dynamic (i.e. it continuously changes its configurations). The relation between different frequencies of neuronal activity can be described as scale-free, scale invariant or self-similarity (i.e. the relationship between the power of the frequencies is the same irrespective of the spectrum of frequencies considered). Related to these concepts (Huang, Obara, Davis, Hap, Pokorny, & Northoff, 2016), different measures of spontaneous activity can be introduced, that are connected to consciousness. The first two are "power law exponent" (PLE) and "detrended fluctuation analysis" (DFA), which indicate the degree of self-similarity or scale-free activity and have been associated with different aspects of consciousness (Northoff and Lamme, 2020). Another measure is entropy, where a higher degree of it may lead to extended consciousness (Carhart-Harris, 2018). Finally, the last index considered by Northoff and Lamme (2020) is complexity, measured through the Lempel Zev Complexity (LZC) index.

To date, there haven't been found functional connectivity patterns and neural networks specifically associated with consciousness. Thus, researchers have been studying alternative measures of the spontaneous activity spatiotemporal dynamics during altered states of consciousness. They have found that the difference between the presence and the absence of consciousness is essentially represented by the relative differences in spatial dynamics (i.e. frequency of particular spatial patterns). In addition, it has been discovered that both a decrease of spatial dynamics and of temporal dynamics signals the absence of consciousness. On the other hand, studies on drug-induced psychosis demonstrated increased spatiotemporal dynamics of the spontaneous activity (e.g. entropy and complexity measures) while using different psychedelic drugs (Atasoy, Roseman, Kaelen, Kringelbach, Deco, Carhart-Harris, 2017). Along with its spatiotemporal dynamic, another aspect of the brain's spontaneous activity is the constant interoceptive input from the own body (Northoff & Lamme, 2020). The interoceptive processing and hence the body is the focus of researchers who posit that embodiment, including subcortical-cortical brain-body relation, is fundamental to consciousness (Azzalini, Rebollo, Tallon-Baudry, 2019; Park & Blanke, 2019). On the other hand, it has to be considered that these theories refer more to the potential neural correlates of the first-person perspective, which may consequently be the heart of consciousness to which phenomenal experience is integrated (Park & Tallon-Baudry, 2014).

Northoff and Lamme (2020) have observed that theories of consciousness agree on what is the role of the operational architecture underlying consciousness: the transition from independent to mutually interdependent (i.e. integrated and distributed) neural activity, reflecting the dynamic/temporospatial basis of consciousness. The integration refers to the fact that neurons are sharing the information with other neurons. Another point of convergence between theories is that they hypothesize that different conscious contents and levels of consciousness are processed in different parts of the brain at different points in time, reflecting a heterogeneous process. Posterior regions may mediate the phenomenal aspects of sensory contents, and then neural activity may be spread to anterior regions that allow for cognitive processing of the same contents, leading to their access, report, knowledge, and meta-cognition related processes (Baars, Franklin & Ramsoy, 2013). A more recent theory to go beyond dichotomies is the Global Brain Activity (GBA), that is related to global effects (Liu & Luo, 2019). This theory posits that, in the pre-stimulus phase, the neural context for the following processing and perception of the visual stimuli is already set. GS (Global Signal) is a measure of GBA, and its level of decrease is related to the level/state of consciousness. Furthermore, it has been observed that there are diverse topographical distributions of GS in different states and in different forms of consciousness (Tanabe, Huang, Zhang, Chen, Fogel, Doyon, Wu, Xu, Zhang, Qin, P, Wu, Mao, Mashour, Hudetz, Northoff, 2020). In summary, GBA reconciles different locations and different dimensions of consciousness, i.e. level and content. Hence, different contents of consciousness (sensory, cognitive, affective, etc.) are all associated with phenomenal experience but are processed in different regions of the brain (Northoff & Lamme, 2020). Furthermore, it may exist a common neural mechanism which leads to phenomenal consciousness. Different theories have tried to explain the nature of this common mechanism.

TTC (Northoff & Huang, 2017) posits that this is represented by pre-post stimulus interaction with temporo-spatial expansion. In addition, GBA could represent another mechanism, where the same contents may undergo different levels of processing, allowing us to access them in different ways (Baars et al., 2013). To summarize, consciousness may be conceived as a heterogenous multifaceted neuronal process with different layers or levels of neuronal activity nesting with each other (Northoff & Huang, 2017). The concept of layers refers to nestedness: consciousness and the brain's neural activity are hypothesized to be characterized by layers that contain (i.e. nest with) each other. This means that both phenomenal and neuronal levels of consciousness are nested: there is a nested organization of the brain's neuronal activity and a nested organization of contents of consciousness. In summary, consciousness gestalt nature comprises a figure (local activity) and a background (global activity). Different theories have proposed different functions of consciousness, that is an integrated, distributed and interdependent kind of processing (Northoff & Lamme, 2020). The function may be sensory (perceptual organization), cognitive (access or prediction), bodily (neural monitoring of bodily input). A recent proposal by TTC theory, is that the function underlying consciousness is a temporo-spatial dynamic process, which mediates sensory, bodily and cognitive functions by operating across different regions, and it is content-unspecific.

The next question addressed by Northoff and Lamme (2020) is the nature of the interaction between the three forms of neural activity related to consciousness (stimulus-related, pre-stimulus, and resting state activity). In fact, the TTC suggests the existence of different temporo-spatial mechanisms that relate to different aspects of consciousness and different forms of neural activity (Northoff and Huang, 2017). The kind of interaction which is postulated by the TTC is reflected in the scale-free nature of consciousness: increased scale-free integration and, hence, temporo-spatial nestedness of the three forms of neural activity will produce consciousness. This must be tested empirically in future research. In addition, there are different dimensions of consciousness: state/level, content, form (Northoff, 2013) and they are related to three kinds of neural activity. The form is a new concept which refers to consciousness' structure or organization at a phenomenal level, like its complex gestalt with figure and background: it encompasses unconscioussness, phenomenal consciousness and reflective consciousness and it is mediated by the spontaneous activity's architecture (i.e. its temporo-spatial nestedness across subcortical and cortical regions). In addition, the content may be associated to post-stimulus and pre-stimulus activity (on the cortical level), and the state/level may be especially related to pre-stimulus activity. Furthermore, different neural activity is linked to different aspects of consciousness. Stimulus related activity is hypothesized to be sufficient for phenomenal consciousness (NCC proper; Koch, Massimini, Boly, Tononi, 2016) and its cognitive aspects or consequences of consciousness (NCCcon; Aru, Bachmann, Singer, Melloni, 2012). On the other hand, pre-stimulus activity is presumed to enable consciousness, so to be a neural prerequisite of it (preNCC). These are all neural correlates of consciousness, i.e. neural features that are present when there is consciousness, and they regard different time points of neural activity (Northoff & Lamme, 2020). PCT proposed a potential hypothesis for their relationship: according to this theory, pre-NCC is linked to NCC proper by the prediction error, that is a modification of the perception influenced by the top-down prediction through the bottom-up stimulus related activity. TCC theory explains this aspect through the presence of temporo-spatial dynamics, which can represent the connection between pre- and post-stimulus activity. In fact, becoming aware of the contents of consciousness requires both spontaneous activity and stimulus related activity. The first one represents the activity independent of specific stimuli or tasks, and it enables the neural capacity or predisposition of consciousness (NPC; Northoff & Huang, 2017). In extreme situations like anesthesia and coma, the basic shape of the spontaneous activity's power spectrum (i.e. its scale free nature) is not preserved, but it is replaced by low power values for both slow and fast frequencies. To summarize, the NPC represents a form of default activity (i.e. the brain's baseline) which, given that is necessary for consciousness, has been related to the brain's scale-freeness (Northoff & Lamme, 2020).

In conclusion the TTC, a novel approach to the study and understanding of consciousness, has hypothesized some basic similarity between neuronal and phenomenal states, that allows for a transformation of the former into the latter. Thus, the relevant innovation of this theory is that neuronal and phenomenal states are no longer conceived as different. This basic similarity has been called "common currency" (Northoff, Wainio-Theberge, Evers, 2019), and it represents the temporo-spatial dynamics of the brain's spontaneous activity, which in turn refers to the NPC: the neural predispositions of consciousness. The concept of "nestedness" can be included in this theorical framework by assuming that all the different types of neural activity related to consciousness are nested into each other, but the larger and overarching temporo-spatial frame is actually the NPC (Northoff & Lamme, 2020). In brief, NCCcon is nested in NCCproper that is in turn nested in pre-NCC, and all those are eventually spatio-temporally contained in the more inclusive NPC. Thus, the NPC provides a way to enable the phenomenal features of consciousness, that is the temporo-spatial dynamics of the brain's spontaneous activity. This latter represents a basic structure, a template (i.e. form) of the brain's inside (i.e. deep interior) that, as common currency, allows for assigning a phenomenal nature to external stimuli from the outside of the brain (i.e. body and world). This movement can be visualized as an "inside-out" and recurrent one: from the brain to the outer world, and vice versa. It is clear that this vision is clinically relevant: if the spontaneous activity's temporo-spatial dynamics no longer provides its neuronal activity basic structure, as in disorders of consciousness, or an altered one, as in psychiatric disorders, also the phenomenal experience of consciousness will be either lost or abnormally altered.

## 1.3 HOW PREDICTIONS MODEL PERCEPTION

Predictions are conceived as prior knowledge about the probabilistic structure of reality, so that they can shape our perception of the world and the way we act in the world (de Lange, Heilbron, & Kok, 2018). Humans, differently from what was the traditional belief about the mind, are anticipatory systems, not merely reactive ones (Rosen, 2012): they construct meanings of incoming data by building predictive models of themselves and their environment. In accordance with this view, also the brain has been conceived as a prediction machine, integrating bottom-up sensory data with top-down expectations (Clark, 2013). Expectations can influence perception in various ways. For instance, and relevant to our study, when sensory input is weak, noisy or ambiguous, predictions are more effective in biasing perception: they can even change also what is perceived (Chalk, Seitz, & Seriès, 2010). Furthermore, the relative impact of expectations depends on the validity or reliability of the predictions (Mumford, 1992): individuals rely more strongly on prior knowledge when expectations are reliable and stimuli are ambiguous. These two aspects of expectations can be summarized in a form of uncertainty weighting.

Predictions arise from statistical regularities of our experience of reality, and specifically from three sources (Seriès & Seitz, 2013): simple frequency distribution of sensory inputs, conditional probabilities and on the basis of one's own actions. In the first case, regularities emerge from the fact that certain aspects of a sensory input appear more often than others. These regularities are stable throughout an individual's lifetime, and thus allow to build prior expectations. These kinds of predictions become encoded in the tuning properties of our sensory cortices (Cloherty, Hughes, Hietanen, Bhagavatula, Goodhill, & Ibbotson, 2016). In some circumstances they can be modulated by recent experiences, implying interactions between long term and dynamic/context-dependent priors. With respect to the second case, it is known that the hierarchical structure of the visual world (that goes from lines and curves to objects and to scenes) is mirrored by the hierarchical organization of the visual cortex (Hochstein & Ahissar, 2002). Nevertheless, this hierarchy does not have a one-way direction, but there is an interaction between bottom-up and top-down processes. When bottom-up signal flows in the visual cortex, there is also a top-down connectivity which enables higher level representations to allow predictions about lower levels features of the reality. Moreover, statistical regularities can be found also between inputs from different modalities, and learning those expectations demands the involvement of higher-order brain regions. To conclude this part, it can be stated that since predictions have a central role in sensory processing, they exist at each level of the cortical hierarchy. However, cortical connections modulate slowly, long-time learning associations, whereas some predictions require to be learnt rapidly and flexibly. One brain region that can serve this purpose is the hippocampus (Lavenex & Amaral, 2000). In addition to this region, there may be an involvement of other higher-order brain regions specialized in forming complex associations, such as the prefrontal cortex (in particular semantic associations). The third case regards how expectations arise from our own actions (Eliades & Wang, 2008). The purpose of these kind of predictions is different from perceptual ones: the goal is to inhibit the representation/perception of expected sensory inputs.

The way in which predictions shape sensory processing is reflected in a phenomenon called expectation suppression: in other words, stimuli that are expected evoke a reduced neural response in sensory regions (Summerfield & De Lange, 2014). This phenomenon represents a general principle of cortical processing, and researchers have proposed two possible explanations that can be integrated. The first one is called the dampening account: the weaker response is due to the brain filtering out the expected components of sensory inputs, so a stronger response can be observed in the case of surprise-inducing stimuli (Friston, 2005). On the other hand, expectations dampen responses in neurons tuned for the expected stimulus. This is in accordance with redundancy reduction theory. The second one is called sharpening account: neurons encoding the unexpected features of the stimuli are silenced, producing a sharper, selective population response with a lower overall amplitude. This kind of response enables a sharpening of the underlying representations of the stimuli (Lee & Mumford, 2003). Ultimately, the authors (de Lange et al., 2018) conclude that these two theories have different roles which are not mutually exclusive: they could take place in separate neural populations encoding errors (sharpening) and predictions (dampening). Regarding when expectations influence perception, de Lange et al. (2018) have highlighted two possible and not mutually exclusive explanations, that could rely on the type of expectation. The first one posits that expectations modulate responses after the earliest components of it (the initial “feedforward sweep”). The second one hypothesizes that this happens even before the stimulus onset, proposing an anticipatory effect.

Research has hence demonstrated that prior expectations influence perception, but one aspect that has to be explored deeply is the conceptualization underneath this influence: what is the role of these modulations? Since the sensory inputs are ambiguous, perception has been conceived as a process of probabilistic inference, which means that our cognitive system chooses the most probable causes of the input data, based on the interaction between those and prior expectations (Gregory, 1980). Bayesian probability theory is a normative theory in accordance with these assumptions. Furthermore, it can be applied to a neural architecture through a hierarchical and recurrent process (Aitchison & Lengyel, 2017), by which the output (or "posterior") from a higher level becomes the input (or "prior") of a lower level. The recurrency of this inference allows the system to progressively reach a globally coherent interpretation of the object or the scene. This process is reflected in the cortex by the gradual constraint of successive areas' inferences in the cortical hierarchy, as the interpretation evolves.

One of the most relevant computational theories is predictive coding, that posits that the brain constructs an internal model of the world by encoding the possible causes of sensory inputs (Friston, 2005). The prediction error- i.e. difference between the input and the prediction- is propagated to higher regions for further processing, and perception arises from minimizing this error and matching the predictions to the input. This does not mean that only prediction errors are processed: as it has been stated before (de Lange et al., 2018), the resulting cortex model contains different neural populations that represent both the current best prediction and the associated error. Furthermore, predictive coding is a general theory about how the brain processes information, and it is not limited to the predictive process. On a clinical level, neurodevelopmental disorders like autism spectrum disorder (ASD) and schizophrenia can be partially explained by an atypical integration of priors and sensory inputs, i.e. aberrant expectations. Nonetheless, these explanations reside at opposite sides of a spectrum. While delusions and hallucinations in schizophrenia may be related to a misperception of inner states due to overly strong expectations (Powers, Mathys, & Corlett, 2017), perceptual atypicalities in ASD are hypothesized to indicate an impaired top-down expectations process (Happé & Frith, 2006). This results in a focus on local details at the expense of a global perception and representation of the reality, that leads to hypersensitivity to unexpected stimuli because the individual is not able to construct stable representations of the world (i.e. everything is new). To conclude, it is important to stress that these considerations about psychopathology are a partial explanation of some correlates of these disorders' pathogenesis and, in particular, they address the neural and cognitive aspects of these conditions. Eventually, they can serve to reach a more complete understanding of the brain function and its alterations (Montague, Dolan, Friston, & Dayan, 2012).

## 1.4 THE RELATIONSHIP BETWEEN PREDICTIONS AND CONSCIOUS PERCEPTION

Within the predictive processing framework (Clark, 2013; de Lange, Heilbron, & Kok, 2018) perception is understood as resulting from an interaction between bottom-up stimulus-related neural activation and top-down prediction-driven biases. In this framework, the brain continually formulates hypotheses (predictions) based on past experience about the hidden causes of its sensory input (Friston & Kiebel, 2009; Hohwy, 2013). Such sensory predictions constrain possible interpretations and shape perception (de Lange et al., 2018). It has for example been shown that predictions facilitate object recognition, they benefit perception when the perceptual context is ambiguous (Denison, Piazza, & Silver, 2011) or when sensory input is weak (Summerfeld, Egner, Mangels, & Hirsch, 2006). So, according to this theory, a double movement can be conceived: on one hand, our representations of the world that build predictions shape the way we perceive the external but also the internal world; on the other hand, the external or internal stimuli influence predictions introducing an error variable, mentioned above as "prediction error" (de Lange et al., 2018).

Alilović, Slagter, & van Gaal (2021) have investigated the relationship between predictions and consciousness, that is the first research question that has been addressed in the present study. Within this main subject, two specific issues have emerged: first, it has been questioned whether predictions can bias perception and increase subjective visibility report of stimuli; secondly, it has been investigated whether non-conscious predictions can affect perception at short time scales, given the quickly-decaying nature of non-conscious processes (Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006). In order to answer these questions, three experiments have been implemented by the authors (Alilovic et al., 2021). In these studies, it has been researched whether predictions induced by a seen or an unseen predictor stimulus (T1) affect subjective measure of awareness (experiment 1 and 2) or discrimination performance of the subsequent stimulus (T2; experiment 3). The results have shown that valid predictions increase T2 perceived awareness or discrimination, and these prediction effects depend on the fact that T1 has been consciously accessed before. To summarize, predictions influence perceptual visibility reports of visual stimuli and the ability to consciously access them, but this is dependent on whether the predictor has been consciously accessed or not. Prediction effects have been measured through the Perceptual Awareness Scale (PAS; Overgaard, Rote, Mouridsen, & Ramsøy, 2006; Sandberg, Timmermans, Overgaard, & Cleeremans, 2010). This scale emphasizes subjective perceptual aspects of awareness that might escape classical dichotomous (e.g., seen/unseen) measures or discrimination response.

Furthermore, the authors (Alilovic et al., 2021) have looked into a possible brain predictive mechanism, which is also reflected in the process underlying consciousness. It has been hypothesized that, at different brain spatial scales, prediction implementation and prediction error signals are enabled by feedforward-feedback interactions (Friston & Kiebel, 2009). At a cortical level, first there is a feedforward activation of sensory cortices, and then sensory neurons receive feedback from higher level cortical regions (Summerfield & De Lange, 2014). As mentioned above, influential models of consciousness posit that exactly this same mechanism is also necessary for conscious access (Dehaene & Changeux, 2011; Dehaene et al., 2006). This mechanism is represented by feedback loops spanning widely-distributed brain areas, enabling sustained activation and global availability of information. The idea that there are both feedforward and feedback signals reflects the above mentioned interaction between top-down and bottom-up processes in shaping our perception of the world (Clark, 2013). Based on these models and on Alilovic et al. (2021) study, both predictive processing and full-blown conscious access to stimuli require a high-level neural feedback (Boly, Massimini, Tsuchiya, Postle, Koch, & Tononi, 2017, 2013). To conclude, this continuity between neural and cognitive models could represent the relationship between predictions and consciousness at a neural level.

In Alilovic et al. (2021) research, the predictions were recently learned and so based on arbitrary stimulus associations. It remains unclear whether the effects found by the authors apply also to environmental priors, that is the ones acquired through life-long experience (developmental perspective; Brodski, Paasch, Helbling, & Wibral, 2015; Chang, Baria, Flounders, & He, 2016). It is possible that while more hard-wired priors acquired through life-time learning do not require conscious access to exert effects on perception, arbitrary associations between stimuli depend on consciousness as they need to be implemented at higher processing levels.

## 2. EMOTIONS AND EMPATHY

### 2.1. THE THEORY OF CONSTRUCTED EMOTIONS

The traditional view of emotions has considered them as separate faculties or mental categories having their own innate essence, that differentiate each of them from other emotions (Barrett, 2017). According to this perspective, emotions and cognition are two distinct categories that are often in conflict with each other. Following an inductive approach, the classical view has posited that the emotions we perceive as sharp concepts have their direct correlate in nature and, specifically, in the brain. On the other hand, Feldman-Barrett (2017) has proposed a novel theory of constructed emotions, that considers them as a nuanced experience of the single individual that cannot be categorized in a definite manner. The phenomenological knowledge of the emotion is viewed as different from the neural manifestation of it. The current view of the brain conceptualizes it as a complex network in place of a set of modules, which creates the possibility to establish complex spatiotemporal patterns throughout the organ (Rigotti, Barak, Warden, Wang, Daw, Miller, & Fusi, 2013). This perspective reconciles with the TTC (Northoff & Lamme, 2020), previously presented. When the concept of degeneracy (i.e. the ability of a set of neurons or representations to create examples of a single category in different contexts, that explains the complexity of our brain; Edelman, & Gally, 2001) is applied to the field of emotions, it can be observed how the classical view is no more consistent. In fact, this means that instances of an emotion are constructed by different spatiotemporal patterns in various neuronal populations, since diverse examples of a single emotion have not the same characteristics (Clark-Polner, Wager, Satpute, Barrett, 2016). Hence, degeneracy explains individual differences. In conclusion, considering the emotions as categories is just an abstraction that is not present in nature, but it is merely a way to represent them in our mind (Barrett, 2017).

According to Sterling and Laughlin (2015), the brain has evolved around a main purpose: allostasis, which refers to the process of predictively regulating the internal environment of the body, allowing the individual to grow, survive and reproduce. In order to accomplish the allostasis task, the brain has to construct an internal model of the body in the world (which is called “embodied simulation”; Barsalou, 2008), so that the organism can thrive and convey its genes to the subsequent generation. The brain creates a model of the “body in the world” physiological needs: it includes both a representation of the external world and another of the internal milieu (Barrett, 2017). This model of reality starts to be implemented even before birth and the representation and usage of these internal sensations is called “interoception” (Craig, 2015), which rises from allostasis. These interoceptive sensations are experienced as lower-dimensional feelings of affect, whose properties - valence and arousal - are considered basic characteristics of consciousness (Damasio, 1999). The brain creates the internal model, also called “affective niche”, based on past experiences that have been relevant for allostasis, in a recurrent manner (Barrett, 2017). This means that, in line with Northoff and Lamme’s (2020) view of consciousness, ongoing brain activity influences how the brain processes incoming sensory information. Hence, the traditional reactive accounts of emotions are no longer able to provide a valid interpretation of the phenomenon: the internal model our brain creates is not reactive, but predictive (Barrett, 2017). Predictions are embodied simulations of sensory-motor experience, and they are constructed in the service of allostasis. So, the brain simulations fine-tune the meaning of the incoming sensory input in a Bayesian manner, constructing perception and emotions (also on a psychophysiological level). Predictions anticipate events in our sensory reality through a top-down process. This hypothesis refers to what has been previously called “predictive coding” framework (Friston & Kiebel, 2009), which posits that the individual consciously experiences the consequences of these simulations (i.e. representations) of reality as an affect. As it has been stated in the previous chapter, predictions are considered inferences about the causes of sensory events, and they guide actions (Friston & Kiebel, 2009). Since they are based on previous knowledge, these representations are already available in the brain as ongoing or spontaneous brain activity (Barrett, 2017). In addition, unanticipated information (called “prediction error”) establishes a bottom-up process that starts from the sensory world and interacts with top-down predictions. These latter and prediction error oscillate at different frequencies in the brain, as it has also been observed by Northoff and Lamme (2020).

According to Barrett’s constructivist view (2017), the brain processes information as a conceptual system. This means that predictions are considered concepts, which are categorizations of the incoming sensory inputs. The meaning constructed by predictions contains actions: hence, the predictions do not trigger actions, but they result from them. The perception or experience of the emotion is constructed from the past experience of that emotion, that serves to categorize the actual sensory one, present in the “now” moment (called instance of emotion). The representation of that category is a concept which, on a brain level, corresponds to a group of distributed patterns of activity. Moreover, the mechanism used by the brain to construct emotional percepts is the same for all other percepts. The brain implements predictions by carrying the related signals via feedback connections that originate in agranular cortical areas (also called limbic cortices or visceromotor regions) that relay them to the internal milieu (Barrett, 2017). Consistently, these areas are responsible for allostasis, they have an emotional function, and they also drive the concepts that constitute the brain's internal model. Then, predictions signals flow from the deep layers of limbic cortices to more developed cortical regions, with a granular structure. In this way, simulations change the spiking of primary sensory and motor neurons, even if the external sensory input has not yet arrived. To sum up, all actions and perceptions are constructed through concepts, which contribute to allostasis and reflect changes in affect. For example, when the brain tries to make sense of an instance of an emotion (i.e. pain), it constructs a set of simulations (potential actions and perceptions), that have a similarity to the current emotional situation, and then it chooses the one whose meaning better reflects that specific context (Barrett, 2017). Simultaneously, there is an inverse movement that accounts for the prediction error: it falls down in a feedforward cortical sweep (as it has also been stated by Alilovic et al., 2021), originating in the more developed and granular layers of the cortex and concluding in the deeper ones with an agranular and less developed structure. They are needed to update the internal model of the stimulus, correcting actions and representations related to it. The information contained in the prediction error is not emotional per se, but it reports the uncertainty of the prediction (Whalen, 1998), helping to fine-tune allostasis. As it cascades, this information is reduced and condensed (Finlay & Uchiyama, 2015), so that the brain can represent a huge amount of data within a smaller neural population, thus reducing redundancy (as it has also been noted by Friston, 2005). This process enhances efficiency, which is also gained because conceptually similar representations reuse the same neural populations (Rigotti, Barak, Warden, Wang, Daw, Miller, & Fusi, 2013). To conclude, new learning represents concept learning, since the brain is condensing redundant spiking patterns into multimodal summaries (Barrett, 2017). These latter can be reused as predictions, which become more detailed and nuanced as they flow to more complex regions, and they represent embodied concepts.

One of the intrinsic networks of the brain, namely the default mode network, has been proposed to be necessary to build the brain's internal model, whose simulations cascade to create concepts which categorize and guide actions (Barrett, 2017). It has been suggested that these simulations represent fully embodied brain states or embodied representations of concepts. When these summaries cascade through various brain areas, they progressively become more detailed, modulating the firing patterns of neural activity. In addition, the salience network fine-tunes the internal model by choosing which prediction errors are relevant to allostasis, namely precision signals (Feldman & Friston, 2010). They indicate the degree of confidence in the predictions, called priors (Barrett, 2017). Furthermore, being a prediction producer, the brain develops simulations in the form of concepts, across different time dimensions, processing not individual stimuli but events across temporal windows. For example, it can construct patterns that last for a long period of time, but it is also able to learn quickly. In conclusion, emotion perception is event perception, not merely object perception.

The theory of constructed emotions is a conceptualization that integrates valid and emerging biological views of the nervous system, proposing an innovative approach to understand the brain basis of emotions (Barrett, 2017). According to this theory, and in line with the predictive coding framework (Friston & Kiebel, 2009), the brain constructs an internal model based on past experiences and operationalized by concepts, in function of allostasis (Barrett, 2017). A concept is an embodied brain state that predicts the meaning of incoming sensory information in our world, the most adequate action to implement in response to it, and the consequences for allostasis, that are consciously experienced as an affect. When the prediction error is minimized, the prediction emerges as the subjective experience of that specific emotion. In an emotional context, the prediction explains its cause, categorizes it, and drives actions for dealing with it. Only once the internal model has been constructed in the form of an emotion concept, this results in an instance of emotion.

Furthermore, Barrett (2017) has described some of the assumptions of constructionism that are relevant to the field of emotions. First, emotion categories are not associated to definite, modular, and dedicated neural correlates. They are real but, as conceptual categories, they require a human perceiver. There are no dedicated neurons for specific emotions, and one neuron can take on multiple functions. In addition, in the discovery of the neural basis of psychological categories, it has been important to take a network-perspective rather than focusing on single neurons. Moreover, networks work through degeneracy, which means that they have the possibility to take over different functions, constrained by their structural form. Furthermore, an instance of emotion is a brain state that makes sense of the sensory reality and mobilizes patterns generators to implement actions in response to the current interoceptive state. Pattern generators are group of neurons that implement the sequence of actions for coordinated behaviours, for example feeding (Barrett, 2017). The default mode and salient networks, as domain-general, represent different psychological categories in the form of multimodal summaries, from which a cascade of predictions develops to become progressively detailed representations. Eventually, the entire cascade represents the instance of a concept (i.e. emotion). Relevant to our study, Barrett (2017) has also stated that individuals construct concepts to categorize sensory information and guide actions in an unconscious and automatic way, without awareness. Another important aspect for the present research is that questions about the nature or the phenomenal aspects of emotions are perceiver-dependent, so they are intrinsically human. One last implication of the constructivist theory of emotions is that the actions are not considered equal to the emotions related to them (i.e. mental inference fallacy).

To conclude, it can be proposed that, in order to unravel the nature of emotions, it is not sufficient to look at the physical manifestation of it, like facial movements and changes in the autonomic nervous system responses (Barrett, 2017). On the other hand, it is necessary to identify the brain bases of an embodied process, which consists in the making sense of physical changes in the body and in the world, and results in an instance of emotion. The theory of constructed emotions posits that these latter must be viewed in an embodied way: mind and body cannot be separated when an individual experiences an emotion. Furthermore, the context in which they develop has to be taken into account. Emotions are considered as constructions of the world and not simplistically reactions to it: they are represented by the dynamics of various networks in the brain, which are the computational heart of the internal working model of the body in the world. Both central and peripheral nervous systems are involved in the constructions of this model, that is a multi-sensory representation of reality in the service of allostasis and interoception. Allostasis is the process of predictively regulating the internal environment of the body, and interoception refers to the internal milieu representation. The process of constructing emotions is a recurrent and bidirectional one: concepts (i.e. prediction signals) fine-tune prediction errors and vice-versa, and the physical (e.g. pain) and the phenomenal facets of emotions are integrated into each other.

## 2.2. INTERSUBJECTIVITY AND ITS COMPLEX EXPRESSION: EMPATHY

Several fields of knowledge, from philosophy to psychology and neuroscience, agree on the notion that the human being is intrinsically social**:** it is hard to contemplate the development of an individual beyond a relational network (Ammaniti & Gallese, 2014). According to the transactional model (Sameroff, 1975), intersubjectivity is defined as the capacity to identify ourselves and the other as subjects in an interaction and, consequently, to share reciprocally subjective inner states. It is impossible to understand intersubjectivity without the study of the cognitive and affective bases of this process, which means the conditions that make the relationship between human beings possible. For instance, it is necessary to consider both the individual and the other person, which are part of a system: the interaction (Ammaniti & Gallese, 2014). Each subject of this interaction contributes actively to it and has to be examined with his/her own characteristics, including the information conveyed by the face and the body (also called "social cues"; Ward, 2017). Different models have been proposed to explain how we process faces and facial expressions, to investigate how we extract social and ethnical groups information, or even to understand how we make inferences on personality traits ("first impressions"). In addition, it is known that the gaze is an important social cue to enable intersubjectivity and relationships: it is considered a window on the other person, and a relationship-modulator (Ammaniti & Gallese, 2014). Furthermore, it can be noted that, since human beings are mammals, they are conceived through a relationship, they are born from and develop inside another person's body. Our brain maps the relationship with the other individual from a period which precedes birth: even the twins' fetal movement seems to be directed from a twin to the other one, as it has been demonstrated in an important study by Castiello et al. (2010), titled "Wired to be social". A recent theoretical framework that has emerged from psychoanalysis is Infant Research, that has been founded by Daniel Stern (Tambelli, 2017). It does not consider the child as passive but as an active subject in the construction of the relationship and the interaction, especially with the caregiver. One of the most important exponents of this current of framework has been Colyn Trevarthen, who has described two forms of intersubjectivity: primary and secondary (Trevarthen, & Aitken, 2001). The transition from primary to secondary intersubjectivity usually occurs between the $9^{th}$ and the $18^{th}$ month of the child development, enabling him/her to recognize the other person as different from himself/herself and to communicate with the external world, including other people or objects in the dyadic interaction (as in joint attention; Bruner, 1995).

There are various higher-level mechanisms that have been considered as mediators of intersubjectivity as, for example, empathetic processes (Ammaniti & Gallese, 2014). Empathy is conceptualized as the ability to understand others' mental states and to relate to the emotional experience of the other person (Ward, 2017). It can be experienced towards a person or a situation that are physically present or that are imagined ("Theory of mind"; Baron-Cohen, 1999). It is a fundamental aspect of the social functioning: in fact, being able to empathize with another person's affective states allows us to anticipate and understand the other's emotions, motivations and behaviours and this, in turn, facilitates the construction of affective relationships, leading to social aggregation and to intersubjective solidarity (Ammaniti & Gallese, 2014). The term "empathy" has been first used in 1907 by the psychologist Titchener to translate the German word "Einfühlung", introduced in 1903 by the philosopher Lipps to refer to the projection of emotions in observed artworks (Stueber, 2008). According to Titchener (1907), empathy can be viewed as the capacity to put yourself in someone else's shoes, in order to understand the other person. Currently this concept can be defined as the ability to comprehend and share others' affective states. It is considered a multidimensional construct, constituted by a cognitive and an affective component (Ward, 2017).

Research in psychology has suggested that there are different lower-level mechanisms which are fundamental for the development of other higher-level ones. In this specific context, it is known that one of the processes at the basis of empathy is the recognition of the emotion of the other person (Ward, 2017). One of the main theoretical models that has been considered relevant to explain how we identify others' emotions is the simulation theory. It is represented by a collection of theories proposed by various authors (Gallese, 2001; Goldman, 2006; Clark & Kiverstein, 2008; Preston & de Waal, 2002): they have suggested that, in order to recognize a facial expression and thus attributing a mental state to it, it is necessary to simulate or reproduce in first person the other's mental state**.** Hence, there has been proposed a close relationship between production (i.e., first-hand experience of an emotion) and recognition. Two types of simulation have been hypothesized: sensorimotor and embodied (also called neural resonance;

Ward, 2017). The first one refers to the notion that observing a certain facial expression leads to the subtle contraction of the same facial muscles that are used to produce it in first person, thus simulating and identifying the facial expression of the other person (Dimberg, Thunberg, & Elmehed, 2000). Using an experimental manipulation tagging facial mimicry, Niedenthal (2007) has noticed that even the facial and physical position we assume influences the emotion we experience in that moment. According with these results, a study has demonstrated that the facial mimicry activation is relevant also for the recognition of the other's emotion (Wood, Lupyan, Sherring & Niedenthal, 2016): when the facial mimicry of the observer is blocked with a hardening gel, this prevents the recognition of the facial emotion expression of the other person. The second hypothesis on simulation regards the fact that the same brain areas are activated during the first-hand experience of a certain emotion and its observation in another person. In the experiment conducted by Wicker, Keysers, Plailly, Royet, Gallese, and Rizzolatti (2003) both the observed and the first-person experienced emotion of disgust have elicited the activation of the left anterior insula (LAI) and, in part, of the inferior frontal gyrus (IFG) and the anterior cingulate cortex (ACC). In this case, the term "neural resonance" refers to the fact that, in order to understand the other person's affective state, one has to resonate with that emotion. These two aspects of simulation have been integrated in a more comprehensive theory, that has been proposed by Wood, Rychlowska, Korb and Niedenthal (2016). They have suggested that, at the presentation of an emotional facial expression, there is the activation of two parallel systems: the visual system (also called "*core system*"), referring in particular to the Superior Temporal Sulcus (STS), that is responsible for the visual analysis or detection of the facial expression (Haxby & Gobbini, 2011); the sensorimotor simulation system, which can involve the facial mimicry component, that is the activation of the motor programs for that specific emotion. Subsequently, in a cascade manner, other brain areas have been postulated to be active, like the ones involved in emotional processing (i.e., limbic system: insula, amygdala and the reward system) and in the ability to reason on others' mental states. This is in line with the actual view of the brain, that considers it as a complex aggregation of networks, abandoning the anachronistic focus on the single neuron or brain region (Northoff and Lamme, 2020; Barrett, 2017). Moreover, and in accordance with Barrett (2017), the recognition of the other's emotion derives from the attribution of a meaning to the stimulus that has been observed, so by inferring an emotion from the face stimulus (Bayesian theory). In fact, this process is also influenced by our past experience on the emotional context. This is in agreement with psychoanalytic theories that state that the understanding and mirroring of the other' experience depends, in part, on the use of our personal repertoire (Semi, 1985). Finally, the comprehensive model postulates that there is an iterative communication between the visual and the sensorimotor systems, such that a better sensorimotor simulation can refine the visual representation of the observed emotional expression in the visual working memory (Wood et al., 2016). This aspect has been studied by Sessa, Schiano Lomoriello, & Luria (2018): they have found that blocking the facial mimicry of the participant during the observation of an emotional expression worsens its representation quality, especially in the visual working memory, as measured by an electrophysiological index called Sustained Posterior Contralateral Negativity (SPCN).

Relevant to the present study, there have been found some individual characteristics that modulate the tendency to use the facial mimicry in the recognition of the other's facial expression of emotions. For instance, some of them are gender, age, empathic tendency (as we have measured through the Questionnaire of Cognitive and Affective Empathy, QCAE; Reniers, Corcoran, Drake, Shryane, & Völlm, 2011), power status, attachment style, and social anxiety. In fact, in the study of Sessa et al. (2018) more empathetic participants showed an enhanced negative impact of blocking their mimicry on the construction of mental representations of emotional expressions. Moreover, the relationship between the observed and the perceiver is important too: in-group members and liked others elicit more facial mimicry than do out-group members or disliked others (Seibt, Mühlberger, Likowski, & Weyers, 2015). In addition, there are conscious and unconscious correlates of emotion processing of faces, and the latter may be conveniently measured by skin conductance response (SCR; Ward, 2017), as it will be investigated in future studies.

The most relevant theories on empathy have suggested that it is a multi-componential construct, that includes at least two comprehensive processes (Zaki & Ochsner, 2012; Baron-Cohen & Wheelwright, 2004; Shamay-Tsoory, Aharon-Peretz, & Perry, 2009): affective and cognitive. The most recent model has been proposed by Zaki and Ochsner (2012), that have indicated the existence of two empathy components. The first one is called "experience sharing", and it refers to the affective nuance of the construct: it includes the automatic and more immediate mechanisms of simulation/neural resonance and the shared self-other representations. In fact, in order to understand what the other person is feeling, it is necessary to resonate with (or simulate) that specific affect in ourselves. This type of empathy develops earlier than the other component, and it is somewhat similar to the concept of "emotional contagion" (Zaki & Ochsner, 2012). This latter is often experienced by children when they observe a peer that is showing a certain emotion, for example pain: they seem to feel and behave in the same way (i.e., they cry when they see another child crying), because they have not developed differentiated representations of themselves and the other person, that also include the affects. The second one is the cognitive component or “mentalizing”: it includes higher level and more controlled aspects of empathy and emerges later in development. It encompasses the ability to infer the other’ mental state by reasoning on it, so the capacity to represent the other person and his/her affects and intentions as separated by their own, and the ability to take the other’ perspective. Finally, the model includes a third concept associated to altruism, which is in some way separated but related to empathy: the “prosocial concern”, that leads to help behaviours (Zaki & Ochsner, 2012). Another aspect that has to be considered when defining empathy is the difference between this construct and sympathy (Ward, 2017): in the first case the observer and the observed person experience a similar condition, perhaps of different intensity; in the second case the observer perceives compassion or preoccupation towards the other’ state. Hence, sympathy is other-oriented, and the state of the observer does not necessarily match the one of the observed person, whereas certain forms of empathy are self-oriented, like personal distress. This latter is similar to the less developed and more automatic way to empathize with someone, namely emotional contagion, and it includes a state of anxiety. Another model that has to be taken into account is the one by Decety and Jackson (2004), which is similar to the theory of Zaky and Ochnser (2012), but it includes an additional component. They have proposed the existence of three aspects in the investigation of empathetic processes: experience sharing, also called affective empathy or neural resonance; top-down control mechanisms, that are sustained by the ventromedial and orbital prefrontal cortex and are somewhat comparable to cognitive empathy. These latter processes can amplify or diminish the affective component activation; thus, they have a modulation function. The last component is called self-other discrimination, which is associated to the right temporo-parietal junction (rTPJ) and regards the ability to represent and distinguish our own and the others’ mental states, like the emotions experienced. This capacity emerges later in development and enables the individual to act in order to help the person who is in a certain affective state, like pain.

The self-other discrimination component is relevant to our study, since only when all these components are developed, the person can fully empathize (Decety & Jackson, 2004). A mature sense of self and self-consciousness are fundamental for complete empathic responses because it enables the observer to go beyond affect sharing, evaluate the other’s emotional state in relation to oneself, and involve in inter-subjective interactions (Geangu, Benga, Stahl, & Striano, 2011). Starting from this assumption, it seems clear that the capacity to differentiate between our own emotions and the other person ones is really important for the intersubjective ability of the caregiver too. Mary Ainsworth (1969a), who has developed the experimental paradigm called "Strange Situation" to assess the different infants' attachment patterns, had already discussed this point. In fact, she has defined the three components of the mother's behaviour that represent the prerequisites to accurately understand the child's needs and communications: awareness of them, freedom from distortion and empathy (Ainsworth, 1969). It can be noted that the present thesis focuses on the investigation of how these three aspects can interact, influence each other, and integrate with each other. In this case and in accordance with Zaki and Ochsner (2012), the freedom from distortion component can be well complemented with the notion of empathy. The first one indicates that the mother has to distinguish her child mental and affective states from her own, and more broadly that she has to be able to separate her identity from the one of her infant, in order to mirror the authenticity of the child Self (Winnicott, 1967). Hence, this capacity influences the caregiver' ability to fully empathize with her child, that means "to share and to explicitly comprehend the other' inner states that in turn enables the individual to respond to the others' needs" (Zaki & Ochsner, 2012). Since the mother-infant dyad is conceived as an interactive and reciprocal system (Sameroff, 2009), the intersubjective capacity of the caregiver, by differentiating and reflecting on the child states, is moved to and interiorized by the infant, who will develop its own ability to understand and interact with the other person. More specifically, it is known that the child learns how to identify, understand and name his/her/others' internal states through the mirroring ability of a "good enough mother", which is a specific nuance of intersubjectivity (Winnicott, 1967). In this context, the intersubjective ability can be broadly conceived as the capacity to understand and regulate another person's internal states, like the emotions. This is carried out by confirming and then mirroring the child's emotion through a special language (the "motherese"), in order to show him/her that the affect experienced is real. Then, the caregiver has to gather the child's rough emotion in order to transform it in a clear, nameable, containable and shareable affect. Wilfred Bion, another remarkable psychoanalyst, has expressed this concept with the term "reverie" (1962). The so called "Beta elements" of the child (i.e., primitive, raw, sensorial, and emotional experiences) can be transformed by the caregiver in "Alpha elements", that represent a psychic experience deriving from a further elaboration and symbolization of the first elements.

To summarize this part, it can be proposed a comparison between the cognitive/neuroscientific approach and the psychodynamic/psychoanalytic account, noting that there is a similarity between the simulation system and the mirroring ability, that can all be conceptualized under the broader umbrella of intersubjectivity (Gallese, 2011). On a developmental level, a "good enough" caregiver should have matured her/his own intersubjective capacity that, in turn, is fundamental to structure the child one. So, for example, when the child falls down and gets hurt, the caregiver comes and, through her/his language helps her/him to understand and name the emotion he/she is experiencing, like pain. In this way, the child progressively achieves conscious affective experience through others' language: in other words, he/she is facilitated in conscious and explicit emotion recognition through his/her caregiver intersubjective ability. The intersubjective capacity can be conceived as being composed by two main nuances: the cognitive one, referred to as mentalization (i.e., the ability to understand one's own and others' mental states, thereby comprehending one's own and others' intentions and affects; Baron-Cohen, 1999); the affective one is represented by empathy, which is relevant to the present study.

Most of the studies on empathy have been focused on a particular type of this construct (Ward, 2017), that is empathy for others' pain, which is important for our research. This preference can be attributed to the fact that the pain matrix (i.e., the brain areas activated when we experience first-person pain) is well-known: it has been easier for the researchers that used functional Magnetic Resonance Imaging (fMRI) to evaluate a possible overlapping between the regions involved in the first-hand experience of pain and the ones engaged in the observation of another person in pain. This would be in line with the discovery of a neural resonance or affective empathy component (Zaky & Ochnser, 2012; Decety & Jackson, 2004): it has been hypothesized that when we empathize with another person's pain we initially and automatically share the other's emotional experience, as there was an affective contagion between the observer and the observed person. In addition, it has been highlighted that the functional dissociation is reflected in a similar anatomical dissociation between the affective and the cognitive components of empathy (Zaki and Ochsner, 2012). The first one is underpinned by the anterior cingulate cortex (ACC), the anterior insula (AI), the premotor cortex (PMC), the inferior frontal gyrus (IFG, which contains mirror neurons) and the inferior parietal lobule; the second one is supported by the medial prefrontal cortex (MPFC), the precuneus, the temporal poles (TP) and the temporo-parietal junction (TPJ). The affective empathy component, also called neural resonance or experience sharing, has been hypothesized to include the mirror system and, in particular, mirror neurons, that do not refer only to the sensorimotor aspect of the simulation process (Iacoboni, Woods, Brass, Bekkering, Mazziotta, & Rizzolatti, 1999; Gallese & Goldman, 1998; Ramachandran, 2000). It is important to state that there is no direct proof that mirror neurons are involved in empathic processes, but it is legitimate to hypothesize that mirroring mechanisms are engaged in this intersubjective ability. In fact, it has been observed that a cardinal component of empathy is represented by mimicry processes that, in turn, are hypothesized to be at least partially a reflection of the mirror neurons activation. In addition, the mirroring system is not involved only in sensorimotor but also in embodied simulation, as it has been seen in the experiment about first-person and observed disgust (Wicker et al., 2003). Moreover, a similar neural resonance mechanism, that is not associated with motor aspects, has been highlighted by Tania Singer and colleagues in two studies (2004; 2006). In this context, the embodied simulation process has been reconnected to the anterior cingulate cortex (ACC) and to the anterior insula (AI).

The first study, published in Science, has been conducted by Singer, Seymour, O'doherty, Kaube, Dolan, & Frith (2004). The authors have demonstrated that both in first-hand and in observed other's pain experience, the ACC and the AI were activated. As it has been seen before, these areas are connected to the neural resonance component of empathy. Moreover, this effect has been noted to be wider when the person who experienced the painful stimulation (i.e. an electrical shock) was the fiancée of the participant, compared to the situation in which the observed other was a stranger. This latter aspect has confirmed the notion that empathy is influenced by relational components between the perceiver and the person observed. Finally, there has been noticed a positive correlation between the level of AI activation (and, in part, also of the ACC) and the score at the Interpersonal Reactivity Index scale called "empathic concern" (Davis, 1980). This instrument refers to affective empathy: hence, the hypothesis that AI and ACC represent the foundation of neural resonance has been supported more significantly. In the second research, published on Nature, Singer, Seymour, O'Doherty, Stephan, Dolan, and Frith (2006) have further investigated how empathy can be modulated by the affective relationship with the other person, including the one based on trust. The participants were engaged in playing the Prisoner's Dilemma game with an experimenter confederate that could act either in a honest or in a distrustful way. Subsequently, the subjects have undergone an fMRI session where they have either received a painful stimulation or observed the confederate in pain. The results were divided by gender: women showed a higher affective empathy towards both types of confederates in terms of ACC and AI activation compared to males, even though the neural response has proved to be wider for the honest confederate. On the other hand, male participants showed a totally suppressed activation for the dishonest other. Indeed, the nucleus accumbens, which refers to the reward system, has resulted to be activated: this suggests a sort of satisfaction that derives from painful punishment to the untrustful confederate. Regarding the cognitive component of empathy, this has been studied by Cheng, Lin, Liu, Hsu, Lim, Hung and Decety (2007). The research, published on Current Biology, has investigated the top-down control nuance of cognitive empathy. The hypothesis that has been supported refers to the fact that shared neural representations (connected to the more automatic and affective empathic processes) can also induce personal distress, that we have observed to derive from a less developed relational and emotional capacity. In fact, affective empathy does not require the capacity to distinguish the Self from the Other. It has proposed that regulative top-down mechanisms may be involved in inflicting painful procedures to others in expert acupuncturists when compared to non-experts. In fact, their medical caring ability could have been subjected to an interference by the more immediate empathic distress, if no other processes were engaged to modulate this latter. More specifically, this was evident when looking at the fMRI results: in the expert participants medial and superior prefrontal cortex were activated when observing another person in pain, whereas in non-experts ACC and AI were engaged. The first regions support the cognitive component, whilst the latter ones are involved in affective empathy.

Departing from this theoretical background, it is possible to move forward toward our research hypotheses. In the field of empathy study, we have proposed to disentangle the relationship between empathy for others' pain and consciousness, referring in particular to the conscious perception of a facial expression of pain. The aim is to understand whether there is an association between the two processes and, if so, in what direction. It could be that a greater conscious perception of the other' emotions leads to a higher tendency to empathize with the observed person. In this case, it can be hypothesized that the multisensorial integration of different sources of stimuli (contextual, for instance a read sentence; visual, like a facial expression; or even imaginative, for example representing internally a certain situation experienced in the past) could facilitate the conscious perception of the emotional stimulus and, in turn, the empathic processes activation. It can be suggested that, in this context, empathy would be involved in terms of its cognitive nuance, since it would result from a higher stimulus awareness and cross-sensory integration, which require a further elaboration of the emotion observed. As has been stated by Barrett (2017), multimodal summaries deriving from concept learning, can be used as predictions and they represent embodied concepts. Another hypothesis would be that a higher tendency to be empathic leads to a greater emotional consciousness of the facial expression. This latter proposal would be in line with the predictive coding models, since the predisposition to empathize (as measured with the QCAE; Reniers et al., 2011) could be considered as part of our expectancies about the world, that are based on our structured representations about the Self, the others, the relationships, and the reality itself. In fact, according to developmental and dynamic psychology, these stable representations (also called Internal Working Models- IWM; Bowlby, 1969) are formed in the infant from the internalization of repeated interactions with the caregiver. Ultimately, we base our relational modalities (including the tendency to empathize) and expectancies on these IWM. Reconnecting to our hypothesis, it is possible that the tendency to be more empathetic could serve as a prediction for the emotional face conscious perception. As it has been seen before (Barrett, 2017), predictions are embodied simulations of sensory-motor experience, and they are constructed in the service of allostasis. To conclude, it could be introduced a particular nuance of our reference theory, that may be referred to as "interpersonal predictive coding". In fact, it is known that empathy arises from shared self-other representations, that could constitute the expectations or simulations in this theoretical framework. In further support of this statement, Gallese and Goldman (1998) have considered empathy as a special case of mental simulation, that involves the usage of our own psychological mechanisms as a model for the other person ones. Hence, it has been proposed that we are able to use our own resources to construct a model of another person and, thereby, to identify with him/her, projecting ourselves imaginatively into his/her situation. In my opinion, it is also possible that both mechanisms come into action, through a bi-directional influence between empathy and consciousness.

To date, it can be noted that there is no complete or specific literature about our hypotheses that, therefore, constitutes a novelty in neuroscience and psychology research. Nonetheless, it is possible to base our study on some related hints provided by other authors and research. First, Evan Thompson (2001) goes even further the simulation theory (ST) conceptualization, presenting five points that support the hypothesis that empathy represents the foundation for consciousness: human consciousness is formed in the dynamic interrelation of self and other, and therefore is inherently intersubjective; the concrete encounter of self and other fundamentally involves empathy, understood as a unique and irreducible kind of intentionality; empathy is the precondition (the condition of possibility) of the science of consciousness; human empathy is inherently developmental; real progress in the understanding of intersubjectivity requires the integration of methods and findings of cognitive science and phenomenology. Thompson (2001) has conceived the mind as characterized by three main aspects, which link it to empathy and intersubjectivity: embodiment, emergence, and self-other co-determination. This means that the mind or cognition is embodied in the whole organism that lives in its environment, that it emerges in the interrelation of top-down and bottom-up processes, and from the dynamic co-determination of self and other. Regarding this last point, it is known that self-consciousness develops from an embryonic and preverbal sense of self that is already present in newborn infants. It is indissolubly associated with the perceptual recognition of other human beings, evident in the phenomenon of imitation (Gallagher & Meltzoff, 1996) and it emerges through the experience of being recognized by the primary caregiver, as exemplified by mirroring (Winnicott, 1967). The hypothesis of self-other co-determination is connected to the re-discovery of the importance of affect and emotion in cognition (Damasio, 1994). The affect can be defined as a prototypical two-organism or self–other event (Thompson, 2001). Another important aspect to understand the interrelation between empathy and consciousness regards a common concept between phenomenology and affective neuroscience (especially mirror neuron system theory), called non-inferential bodily pairing (Gallese, 1998; Husserl, 1973). It refers to the fact that in order to recognize others' mental states, it is not necessary to implement an explicit and cognitive evaluation, but it is sufficient to establish an immediate pairing or matching of bodies and actions of Self and Other. This could imply that empathy is at the basis of consciousness because it is founded on an automatic bodily recognition of the other that precedes a conscious evaluation of mental states. In line with this hypothesis, it can be observed that intersubjectivity and the relationship with the significant Other comes before and modulates the Self birth and, hence, the child's own inner states consciousness and the possibility to reflect on them. Another relevant affective neuroscience concept concerns the proposal that gesture recognition mirror neuron system is at the basis of the language development, literally "the neural prerequisite for the development of interindividual communication and finally of speech" (Rizzolatti & Arbib, 1998, p. 190). This aspect is meaningful for our study because it could imply that the linguistic context intervenes subsequently to the automatic emotion recognition and empathic processes, hence both the context and immediate empathy may be integrated to facilitate conscious affective perception. Departing from these assumptions, Thompson (2001) goes even further the ST, that begins from the individual and then projects to the other through mimicry and imaginative projection. On the other hand, according to phenomenology, the Self is already 'intersubjectively open' in its very structure for these mechanisms to function effectively. In fact, mimicry and the imaginative transposition of oneself to the place of the other are certainly empathy components, but they are based on more fundamental pre-reflective matchings of Self and Other at the level of the lived body. In this sense, the pre-reflective and not-voluntary initiated experience of the Other as an embodied being like oneself lays the foundations for mimicry and the more complex mental act of imaginative self-transposal. Another point of questioning raised by Thompson (2001) regards the conceptualization of the Self and the Other, which are not viewed as an observer and an observed person since the intersubjective relationship nature is mutually defining (Gallagher, 1997).

To understand better the relationship between empathy and consciousness, Thompson (2001) has reviewed the phenomenological point of view on the question. According to these authors (Husserl, 1973; Zahavi, 1997), consciousness is intrinsically ‘intersubjectively open’, that is, it is structurally open to the Other prior of any real, concrete meeting of Self and Other (this concept seems to be in line with the predisposition for consciousness proposed by Northoff and Lamme, 2020). Furthermore, one’s consciousness of oneself as an embodied individual placed in the world depends on empathy, in particular on one’s empathic understanding of the Other’s empathic grasp of oneself (this aspect is consistent with dynamic developmental psychology research, that considers the formation of the self in primary relationships; Winnicott, 1967). Regarding the first point, Zahavi (1997) has stated that our personal experience or perception of another subject as embodied is founded upon an a priori reference to the Other, which means that the open intersubjectivity constitutes an *a priori* structure to perceive the other as an embodied subject. Nonetheless, it is not the only factor involved in our embodied experience of other subjects: *concrete* intersubjectivity is also necessary, that is the concrete bodily experience of the other person. One aspect of this latter construct is empathy (Thompson, 2001) which, as an imaginative self-transposal, presupposes the open intersubjectivity of consciousness. This means that empathy allows us to acquire a new spatial perspective on the world, that of the Other. Simultaneously, we can maintain our own centre of spatial orientation. Hence, the open intersubjectivity of consciousness and its actual expression in empathy are the pre-requisites for our experience of living in a common, intersubjective, spatial world. It is through empathy as the experience of oneself as an Other for the Alter-Ego that one gains a perspective of oneself as embodied beyond the first-person singular point of view. This point has been elaborated by Stein (1964) as “reiterated empathy”, that means that I see myself from your (empathetic) perspective of me. In this way, one’s sense of self-identity, even at the most basic layer of embodied agent, is embedded with the recognition by another person, and from the capacity to understand that recognition empathetically (this concept is again in line with the one of “mirroring” between the infant and the caregiver, as expressed by Winnicott, 1967). In conclusion, the link between the open intersubjectivity of consciousness and empathy can be conceptualized by the fact that the Self is structured in many parts that represent various inner openings to the others, hence consciousness cannot be separated from the confrontation with the Alterity (Thompson, 2001). This link is called by Thompson (2001) “core dyad”, and it has two sides. The first one is phenomenological, and it introduces the term “reciprocal empathy”, that is considered the condition of possibility for consciousness, viewing the Self and the Other as concretely co-determined. According to this view, the embodied mind, which is the object of study of cognitive science, can be defined through reiterated/reciprocal empathy and on the basis of the lived body (Merleau-Ponty, 1962). The second side is represented by cognitive science, in which empathy is considered as an evolved, biological capacity of the human and other mammalian species. Thompson (2001) has suggested the necessity to integrate these two forms of understanding, namely phenomenology and cognitive science, since the lived body and the embodied mind cannot be separated, rather they constitute two sides of one single spatio-temporal individual (recalling Northoff and Lamme, 2020). In fact, the lived body is the precondition of our capacity to know and perceive reality, and especially of our ability to know anything about the embodied mind as an object of biological and cognitive scientific study. On the other hand, abandoning biology and cognitive science would imply the impossibility to understand what stands before the living body, that is the organism in the broad life complexity (Thompson, 2001).

Gallese (2011), who is one of the most representative authors of the simulation theory, the mirror system theory and intersubjectivity in general, has supported the integration between different areas of research, specifically cognitive neuroscience and psychoanalysis, in the context of embodied simulation (ES) theory and departing from a paper by Hustvedt (2011). As also Thompson (2001) has noticed, Gallese has observed that the bodily affective Self (which is primordial, proprioceptive and preverbal) is at the roots of the narrative Self (which is more developed, conscious and verbal). In addition, the author has shown that, when we read or listen to narratives, we embody them by activating our sensorimotor system, enabling the so called “Feeling of Body”, which is the physical experience of the mind (Gallese & Wojciehowski, 2011). This aspect is relevant for our study since one of the independent variables is the linguistic context: the participant has to read emotional sentences that we propose could contribute to empathic processes and consciousness through embodiment. This last consideration becomes important for the link between neuroscience and psychoanalysis since communication and non-verbal language are a significant part of the relationship between the patient and the therapist: the affective and corporeal quality of communication is prominent in transfert processes activation and, in general, for the possibility to understand and, thus, empathize with the patient (Mancia, 2006). In order to support the proposal of a bi-directional influence between empathy and consciousness, it can be noticed that according to Gallese (2011), the embodied simulation process is triggered by a perception, and it can also occur when we imagine doing or perceiving something. This has two implications: it could be that also conscious perception (including multisensorial one) could enhance empathy and that this could happen also if the situation is only imagined, like when reading an emotional sentence describing a person in pain. In fact, both visual and motor mental imagery involve the activation of sensorimotor brain regions, hence they can be qualified as further forms of ES. Following Gallese (2011): "at the core of our perception, understanding, and imagination is the body". According to the author, both being immersed into fictional narratives and empathizing with someone represent intermediate worlds, because they evoke a Feeling of Body by identification and mirroring mechanisms, and because they recall our past memories and experiences. It is possible to draw a parallel between intermediate worlds and the transitional space theorized by Winnicott (1967): empathy and ES can be viewed as having a transitional value, since they are found between me-not me, reality-fantasy (Beesley & Tyrell, 2012). In fact, when I empathize with someone, I partially identify with the situation lived by the other person, hence reconnecting to my personal repertoire and imagining something that is not present at the actual moment but that has happened to me in the past or could have possibly occurred. Thus, this process enables to build a bridge between the Self and the Other. Going further, it could be hypothesized that empathy facilitates consciousness or emotional conscious perception. In fact, the child, through the elaboration of the transitional object, progressively manages to shift from an unconscious, automatic and visceral emotional experience (loss sensation or separation anxiety) to a more conscious and explicitly representable, hence more tolerable, affective experience. Therefore, it can be hypothesized that empathy represents the resolution of the transitional field, enabled by a process of decathexis (Beesley & Tyrell, 2012): this concept means that the child abandons the transitional object, internalizes it as a representation that stands for the caregiver and allows the child to feel safe without the actual presence of the mother. It can be noticed that intersubjectivity enables this transition: it is from the repetitive reliable relationship with the caregiver that the child internalizes a sense of internal safety, shifting from the need for the mother to the desire of her, thus processing separation and moving to a progressively greater autonomy and individuation (Litt, 1986). In fact, Spitz and Metcalf (1978) have considered this passage as a development from the recognition to the evocative memory: the transitional object serves to evoke the mother when she is absent, in order to call her representation to consciousness. In our research case, it can be hypothesized that if empathy or ES has a transitional value, it could facilitate the pain emotional stimulus (i.e. noisy face) awareness and symbolization. To summarize, the concept of the ‘intermediate area’ of experience has been associated to the development of the capacity for empathy and of meaningful object relationships by various writers (e.g. Horton, Louy & Coppolillo, 1974).

Finally, Kagan (1982) has proposed a relevant conceptualization of the transitional object role in development. According to the author, early psychological development is characterized by a progression of cognitive abilities that mature in a regular way: recognition, retrieval, and active memory. These latter relate to universal fears of infancy and the more advanced capacities to infer causes of events (e.g. as the construction of emotions; Barrett, 2017), empathize with another person feelings, and recognize the self as distinct from others (that is an empathy component). Since children develop attachments to transitional objects while they are developing these critical cognitive skills (like conscious emotional experience and Self) it can be suggested that the two processes are associated or at least that the appearance of the attachment object enables the access to the requisite cognitive structures. Lastly, Kagan (1982) has also proposed that the obtainment of the ability to infer the causes of the events brings to a sensation of uncertainty that, in turn, leads to motivation to mental development to reduce the anxiety. This leads to the use of the transitional object, in its emotional regulation function.

## 3. THE RESEARCH

### 3.1. INTRODUCTION: THE EXPERIMENTAL HYPOTHESES

The present study can be considered as part of two broader areas of research: emotions and consciousness. As it has been mentioned before (Damasio, 1994), only recently there has been the re-discovery of the importance of affect and emotion in cognition. In fact, according to Damasio (1994), emotions are integral to the process of reasoning, since reasoning arises from the need of the organism to regulate itself in the service of allostasis, by the construction of emotions (in line with Barrett, 2017). Indeed, Damasio (1999) and Panksepp (2000) in his review have tried to disentangle how the mind can emerge from brain-body dynamics ("primordial self"), namely embodiment processes. Hence, it could be speculated that even empathy involves the regulation of the internal milieu but, in this case, specifically in function of the relationship with the other person. If emotions are constructed in the service of allostasis, it could be that also the Other's emotions are built through the same mechanisms, in order to serve a social purpose. As it has been hypothesized in the previous chapter and according to Thompson (2001), human consciousness is formed in the dynamic interrelation of Self and Other and, therefore, it is inherently intersubjective and founded on empathy. Again, lower-level processes reveal to be at the basis of higher-level ones (Ward, 2017).

Based on these premises, it is possible to illustrate our research hypotheses. The main theme of the study concerns the relationship between empathy and consciousness, and how these constructs are influenced by emotional contextual information, referring specifically to language cues. As it has been described in the first two chapters, our theoretical background lies on the Predictive Coding Theory (PCT; Friston & Kiebel, 2009), the Temporo-Spatial Theory of Consciousness (TTC; Northoff & Lamme, 2020) and the Theory of Constructed Emotions (TCE; Barrett, 2017). In general, it can be assumed that perception is the result of an interaction between bottom-up stimulus related neural activation and top-down prediction-driven biases and that predictions benefit perception when the sensory input is weak (Clark, 2013). So, it can be stated that predictions can modulate subjective perceptual report of visual stimuli and the ability to consciously access them (Alilovic et al., 2021). In addition, studies have reported that predictions modulate conscious access in a perceptual way, influencing sensory responses and sharpening neural representations in sensory brain regions. Furthermore, it can be postulated that emotions, like all other perceptions, are constructed integrating somatic feedback and external (contextual and visual) information, in the service of allostasis (Barrett, 2017). Starting from these theories, two main first hypotheses are presented in this research.

First, we have investigated whether congruent contextual information could heighten the conscious processing of emotional (painful) faces. In our case contextual information is operationalized as sentences. So, we have hypothesized that the presentation of emotional painful sentences could boost the level of awareness of facial expressions of pain. This is in line with Barrett's (2017) theory, for which emotions are constructed through predictions (activated by the contextual information) and the involvement of a large brain system, in the service of allostasis and interoception.

Second, we have examined if anticipatory somatic feedback activated by congruent contextual information is somehow related to conscious access of emotional (painful) face expressions. More specifically and moving further, we have recorded two psychophysiological indexes: electromyography (EMG) and skin conductance (SC). We have hypothesized that the EMG mimicry response will be congruent with the emotional content of the context, such that a larger activity of the corrugator should be observed for written sentences with painful content. Furthermore, we expected that EMG response to emotional context would predict conscious access to the target face. Specifically, SC responses will be analyzed in future research, so they are not reported in the present thesis. A further hypothesis is that the emotional contextual information, with respect to the neutral one, could enhance the overall empathic ratings expression, even if the painful facial expression is not completely visible and, hence, consciously accessible. This observation would be useful to understand if and how an individual is prone to understand and then help a person in a painful situation, where the dangerous context is ambiguous and not totally clear. The question at issue is relevant for different points of view and fields of knowledge. First, for the neuroscientific and cognitive one, since it would be important to understand if cortical mechanisms and conscious perception of a stimulus are necessary to activate more subcortical and emotional processing, or vice versa. Second, for the psychodynamic area of research, it would be significant to have an insight into intersubjective and unconscious defense mechanisms in the prosocial and empathic contexts. For instance, if we are less prone to help a person in a presumptive (because not completely visible) painful situation, it could be that unconscious defense mechanisms are getting involved, in order to protect ourselves from a likely dangerous context. Another hypothesis that we have elaborated regards the possible influence of a higher trait affective and cognitive empathy on the predicted effects (situational empathy, conscious perception of the facial expression, and psychophysiological indexes' magnitude) as measured by the QCAE (Reniers et al., 2011). This latter hypothesis is relevant also for the psychodynamic point of view since we are considering the importance of subjective and individual differences in psychological and neural processes. Lastly, and as has been previously described, we aim to investigate if consciousness could enhance empathy, vice versa, or if bi-directional mechanisms are involved. It could be that a greater conscious perception of the other' emotional expression leads, through multisensorial integration, to a higher tendency to empathize with the person observed. On the other hand, it is also possible that being more prone to empathize, as measured by a situational subjective empathy rating, brings to a greater emotional consciousness of the painful facial expression. As it will be explained afterwards, it is only possible to have a theoretical insight into the question, but it is statistically not possible to disentangle the direction of the association.

## 3.2. METHODS

### 3.2.1. PARTICIPANTS

The experimental subjects have been selected from a medium size sample of 48 participants, through online portals or distributed physical flyers. The contexts that have been chosen for the recruitment were mainly university ones (e.g. Facebook groups, libraries). The selection has been guided by the information provided in the advertisement, where the participant was required to own specific characteristics. We have recruited subjects regardless of the gender, between 18 and 35 years old, with good knowledge of English and a mother tongue level of German (i.e. minimum C1), and with no previous of actual psychiatric or neurologic diagnosis. This because the experiment (instructions, sentences, and questions on the computer screen) was in German, whereas the experimenter spoke in English. The final sample was composed by 39 participants, since the first 9 subjects have been excluded after an improvement of the first phase of the experiment, called calibration. The participants were 22 females (average age of 23.36 years old) and 17 males (average age of 24.64 years old). All the subjects have read and signed the consent form to take part in the experiment, in agreement with the Declaration of Helsinki.

### 3.2.2. STIMULI, TASKS, AND PROCEDURES

The experimental session has been conducted in a dedicated laboratory at the University of Vienna's Faculty of Psychology, Clinical and Social Neuroscience Unit, under the supervision of professor Giorgia Silani and Dr. Claudia Massaccesi. The study has consisted in a task carried out on a computer, and it has involved the vision and reading of stimuli (faces, sentences) and some judgments, while 2 psychophysiological signals were recorded. At the end, a questionnaire has been administered, always on the same computer. When the participant arrived, he/she has been welcomed and got comfortable, while reading and signing the consent form to participate in the present study. Then, the experimenter has explained the psychophysiological procedures that have been prepared and implemented during the first part of the experiment, namely EMG and SC, reassuring the participant that these recordings are not invasive or painful. So, first, the researcher has applied five electrodes on the participant face: one on the forehead, close to the hairline, that represents the baseline of the EMG; two on the left upper extremity of the eyebrows, close to the nose, that measure the Corrugator Supercilii (CS) muscle's activity, which is active during the facial expression of pain; two on the left cheek, that are used to record the zygomaticus major (ZM) muscle activity, which is active during the facial expression of happiness. Being an empathy for pain study, we expect the first one to be particularly operating, whereas the second one is predicted to be quite silent. Before the electrodes' application, the skin has been prepared by scrubbing the interested areas, in order to reduce an index called "impedance" (Bora & Dasgupta, 2020). This latter is defined as the response of a specific skin region to an externally applied electrical current (or voltage); it is indicated with the symbol Z and measured in Ohm (Ω; Merletti, Botter, Troiano, Merlo, & Minetto, 2009). It is a function of the skin's structure and composition, and it is a complex quantity consisting of resistance and capacitance whose values depend on the frequency of the applied voltage for the measurement. In our experiment, we have considered as valid the measurements where the impedance was lower than 20 Ω. In addition, it has been applied an electroconductive gel underneath each electrode, through a precision syringe, in order to improve the conductance of the electrodes and, thus, to obtain a better acquisition of the psychophysiological signal.

After the preparation phase, the experimenter has explained the first part of the experimental design, which is called calibration phase. The stimuli for the calibration phase are painful facial expressions, both female and male, and they have been taken from a database used in a study by Sessa, Meconi and Han (2014; Figure 3.1.).

*Figure 3.1. Painful facial expressions used as face stimuli.*

This phase concerns the calibration of the amount of Gaussian noise necessary to identify threshold stimuli, and it has been done for each subject and for each face stimulus. In particular, we have defined threshold stimuli as those that are consciously perceived/seen in 50% of the trials. We used a 1up-1down staircase procedure (Leek, 2001; García-Pérez, 2001) increasing or decreasing the gaussian noise respectively after a “seen” (PAS 1) response or “unseen” (PAS 2-4) response. The final noise level was estimated averaging all reversals excluding the first two. In the second phase, or experimental session phase, the participants have been presented with 16 emotional and 16 neutral sentences (which have acted as contextual information) and, afterwards, with the facial expression at the individual threshold previously identified. The sentences could describe a person in either a neutral (e.g.: “This person drank green Japanese tea”) or painful situation (e.g.: “This person smashed his/her finger with a hammer”), and they were in German. In addition, the participants have been told that the face belonged to the person described in the sentence. After a mask, the participants have been requested to do a perceptual rating through the Perceptual Awareness Scale (PAS; Overgaard et al., 2010; Table 3.1.): they have been asked to judge whether they have seen the facial expression or not on a scale from 1 (“I did not see the face at all”) to 4 (“I clearly saw the face”).

| Category | Description |
|---|---|
| No experience | No impression of the stimulus. All answers are seen as mere guesses. |
| Brief glimpse | A feeling that something has been shown. Not characterized by any content and this cannot be specified any further. |
| Almost clear experience | Ambiguous experience of the stimulus. Some stimulus aspects are experienced more vividly than others. A feeling of almost being certain about one's answer. |
| Clear experience | Non-ambiguous experience of the stimulus. No doubt in one's answer. |

*Table 3.1. The Perceptual Awareness Scale (PAS; Overgaard et al., 2010) implemented during the experimental session phase.*

Finally, the subjects had to express how much they have empathized with the person depicted in the facial expression (considering the situation described by the sentence) on a Likert scale from 1 (“Not at all”) to 6 (“Very much”). Both the calibration and the experimental session phase have been preceded by practice trials, to allow the participant to familiarize with the tasks.

Prior to the starting of the actual study, a preliminary short behavioural pilot has been conducted, in order to identify the most neutral and painful sentences, that means to check for the arousal and valence of the stimuli. These stimuli have been taken from the same database of the facial expressions (Sessa et al., 2014). Eventually, all the sentences have been used for the experiment because we have deleted one stimulus type from the original facial expression database (that is, we have kept only painful faces, whereas we have not included the neutral ones, in order to reduce the overall duration of the experiment). As a result, this decision has allowed us to maintain a certain variability of the responses. In fact, we have been able to maximize the effect we aimed to observe by balancing two factors: duration of the experiment and number of participants (with the aid of a Power Analysis).

Regarding other details of the experiment, the face stimuli have been scaled using an image-processing software so that each face fit in 2.9° × 3.6° (width × height) rectangle. The sentences have been presented on three lines at the center of the computer screen in a 1.73° × 3.9° (width × height) virtual rectangle from a viewing distance of approximately 70 cm. In total the participants have performed 197 trials, including the catch and the practice ones. Catch trials have been implemented as non-face stimuli, in order to facilitate the estimation of the level at which a participant is guessing when no stimulus is present (VandenBos, 2007). Each trial has begun with the presentation of a fixation cross at the center of the screen (2000 ms), followed by a sentence (4000 ms). After a blank interval (8000 ms), jittered in steps of 100 ms, a face at the individual threshold (i.e., target stimulus) has been displayed for 50 ms and masked for 300 ms (Figure 3.2.).

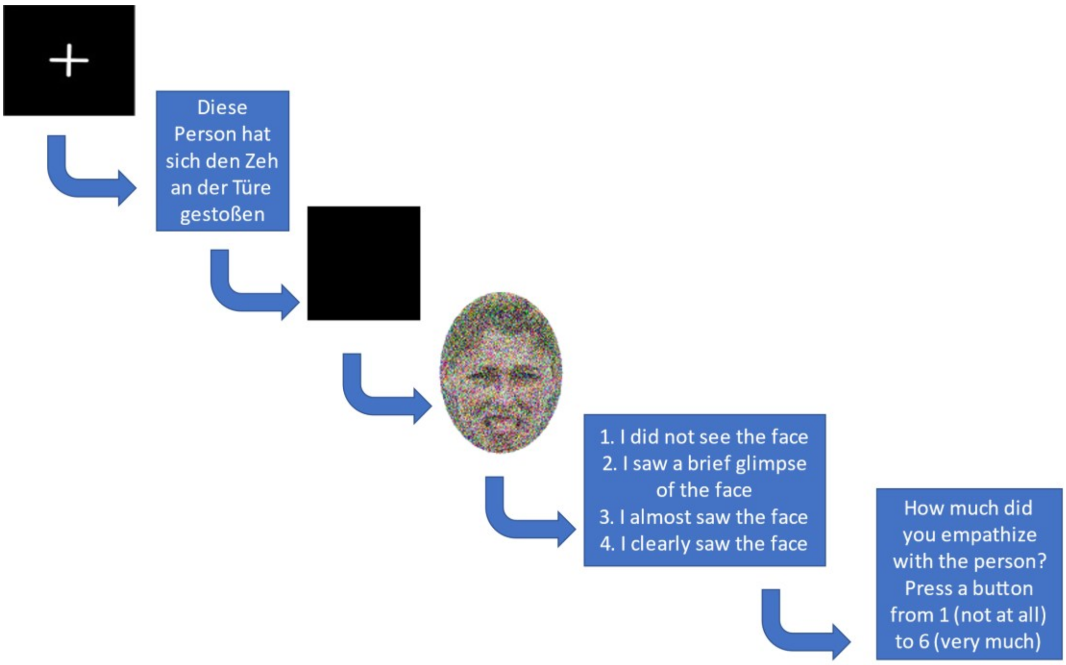


*Figure 3.2. The experimental paradigm for the second phase of the study.*

To summarize, the independent variables were 5, where 3 of them were of behavioural nature and 2 of them were psychophysiological. In particular, we have measured the visibility rating by means of the Perceptual Awareness Scale (PAS; Overgaard et al., 2010) on a 4-point scale (Table 3.1). As it has previously observed in the present thesis (Alilovic et al., 2021) the PAS is used to measure prediction effects and it emphasizes subjective perceptual aspects of awareness that might escape classical dichotomous measures or discrimination response. In fact, the scale is structured as a *gradual* perceptual report. This means that the PAS scale is less likely influenced by response biases, i.e. tendency towards reporting (or guessing) the more probable stimulus category, as opposed to a discrimination response. This latter aspect is further enhanced by the inclusion of catch trials, as it has been stated in the previous paragraph. Furthermore, the PAS has allowed us to express the effects on a more fine-grained scale (with respect to seen/unseen dichotomous responses). The other behavioural measures that have been implemented are: the empathy subjective rating on a 6-point scale as a function of the PAS; a final questionnaire, after the experimental session phase, which is called QCAE (Questionnaire of Cognitive and Affective Empathy; Reniers, et al., 2011). This last index has been used to understand if a higher score at the questionnaire is associated with higher PAS values, empathy subjective rating and psychophysiological activity (EMG). The QCAE is a 31-item self-report multi-dimensional questionnaire assessing cognitive and affective empathy on a 4-point Likert scale, and it is composed by five subscales: perspective taking, online simulation, emotion contagion, proximal responsivity, and peripheral responsivity. The first two subscales measure the cognitive nuance of empathy, whereas the last three indicate affective empathy. Cognitive empathy is defined as the ability to construct a working model of the emotional states of others, while affective empathy is referred to as the ability to be sensitive to and vicariously experience the feelings of others (Reniers et al., 2011). More specifically, the subscale “perspective taking” is defined as the experience of intuitively putting oneself in another person’s shoes in order to see things from his/her perspective; the subscale “online simulation” indicates an effortful attempt to put oneself in another person’s position by imagining what that person is feeling; the subscale “emotion contagion” is referred to as the automatic mirroring of the feelings of others; the subscale “proximal responsivity” is defined as the affective response when witnessing the mood of others in a close social context; finally, the subscale “peripheral responsivity” indicates the affective response when witnessing the mood of others in a detached social context. Ultimately, the psychophysiological independent variables were the EMG and the SC response: the first one indicates the mimicry response measured during the presentation of the sentence (see the simulation theory in the previous chapter; Gallese, 2001); the second represents an automatic activation measured during the whole trial, and it indicates the level of arousal when the participant sees an emotional stimulus.

### 3.2.3. STATISTICAL METHODS

*Behavioural data*

Behavioural data have been analyzed through the R software (R Core Team 2022). First, it has been conducted an Exploratory Data Analysis (EDA; Behrens, 1997) through bar charts to represent the frequencies associated with the variables’ modalities. The EDA is used as an initial descriptive, explorative, and graphical overview of the data and of the associations between the variables. Afterwards, a series of inference analyses have been carried out, through a regression method: the aim was to understand if an independent variable could function as a predictor for the variable whose variation we want to observe, namely the dependent one. Inference refers to a logic process for which, starting from one or more assumptions, it is possible to draw a conclusion. More specifically, it has been used a multilevel ordinal logistic regression analysis (Agresti, 1989; 1999), that is a regression model for ordered categorical dependent variables. .

We considered participants as random effects (intercept) including also random slopes when relevant. Hypothesis testing on regression parameters has been conducted using Wald z-tests using an alpha level of 0.05 and reporting 95% confidence intervals. For other analysis a linear mixed effect model or paired t-test have been used with an alpha level of 0.05.

*Psychophysiological data*

*EMG*

Physiological responses to written sentences have been measured during the experiment through the facial EMG. Participant's face areas have been prepared using alcohol, water and an abrasive paste. Recyclable Ag/AgCl electrodes have been attached bipolarly according to guidelines (Fridlund & Cacioppo, 1986) on the left corrugator supercilii (CS) and zygomaticus major (ZM) muscles. Generally, the CS knits the brow into a frown, a sign of negative affect and/or pain; the ZM raises the lip, which usually results in a smile (Massaccesi, Korb, Skoluda, Nater, & Silani, 2021). In addition, a ground electrode has been attached to the participant's forehead. The EMG data were sampled at 2048 Hz with impedances below 20 kΩ. EMG data have been preprocessed in Matlab (www.themathworks.com), in part adopting the EEGLAB toolbox (Delorme & Makeig, 2004). The data have been first filtered with a 20 to 400 Hz bandpass filter and a 50 Hz notch filter, then rectified and smoothed with a 40 Hz low-pass filter. In order to examine facial reactions to the reading of emotional/painful and neutral sentences, 4 epochs have been extracted from the onset of the presentation of the sentence (which lasted 4 seconds), with EMG activity averaged over 1 s time-windows. For each trial, values in these four epochs have been expressed as percentage of the average amplitude during the last 0.5 second of the fixation cross at the beginning of that trial, which represented the baseline measure for the facial muscles' activity. Thus, values above the baseline represent an increase in the activation of the muscle, whereas values below the baseline are considered an expression of the muscle's relaxation. Outliers in baseline values were identified as values greater or smaller than 3 SDs from the subjects' average baseline. These outlier baselines were substituted with the average amplitude of the baseline preceding and following that trial (on average, 0.9% and 2.0% per subject for the CS and ZM respectively). Due to technical failure, the activity of the CS for two participants has not been included in the statistical analyses (N = 36): in one case the electrodes were too noisy, whereas in the other case technical issues prevent recording of the muscle's activity. To reduce the effect of non-experimental movements, trials with activity of the corrugator (CS) and zygomaticus (ZM) muscles in one of the four epochs greater or smaller than three SDs compared to the subject's mean (i.e. outliers) were excluded from further analyses (on average, 2.1% and 2.0% per subject for the CS and ZM respectively). Despite this data's artifacts cleaning procedure, skewness (S) and kurtosis (K) values were still high and EMG data were consequently transformed using the natural logarithm (log).

## 3.3. RESULTS

### 3.3.1 BEHAVIOURAL RESULTS

#### 3.3.1.1. PAS ~ context

As it has been stated before, the analysis of behavioural data has been done in two phases. In the first phase, an EDA has been conducted, in order to provide a first description of the data and of the association between behavioural variables. The responses to the empathy rating and to the PAS have been considered as ordered factors. First, it has been studied the relationship between the responses to the PAS and the context, operationalized as sentences. It has been created a bar plot of the PAS distribution for valid and catch trials to check how participants have been using this scale. The PAS responses are on the x axis, while the frequencies of the responses for each modality are on y axis (Figure 3.3.).

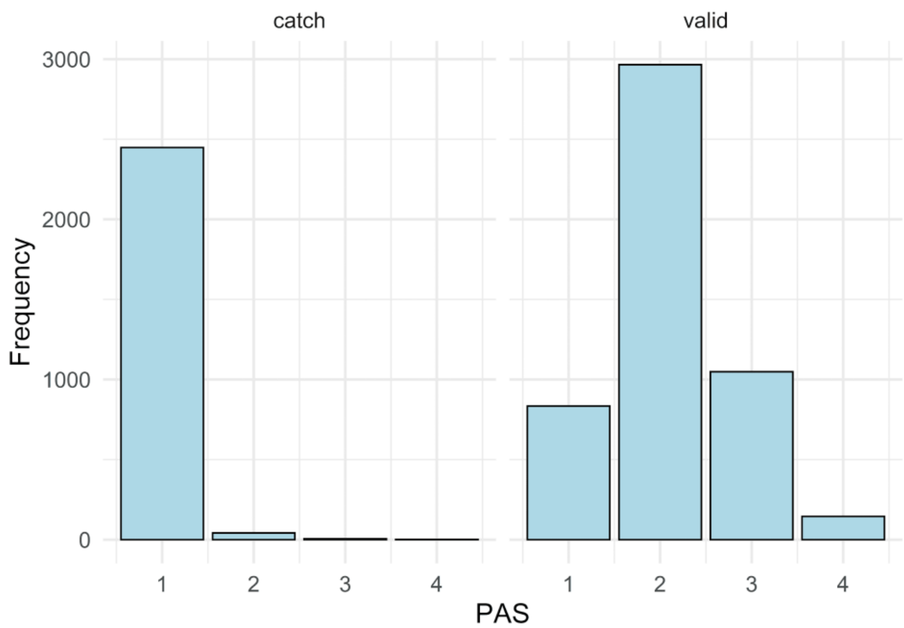


*Figure 3.3. The PAS distribution for valid and catch trials.*

It can be observed that during catch trials the majority of responses is concentrated on the 1 modality (as expected), whereas during valid trials the responses are focused on the 2 modality of the PAS, with few 4 and a good amount of 1-3. Afterwards, it has been checked the PAS distribution as a function of the context (sentences) to see if there is a change in the first variable depending on the second. Also in this case, the plot represents the responses to the PAS (x axis) as a function of the context, which could be emotional/painful or neutral. On the y axis we can see the frequencies of the responses for each modality of the PAS (Figure 3.4.)

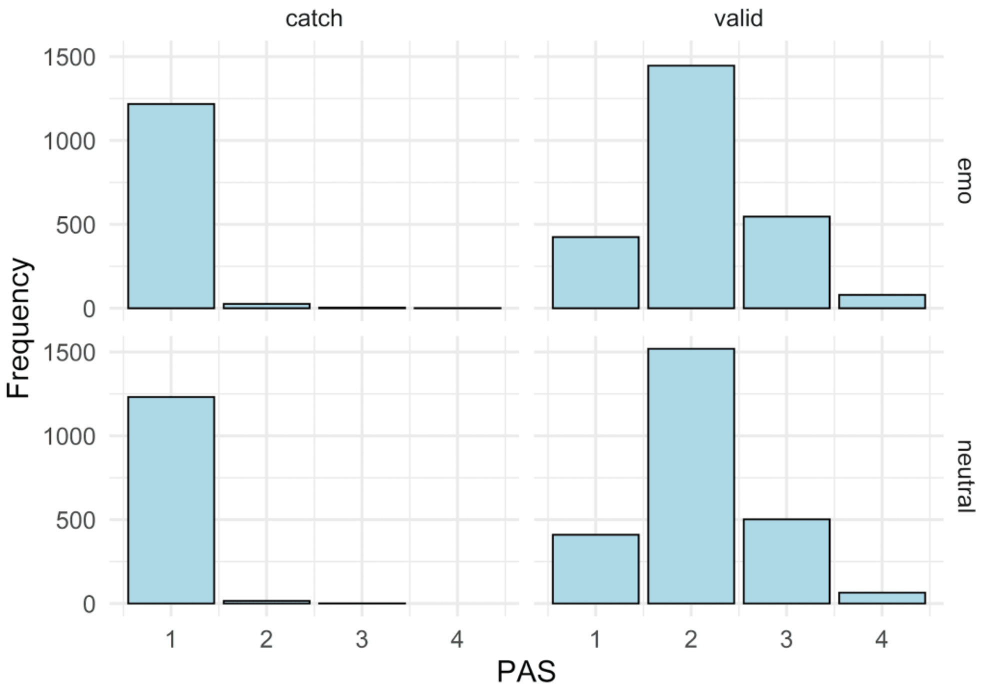


*Figure 3.4. The PAS distribution as a function of the context (neutral and painful sentences)*

As expected, for catch trials the distribution is the same regardless of the context, since they represent a condition with no stimulus, so the participants have responded that they have seen no face. Regarding valid trials, the hypothesis of the possible influence of the context on the conscious perception of noisy facial expressions has been graphically explored. No difference between the neutral and the emotional context has been detected with respect to the PAS responses. In fact, the responses are mainly concentrated on the second modality ("I have seen only a brief glimpse of the stimulus") in both contexts. Thus, the context is not influencing the PAS distribution at an exploratory level.

Afterwards, the relationship between the PAS responses and the context has been further analyzed with proper inferential statistical methods. In particular, it has been checked if the PAS distribution varies as a function of the context (neutral; emotional/painful). The aim of this first part of the investigation is to understand if the emotional context could contribute, as a single factor, to the enhancement of the emotional conscious perception of the painful stimulus, operationalized as an explicit subjective report of the facial expression. The regression model has been fitted using the ordinal package (Christensen, 2019).

| | **pasf** | | |
|---|---|---|---|
| *Predictors* | *Odds Ratios* | *CI* | *p* |
| context [neutral] | 0.91 | 0.81 – 1.03 | 0.156 |
| **Random Effects** | | | |
| $\sigma^2$ | 3.29 | | |
| $\tau_{00}$ participant | 3.36 | | |
| $\tau_{11}$ participant.contextneutral | 0.00 | | |
| $\rho_{01}$ participant | -0.85 | | |
| N participant | 39 | | |
| Observations | 4992 | | |

*Table 3.2. No significant difference has been found in the mean of the PAS responses as a function of the context.*

The potential mean difference in the distribution of the PAS based on the diverse contexts has been expressed in terms of Odds Ratios (ORs, Table 3.2.). In this case, there has been found no significant difference in the mean of the PAS responses on the basis of the variation of the context (p>0.001). Also, the OR associated with the variable "context[neutral]" is approximately 1, suggesting a small to absent difference between the two contexts (Figure 3.5.).

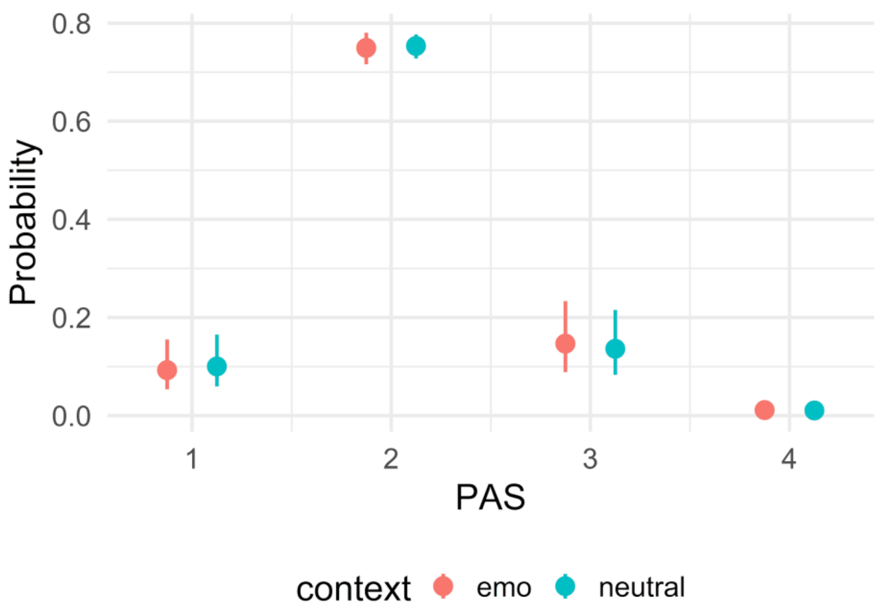


*Figure 3.5. The PAS responses as a function of the context.*

3.3.1.2. Empathy ~ context

Regarding the second hypothesis, it has been inspected if the context influences the responses at the empathy subjective rating. First, it has been checked if and how empathy responses are distributed as a function of the trial type. Given that the answer to the empathy question requires to consider both the context and the face, it can be observed that even for catch trials the sentence read influences the empathy rating. A plot has been constructed, where on the x axis there are the empathy responses and on the y axis the frequencies for each modality of the rating (Figure 3.6.).

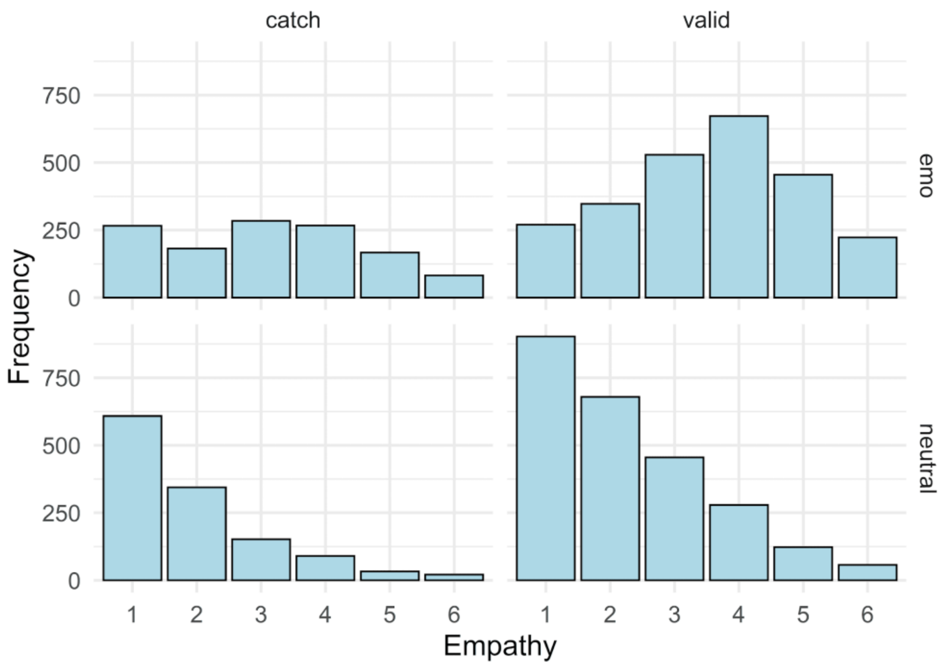


*Figure 3.6. Distribution of the empathy responses as a function of the trial type.*

For the emotional context and valid trials, it can be observed that empathy ratings tend to be higher compared to the neutral context. For catch trials, it can be noticed that empathy ratings are right skewed for the neutral context and pseudo-random for the emotional context. Thus, reading emotional sentences is associated with higher values on the empathy subjective rating, compared to reading neutral sentences.

Moreover, it has also been checked the conditional distribution of PAS ratings and empathy ratings as a function of the context. This concept can be observed in the following plot, where each PAS distribution is plotted separately for each modality of the empathy rating. In this case the relative frequencies of the PAS responses are plotted separately for each context (the sum of the bars for each context is 1; Figure 3.7.).

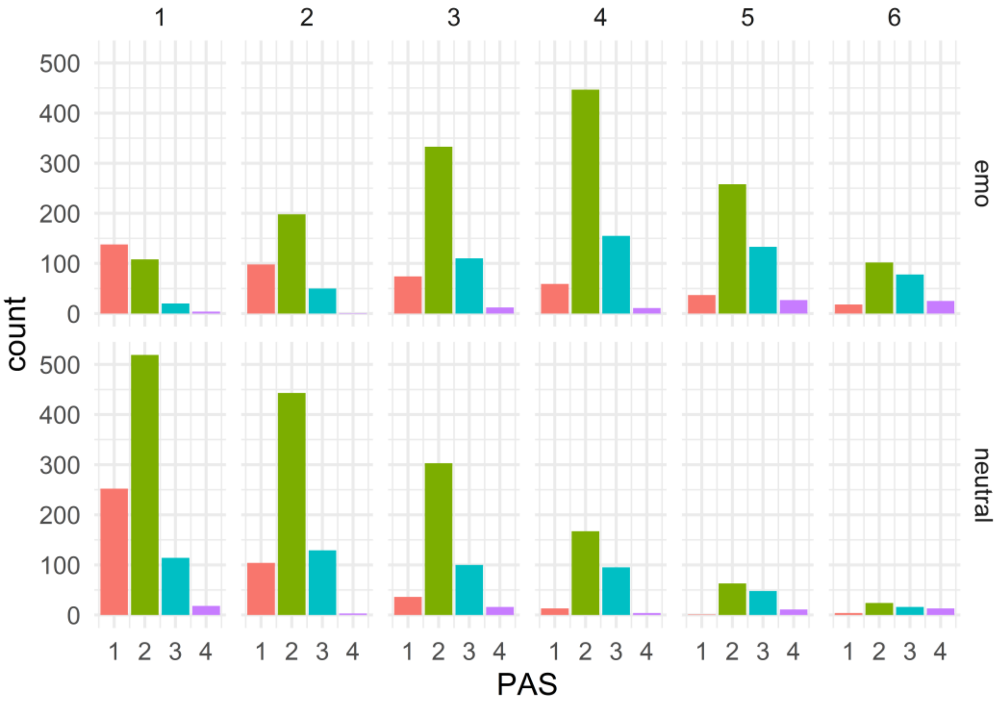


*Figure 3.7. The conditional distribution of PAS and empathy ratings as a function of the context. PAS ratings on the x axis.*

It can be noted that, in the emotional painful context, the responses to the PAS tend to increase with the increase of the responses at the empathy subjective rating, when compared to the neutral context. More specifically, it can be observed that in the emotional context the probability to answer with PAS 2 and 3 tend to increase with higher values at the empathy subjective rating, whereas there is a decrease of the PAS response 1. Instead, in the neutral context, it can be noticed that there is an opposite trend, in particular related to the PAS responses 2 and 3.

The opposite conditional distribution can be observed in the graphic above. This bar plot represents the empathy distribution within each PAS level for valid trials (Figure 3.8.).

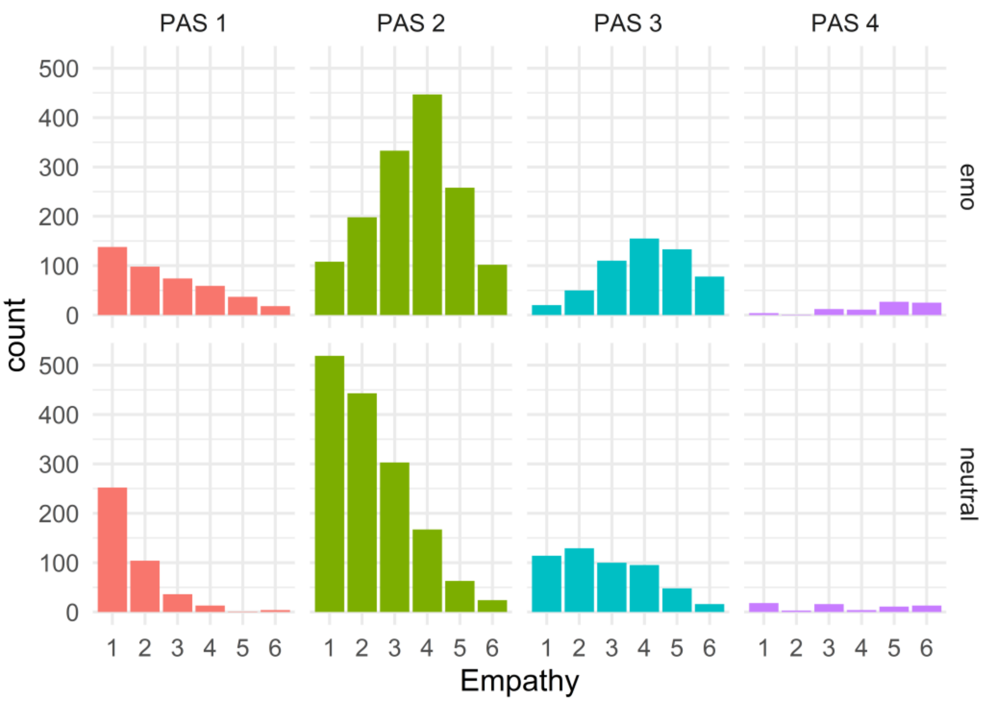


*Figure 3.8. The conditional distribution of PAS and empathy ratings as a function of the context. Empathy ratings on the x axis.*

In this case, it can be seen that in the painful emotional context the responses at the empathy subjective rating tend to increase with higher responses at the PAS rating. On the other hand, in the neutral context it seems that the responses at the empathy subjective rating tend to decrease when the subject's PAS answer was 1 or 2, and they tend to distribute quite homogeneously for the PAS responses 3 and 4.

To summarize, in the last two graphics it can be observed that the context influences the way empathy and PAS rating interact with each other.

### 3.3.1.3. Empathy ~ PAS

Going further, the relationship between PAS responses and empathy ratings has been investigated, considering the former as a possible predictor of the latter distribution. They are both subjective measures: one refers to the tendency to empathize with the actual situation that is displayed (facial expression and sentence), and the second to the level of perceptual awareness of the stimulus (facial expression).

Since the present research is an exploratory analysis, the PAS has been considered as a numeric variable, and this is obviously questionable. Given that the PAS rating is expressed trial-by-trial, it is possible to consider this factor's effect in two terms (Enders & Tofighi, 2007):

1. A within-participant effect: the question is if, within each participant, it is possible to find a relationship between PAS responses and Empathy ratings. This means to investigate if the tendency of a participant to use higher or lower PAS responses influences Empathy ratings within the single experiment.
2. A between-participant effect: in this case, the aim is to understand if, across participants, there is a relationship between PAS and Empathy. In other words, the purpose is to understand if the fact that a participant has a higher or lower *average* PAS rating has an effect on Empathy ratings.

Again, it was possible to fit a multilevel ordinal model using the by-participant average PAS rating (i.e. between effect) and the within-participant centered score (i.e. subtracting each PAS rating from the average rating of each participant). PAS cm indicates the between effect, whereas PAS cc represents the within effect. The aim of the within model is to see if, within each subject, the variation of the PAS rating predicts a variation of the Empathy rating. The purpose of the between model is to understand if, between all the subjects, the variation of the PAS rating predicts a variation of the Empathy rating.

| | empathyf | | |
|---|---|---|---|
| *Predictors* | *Odds Ratios* | *CI* | *p* |
| pas cm | 4.55 | 2.30 – 8.98 | **<0.001** |
| pas cc | 1.80 | 1.35 – 2.41 | **<0.001** |
| **Random Effects** | | | |
| $\sigma^2$ | 3.29 | | |
| $\tau_{00}$ participant | 0.78 | | |
| $\tau_{11}$ participant.pas_cc | 0.69 | | |
| $\rho_{01}$ participant | -0.08 | | |
| N participant | 39 | | |
| Observations | 4992 | | |

*Table 3.5. The PAS responses have a significant effect on the Empathy distribution.*

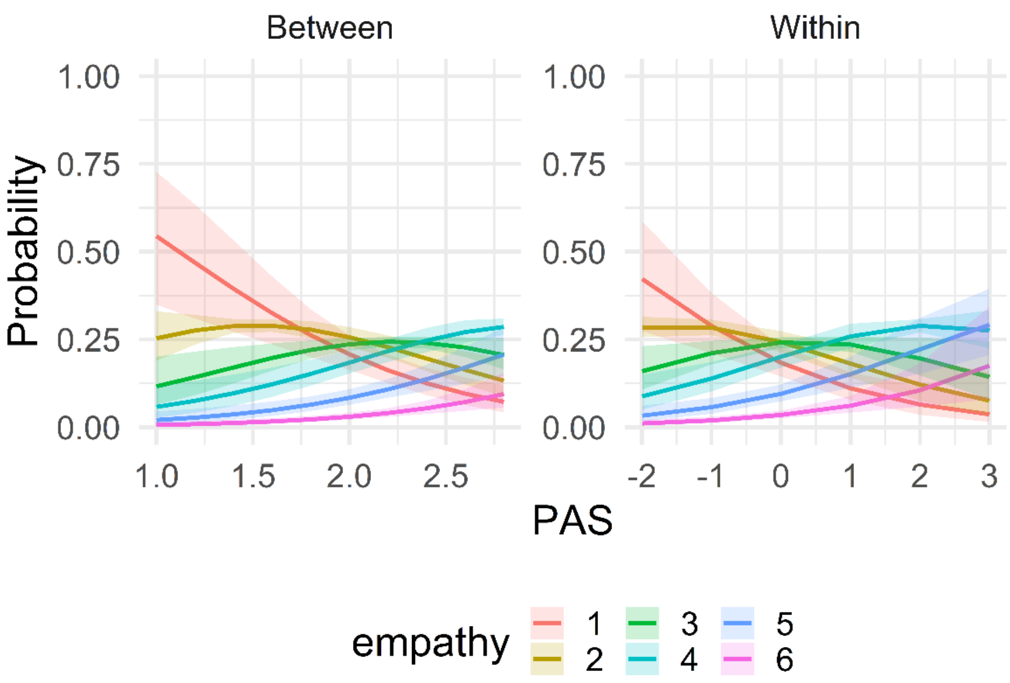


*Figure 3.9. Empathy ratings as a function of the PAS ratings, between and within subjects.*

The model shows that the PAS responses have a significant effect on the Empathy distribution ($p<0.001$; Table 3.5.). In particular, higher PAS responses predict higher Empathy ratings, since the Odds ratios are higher than 1 (i.e. positive association: one unit increase in the PAS rating is related to an odds increase of the Empathy rating of 4.55 for the between effect and 1.80 for the within effect; Figure 3.9.). This means that there is a higher probability to answer with a higher Empathy rating if the participant has perceived more clearly the painful facial expression, both between and within subjects.

To summarize, it can be affirmed that the PAS responses and the Empathy ratings are significantly associated with each other: an increase in one of these two factors implies a corresponding increase in the other one.

3.3.1.4. Questionnaire of Cognitive and Affective Empathy (QCAE)

The fourth and last part of the statistical analysis concerns the study of the results of the QCAE (Questionnaire of Cognitive and Affective Empathy; Reniers et al., 2011), a multi-dimensional self-report measure of empathy. First, it has been checked the correlation structure of the QCAE, of its subscales and of the average PAS and Empathy ratings for each subject. As expected, it emerged no relationship between the questionnaire responses and the average PAS and Empathy ratings.

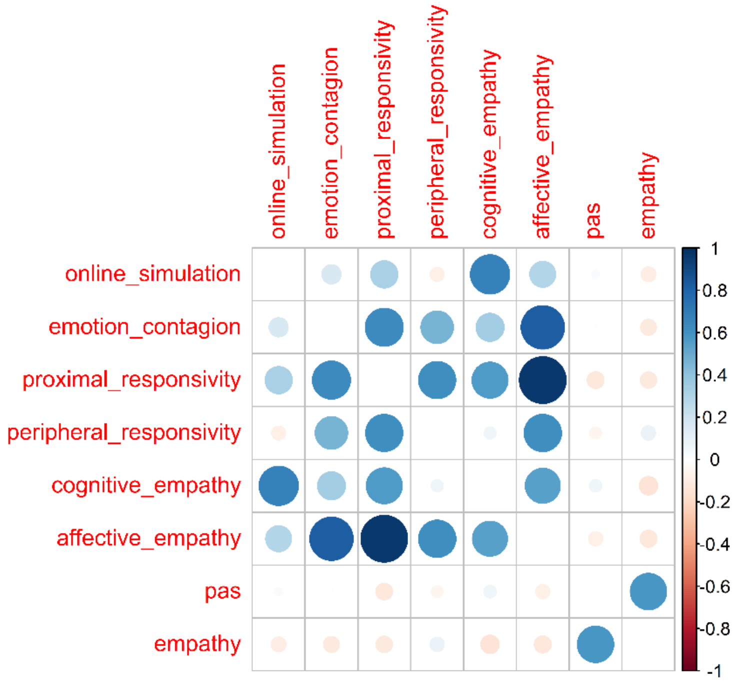


*Figure 3.10. Correlation matrix of the QCAE, of its subscales and of the average PAS and empathy ratings for each subject.*

In the correlation matrix it can be noted that weak correlations have been found (Figure 3.10.). Then, through a multiple linear regression analysis, it has been checked if the average empathy subjective rating's score for each subject correlates with the empathy questionnaire responses, including the interaction with the context, which is a within-subject factor (Table 3.6.; Figure 3.11.). This means that we consider the effect of the context in each subject's experiment.

| | empathy | | |
|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* |
| (Intercept) | 2.91 | 2.69 – 3.14 | **<0.001** |
| cognitive empathy0 | -0.01 | -0.06 – 0.03 | 0.562 |
| affective empathy0 | -0.01 | -0.05 – 0.03 | 0.743 |
| context1 | 1.26 | 0.94 – 1.59 | **<0.001** |
| cognitive empathy0 * context1 | -0.02 | -0.09 – 0.04 | 0.459 |
| affective empathy0 * context1 | 0.02 | -0.04 – 0.08 | 0.455 |
| **Random Effects** | | | |
| $\sigma^2$ | 0.51 | | |
| $\tau_{00}$ participant | 0.24 | | |
| ICC | 0.32 | | |
| N participant | 39 | | |
| Observations | 78 | | |
| Marginal $R^2$ / Conditional $R^2$ | 0.359 / 0.564 | | |

*Table 3.6. No relationship has been found between the self-reported empathy of the QCAE and the average trial-by-trial empathy.*

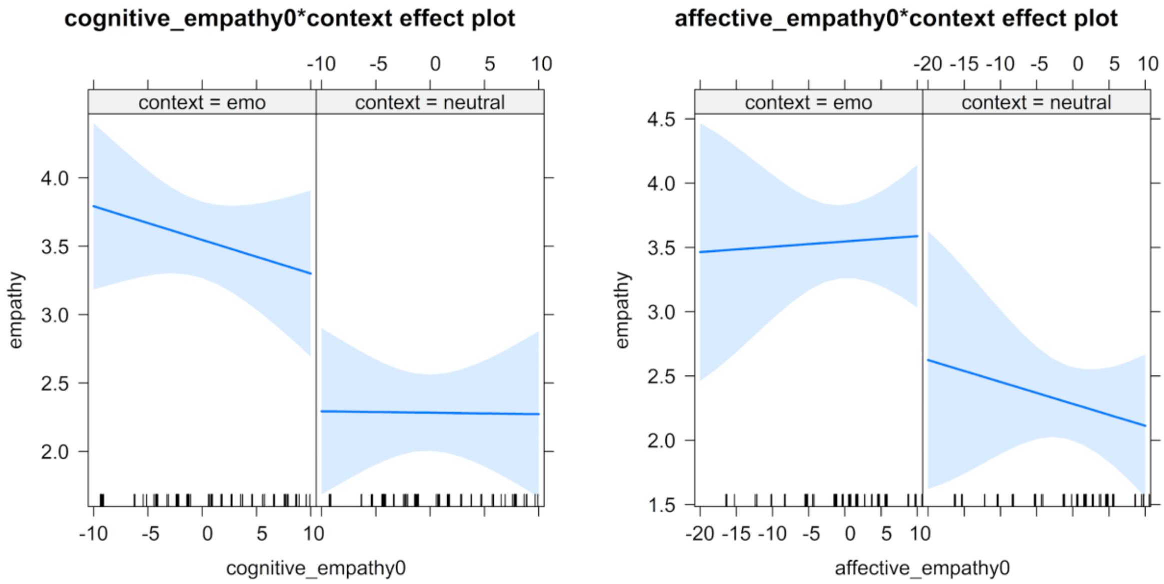


*Figure 3.11. Trial-by-trial empathy ratings as a function of cognitive and affective empathy of the QCAE, with the interaction of the context.*

It has been found that there is no relationship between the self-reported empathy of the QCAE and the average trial-by-trial empathy (experimental session phase). Instead, only the main effect of the context has emerged, which is the same that has been seen in the previous models without considering the questionnaires: in the emotional painful context, compared to the neutral one, the empathy ratings are shifted to higher values (Figure 3.6.). In other words, the empathy ratings differ between the two contexts when we consider the average rating of the QCAE. These results can be observed in the graphics considering that the slope does not significantly change between the two contexts (Figure 3.11.). The same scenario can be noted when analyzing the PAS ratings predicted by the QCAE responses (Table 3.7.; Figure 3.12.).

| | | pas | |
|---|---|---|---|
| *Predictors* | *Estimates* | *CI* | *p* |
| (Intercept) | 2.10 | 1.96 – 2.24 | **<0.001** |
| cognitive empathy0 | 0.01 | -0.02 – 0.04 | 0.407 |
| affective empathy0 | -0.01 | -0.04 – 0.01 | 0.373 |
| context1 | 0.02 | -0.01 – 0.05 | 0.125 |
| cognitive empathy0 * context1 | -0.00 | -0.01 – 0.00 | 0.577 |
| affective empathy0 * context1 | 0.00 | -0.00 – 0.01 | 0.084 |
| **Random Effects** | | | |
| $\sigma^2$ | 0.00 | | |
| $\tau_{00}$ participant | 0.19 | | |
| N participant | 39 | | |
| Observations | 78 | | |

*Table 3.7. No relationship has been found between the self-reported empathy of the QCAE and the PAS ratings.*

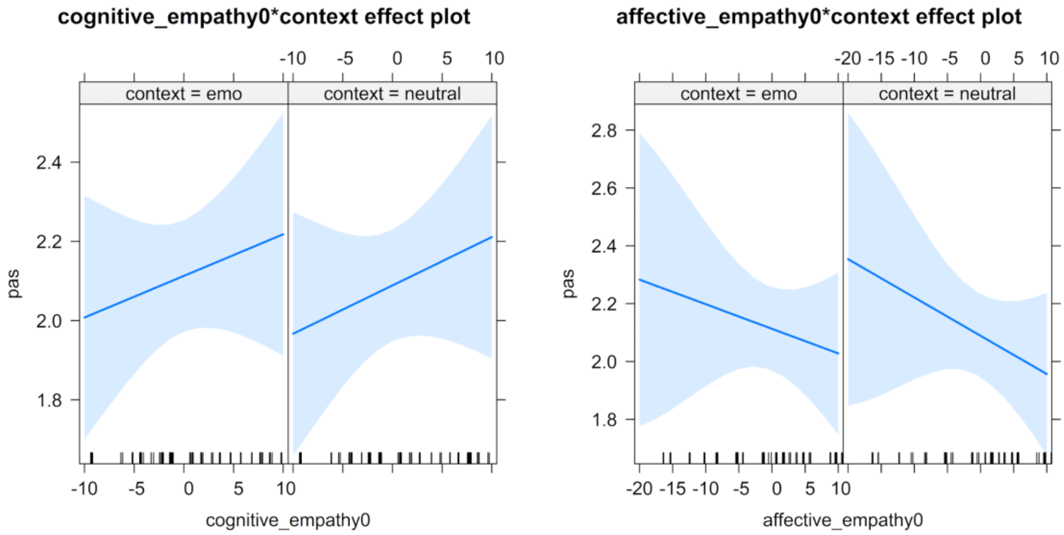


*Figure 3.12. PAS ratings as a function of cognitive and affective empathy of the QCAE, with the interaction of the context.*

It has been found that the QCAE responses do not significantly predict the PAS ratings.

### 3.3.2 PSYCHOPHYSIOLOGICAL RESULTS

*EMG*

Facial muscle activity of the ZM and CS have been recorded during the whole experimental session phase task, and markers have been placed along the EMG track in specific moments. In fact, the aim was to observe the eventual variation in facial muscles' activity in response to certain stimuli as, in this case, the presentation of the sentences. It is known that the ZM muscle's activity is related to positive affect, whereas the CS activity is associated with negative affect or pain. The main purpose has been to test if these facial muscles' activity can represent a valid and reliable mapping of the context's emotivity. In particular, we aimed to observe if there is a difference during the reading of emotional/painful and neutral sentences in the activity of the two muscles. The hypothesis we have elaborated is that, during the presentation of painful sentences, the activity of the CS is higher compared to the presentation of neutral sentences, since it is supposed to reflect a higher negative/painful affect. Furthermore, we have expected that the ZM muscle's activity does not significantly change between the two contexts. In order to test these hypotheses, the EMG signal has been extracted from the moment in which the sentences were presented, divided in 4 epochs that have been then averaged. The signal has not been extracted from the time when the painful faces were displayed, since the stimuli were only painful and not also neutral, and because the face has been presented only for 50 ms. In fact, it would be necessary to display them for at least 150 ms to process the EMG signal.

We have previously seen that the context itself has not been revealed to be a good mediator of others' emotional expressions' awareness, so we have thought to examine if more subtle psychophysiological mechanisms related to the context are involved. The future purpose will be to investigate if the facial muscles' activity could mediate the possible effect of the context on the conscious perception of the facial expressions of pain. In fact, we suppose that embodiment processes that may be involved during the reading of emotional sentences through sensorimotor simulation, could modulate our capacity to become aware of others' emotions.

Since the aim of the study was to understand if, during the 4 seconds' sentence presentation, the CS and ZM muscles activity varies between painful and neutral sentences, the data have been expressed in long format: this allowed us to split each muscle's activity variable in the 4 epochs of time for each participant. Afterwards we have checked for the normality of the data, since it is known that EMG data tend to be skewed, but this is not ideal to conduct an Anova statistical analysis. Figure 3.13 depicts the distribution of EMG CS activity as a function of the context.

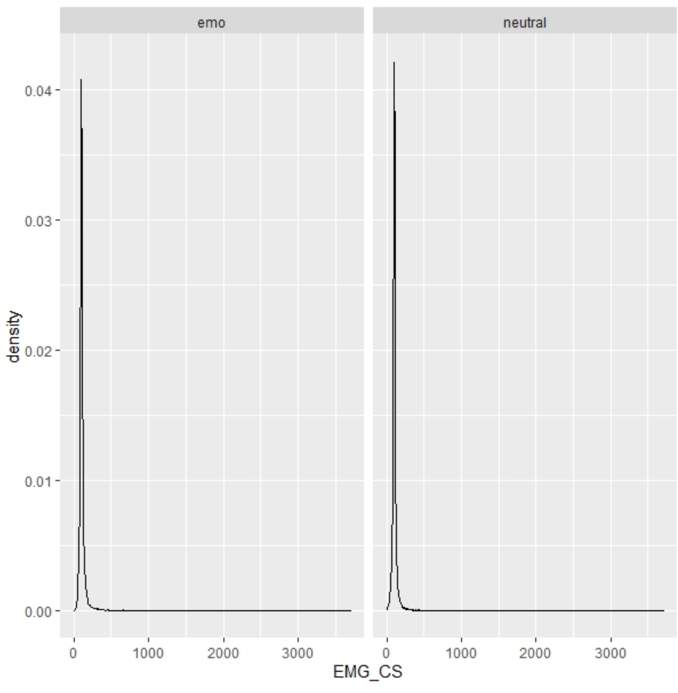


*Figure 3.13. Distribution of the CS activity as a function of the context. The data are right-skewed.*

Firstly we removed the outliers, that are values higher than 3 standard deviations with respect to the mean of the activity for a specific muscle (ZM or CS) in a specific context (painful or neutral) and in a specific epoch of time (that were 4 in total). It has been seen that, in general, the percentages of outliers were not so high that they had to be taken into consideration. In particular, at the baseline level of activity outliers were present in the 2.0 % of the trials for the ZM and in the 0.9 % of the trials for the CS. In addition, they were found to be 2.0% for the ZM and 2.1% for the CS during the presentation of the sentences. Afterwards, to further correct the skewness of EMG data, we have applied a logarithmic transformation (Figure 3.14.).

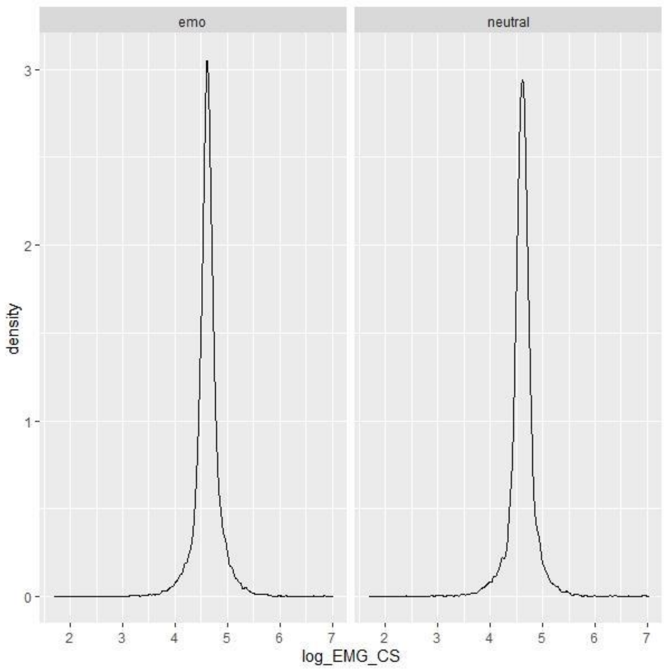


*Figure 3.14. CS activity as a function of the context. The data's skewness has been corrected.*

Subsequently, the proper inferential statistical analysis has been conducted, and it has been divided into two main parts: a t-test and an Anova. The first t-test consisted in studying the difference of the muscles' activity (ZM, CS) in the two contexts (painful, neutral) for each subject, averaging the 4 epochs of time. For the CS, two participants have been excluded from the analysis due to technical issues, so that we had 37 out of 39 participants. The results of this first analysis are that the activity of the ZM in the emotional context is significantly different compared to the neutral context (t= 2.12, df= 36, p-value= 0.041). The mean of the differences is 0.012, thus it is quite low. Nonetheless, it is important to remember that the aim of this part of the investigation is to observe subtle differences in the activity of the facial muscles, according to simulation theory (Gallese, 2001), and also that the data have been logarithmically transformed. In addition, the difference sign is positive, thus during the emotional sentences' reading, the activity of the CS is higher than the one during the neutral sentences' presentation. On the other hand, when observing the results for the ZM muscle activity, it can be seen that there is no significant difference between the two contexts (t = 1.88, df = 38, p-value =0.07, mean of differences= 0.006). Taking into consideration only the last 2 seconds of the 4 epochs, the same results have emerged.

The second analysis refers to the Anova, where the activity of the ZM and of the CS muscles has been separately associated with the different contexts. This time the 4 epochs have been individually investigated. For the CS, it has been found that there is a main effect of the context, a main effect of the time and a significant interaction between context and time. This means that the activity of the CS has been discovered to be different in the two different contexts independently of the time, considering time in the four epochs, and in the two contexts in every phase of the sentence's presentation. So, the results for this first model are the following: the p-value for the context main effect was 0.041; the p-value for the Time main effect was minor than 0.001; the p-value for the context:Time interaction was 0.048. On the other hand, the activity of the ZM has not been found to vary with the context, with time, and with the interaction between context and time: the main effect of the context (p-value= 0.068), the main effect of Time (p-value= 0.479) and the interaction context:Time (p-value= 0.289) have not resulted to be significant.

To summarize, the t-test and Anova results have been found to be in accordance with each other. In order to better visualize the results, two different graphics have been created for the interaction effect of context and Time (predictors/independent variables) on the CS and ZM muscles' activity (dependent variables).

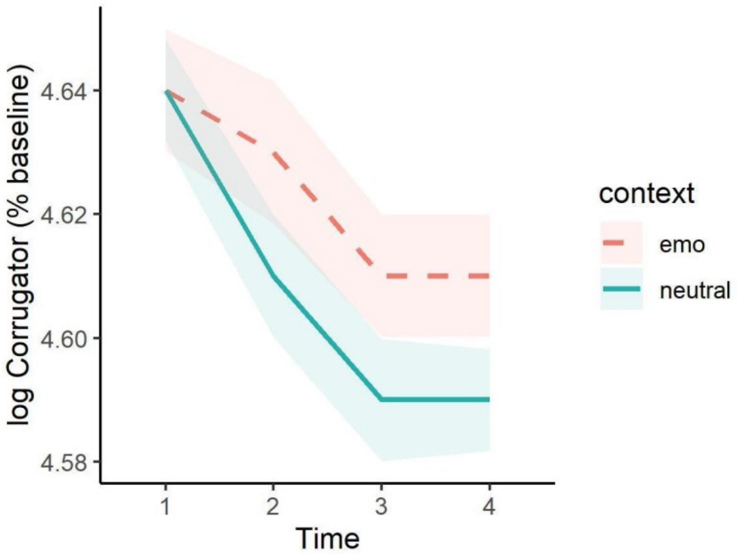


*Figure 3.15. CS activity as a function of the context.*

In the first graphic (Figure 3.15), the activity of the CS muscle (logarithmized in percentage with respect to the baseline, which is approximately 4.6) has been plotted on the y axis and the Time on the x axis, whereas the two colors indicate the different contexts that have been examined. It can be noted that the activity of the CS muscle decreases in both contexts during the 4 epochs of the sentence presentation, but from the last 2 seconds it clearly differentiates: in the neutral context the reduction is more pronounced than in the emotional context, as it can be observed in the confidence intervals that separate from each other. Moreover, it can be seen that the CS activity during the neutral sentences' presentation is below the baseline threshold, indicating that the muscle starts to relax. On the other hand, during the painful sentences' presentation, the CS activity is above the baseline threshold, suggesting that the muscle is more active than the average.

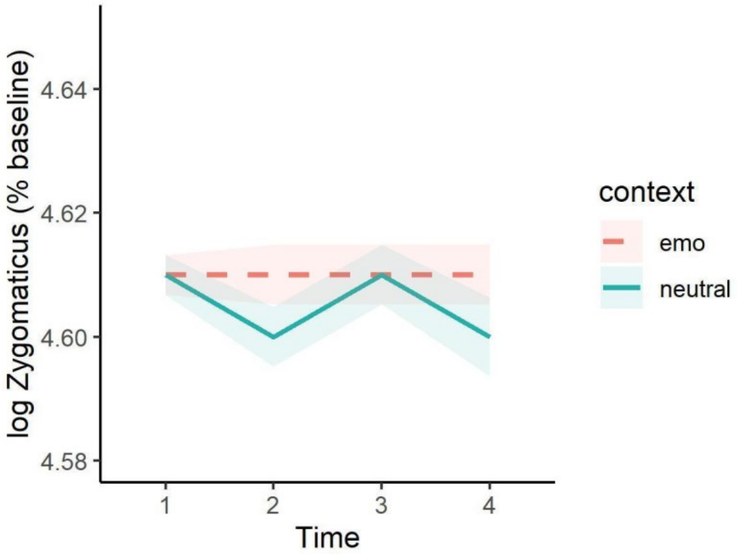


*Figure 3.16. ZM activity as a function of the context.*

In the second graphic (Figure 3.16.), it can be noted that there is no difference in the activity of the ZM between the two contexts over time.

The reason for the first reduction in the CS activity during the initial presentation of the sentence is not known: nonetheless, it could be hypothesized that the participant was concentrated to read and understand the meaning of the sentence and, only then, he/she realizes the emotional nature of the context.

Ultimately, a barplot has been built specifically for the last fourth epoch during the presentation of the sentences, since the effect has been noted especially between the third and the fourth seconds of time. In fact, the same results have been found (Figure 3.17.; Figure 3.18.) and the horizontal bar for each context represents the standard error of the mean (which have been used also for the confidence intervals computation).

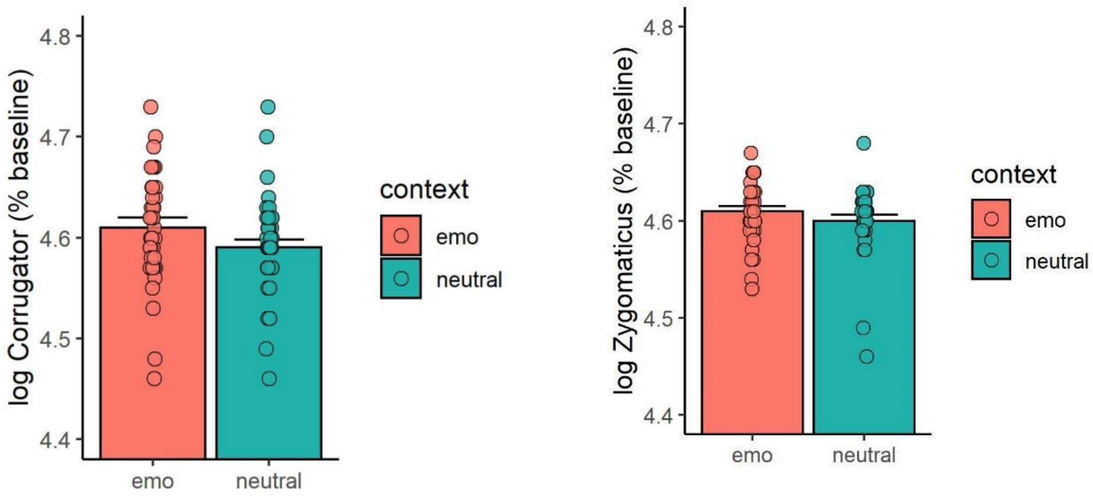


*Figure 3.18. CS and ZM activity as a function of the context, in the last fourth epoch of the presentation of the sentence.*

Finally, it can be stated that the CS facial muscle's activity seems to be a good psychophysiological measure sensitive to differences between painful and neutral linguistic contexts.

# 4. DISCUSSION

## 4.1. THE STUDY AND THE HYPOTHESES

The purpose of the present study has been to investigate the association between empathy and consciousness within a linguistic contextual framework. The levels of analysis involved both behavioural and psychophysiological measures. In particular, the aim has been to observe the influence of emotional (i.e., painful) and neutral sentences on both empathy for pain and the conscious perception of facial expressions of pain. These latter stimuli were not completely visible due to experimental manipulation.

The first hypothesis of this research was to understand if emotional contextual information, operationalized as painful sentences, could enhance the conscious perception of facial expressions of pain and empathetic processes related to the overall situation of pain (i.e., face and contextual stimulus) when compared to neutral sentences. The responses to the Perceptual Awareness Scale (PAS; Overgaard et al., 2006) have been used as a measure of the conscious perception of the facial expression, intended as a gradual subjective explicit report of the visual stimulus. Instead, empathy has been measured through a 6-point Likert scale that represents a subjective rating of the empathic experience related to the facial expression and the situation described by the sentence. Both measures were collected during the experimental phase.

A second hypothesis of the study was to disentangle the possible role of psychophysiological measures in mediating the influence of the context on the dependent variables (painful faces' conscious perception and subjective empathy rating). For the present research, which has an exploratory and preliminary nature, only one measure has been implemented, that is electromyography (EMG), which registers the facial muscles' activity, and it has been recorded during the presentation of the sentences. In line with the studies conducted under the simulation theory framework (Wood et al., 2016), EMG has been used to measure facial mimicry. In fact, the underlying assumption is that an individual imperceptibly mimics another person's observed emotional facial expression in order to recognize the other's emotion (Dimberg et al., 2000) and deeply understand the affective experience of the other person (Gallese, 2001). The two processes can be conceptualized as emotion recognition and empathy, and the former is the basis of the latter. Specifically, it has been hypothesized that the EMG mimicry response will be larger in congruent context-face emotional trials than in incongruent trials (i.e., neutral context and painful facial expression).

As it has been stated above, this suggestion is in line with the simulation theory (Gallese, 2011), for which an individual can empathize with another person's emotion by reproducing internally and simulating in a sensori-motor manner the affective experience observed in the other person. Moreover, this hypothesis is in accordance with the embodiment theory (Damasio, 1994; Gallese, 2001), for which empathy has not only mental and cognitive nuances, but it also involves more immediate and "whole-body" processes. In addition, following the predictive coding framework (Friston & Kiebel, 2009) and the theory of constructed emotions (Barrett, 2017), the perception of a stimulus, like an emotion, is constructed through the integration of top-down expectations and bottom-up processes. The first ones can involve visual stimuli (like emotional sentences) or anticipatory somatic feedbacks (like mimicry processes or arousal responses), which then are connected to our past experience associated to the actual sensory situation. The second ones are related to external sensory information, like painful facial expressions, that in the present study's case were not completely visible. For example, it could be hypothesized that, when the participant has to construct the painful emotional stimulus, the reading of painful sentences and mimicry responses act as anticipatory information or simulations that, together with the actual bottom-up stimulus related activity, contribute to the perception of the face.

A third hypothesis regards the relationship between empathy and consciousness, that in the present research has been considered as an explicit subjective report of the emotional stimulus and has been measured with the PAS (Overgaard et al., 2006). This aspect has been investigated in two ways: by looking at the association between the empathy subjective rating implemented during the experiment and the PAS responses, in order to analyze if a greater conscious perception of an emotional stimulus could boost the predisposition to empathize with the stimulus and the contextual situation related to it, or vice versa; by observing the possible link between the Questionnaire of Cognitive and Affective Empathy (QCAE; Reniers et al., 2011) and the PAS responses, to see if an higher trait cognitive and affective empathy could enhance the conscious perception of an emotional stimulus, or vice versa.

## 4.2. THE EXPERIMENTAL DESIGN: BEHAVIORAL AND ELECTROMYOGRAPHIC MEASURES

In order to test these three main hypotheses, an experimental paradigm composed by 2 phases (calibration and experimental session) has been implemented and a conclusive questionnaire has been administered to a final sample of 39 participants. The calibration phase has enabled us to identify, for each subject, threshold stimuli for the painful facial expressions, that are defined as those that are consciously perceived in 50% of the trials. In the experimental session phase, the participants had to read some sentences, either neutral or painful and, afterwards, they were presented with a painful facial expression, at the individual threshold. Then, they had to perform two judgments: evaluating whether they have seen the face or not on a scale from 1 to 4, through the PAS (Overgaard et al., 2006) and, afterwards, expressing how much they have empathized with the person depicted in the face in the situation described by the sentence, on a scale from 1 to 6. During the presentation of the contextual information, the EMG has been recorded in order to measure the Zygomaticus Major (ZM) and the Corrugator Supercilii (CS) activity. In addition, the SC response was registered during the whole experimental session phase, but but the SC analyses are currently still ongoing. In both phases, catch trials have been included, that represent non-face stimuli used to estimate the level at which a participant is guessing when no stimulus is present (VandenBos, 2007). In the last part of the study, the QCAE (Reniers et al., 2011) has been administered to the participants, to measure cognitive and affective empathic traits that could predict or be correlated with the main effects that have been hypothesized above (increased conscious face perception and empathic subjective ratings).

The analysis has been divided into two parts: one for the behavioural data and another for the EMG data.

### *4.2.1.* Discussion of the behavioral results

*The impact of the context on the conscious perception of emotion*

Behavioural data have been first inspected graphically, through an exploratory data analysis (EDA; Behrens, 1997) and, afterwards, using linear regression models. The first result concerns the relationship between the PAS responses and the contextual information, where the latter has been hypothesized to predict the former. As expected, the EDA has shown that the distribution of the PAS was the same regardless of the context. Indeed, also for valid trials the responses of the PAS were concentrated on the same modality in both emotional painful and neutral contexts. Going further, a multilevel ordinal logistic regression (Agresti, 1989; 1999) has been used for the inferential part, since the dependent variable is categorial and ordinal. We included by-subjects random intercept and slopes.

The model has been implemented using the ordinal R package (Christensen RHB, 2019). The context effect is evaluated using a Wald z test. Results on Table [3.2.] are reported in terms of odds ratios.

Again, it has emerged a not significant difference in the distribution of the PAS between the painful and neutral contexts. So, the context, operationalized as written sentences, has no significant effect on the conscious perception of facial expressions of pain.

This first outcome could be accounted for in different ways, that probably have to be integrated. Primarily, it could be that the calibration phase has not been implemented by the researchers or understood by the participant in a correct way, with the result that the threshold stimuli were not good defined, and they were not a reliable representation of what we wanted to measure. In fact, many of the participants have reported that they were not sure to have dealt with the calibration phase in the right way, and they have complained that the stimuli in the experimental session phase were very difficult to perceive. Indeed, the distribution of the PAS in both neutral and painful contexts was concentrated on the PAS 2 responses ("I have seen only a brief glimpse of the face"), which indicates a poor perceptual experience of the stimulus. A second possible explanation could refer to the fact that reading sentences is not so effective in its emotional impact on the subject as seeing pictures of emotional situations (e.g., faces or objects; Conte, Brenna, Ricciardelli, & Turati, 2018; Carroll & Young, 2005; Ibáñez, Hurtado, Lobos, Escobar, Trujillo, Baez, & Decety, 2011) or hearing expressions of pain with the integration of the prosody (Sessa et al., 2018). It has to be noted that the majority of priming studies that have investigated how emotion recognition can be facilitated by previously presented stimuli regards other types of emotional experiences (e.g., anger, happiness etc.), and not specifically painful ones. A third account could concern the fact that the contextual information alone is not sufficient to enhance the conscious perception of the emotional stimulus but, according to the multimodal integration hypothesis (Barrett, 2017), different sources of information are needed to construct the perception of an instance of emotion. Apart from the theory of constructed emotions, this suggestion is also in line with Northoff and Lamme's Temporo-Spatial Theory of Consciousness (TTC; 2020). The authors have hypothesized that there are different ways to conceive consciousness, since it is a multilevel concept. In our case, the target of the study is the conscious explicit subjective report of a stimulus, which necessitates a multisensorial integration of various variables. This is a close concept to what Northoff and Lamme (2020) called "nestedness": with this term they meant that consciousness and the brain's neural activity are hypothesized to be characterized by layers, that contain (i.e., nest with) each other. Going further in the explanation, consciousness can have different functions, that can be integrated into each other to reach a more explicit and detailed representation of reality: the function may be sensory (perceptual organization), cognitive (access or prediction), bodily (neural monitoring of bodily input). The innovative proposal of TTC theory, is that there is a basic function underlying consciousness: it is a temporo-spatial dynamic process, which mediates sensory, bodily and cognitive functions by operating across different regions of the brain. Thus, these latter considerations further support the idea of multisensory and multimodal integration mechanisms that contribute to enabling consciousness.

Nonetheless, it is relevant to consider that the same sentences in the Italian version have modulated a differential effect on an electrophysiological level, using the ERP (event-related potentials; Sessa et al., 2014) as the method of measurement: painful contexts elicited a larger P3 ERP component than neutral contexts. As we will specify in the study's limitations, in the cited study the target was only empathy for pain and not the conscious perception of an emotional stimulus.

*The impact of the context on empathy ratings*

The second result regards the investigation of the association between empathy, operationalized as a subjective rating during the experimental session phase, and the contextual information, considered as the predictor of the former. It has been demonstrated that the context influences empathic responses both for catch and for valid trials. This outcome has been observed at a graphical level through the EDA analysis: in catch trials empathy ratings are left skewed for neutral context and pseudo-random for the painful context; indeed, for valid trials empathy ratings are higher in the emotional context than in the neutral context. Thus, written painful sentences have been found to be associated with higher values on the empathy subjective rating, when compared to written neutral sentences. The context seems to be a factor that, alone, can influence the empathic experience of the participant. Various possible implications can be drawn from this result: for instance, when another person is in a painful situation, but the individual is not able to see it completely, only the fact to know something about the context can facilitate his/her ability to understand and feel what the other is experiencing. This could happen if the person is not present in the actual painful scenario, but he/she reads about it in the newspaper or someone else informs him/her. The relevant aspect of this outcome is that a higher empathic response could be connected to the tendency to help the person that is in a dangerous situation. This suggestion is supported by studies that have investigated the possible relationship and transition process from empathy to altruism. For instance, Waytz, Zaki and Mitchell (2012) have proposed that altruism is mediated by the cognitive component of empathy, namely mentalizing. In this study, participants received some written biographical information and photos of other people. Afterwards, they had to judge some sentences on possible attitudes of these people depicted in the pictures and described by the written biographical information. Finally, they had to decide how much money they would like to donate to each of them. They have found that with the increase of the medial dorsal prefrontal cortex (dorsal MPFC) activity during the task, there was also an increase in the probability of donating. It is known that this brain area is associated with the cognitive empathy component, thus this aspect of empathy could be important in predicting a possible aid behaviour. In particular, in our study, this finding is relevant because the participants are presented with sentences that describe dangerous and painful situations for another person. When the subject understands the contextual information, then this could facilitate cognitive empathic processes of reasoning on the other's affective experience and, hence, prosocial behaviours. On the other hand, Hein, Silani, Preuschoff, Batson and Singer (2010) have proposed that more automatic, implicit, and unconscious affective empathy mechanisms can predict aid behaviours. In their experiment, participants in the fMRI scanner saw members of their ingroup or of their outgroup receiving an electric shock. The subjects could decide if to ease the other person's pain by undergoing the electric shock themselves. The first result of the study is in line with the assumption that the activation of the affective empathy component areas (AI and ACC) is modulated by the psychological, social, and affective closeness between the observer and the observed person: there was an increase in the AI and ACC activation when the other person in pain was part of the ingroup compared to the outgroup. Moreover, with the rise in the difference of the AI activation between ingroup and outgroup, it has been observed a correspondent increase in the probability to help the other person by experiencing the painful shock in first person. So, these outcomes suggest that, at least in part, also the affective component of empathy could mediate prosocial behaviours. Thus, it can be assumed that in our study both cognitive and affective empathy's components could come into action to predict helping behaviours. On one hand, the mentalizing process could act when the participant reads the sentences; on the other hand, it is also important to consider that the presentation of the sentence and of the facial expression are quite fast, thus not enabling a deep reasoning on the painful situation. Hence, it could be hypothesized that, in this specific case, the cognitive component could not be sufficient to activate empathic responses and prosocial behaviours in absence of a more automatic and affective component. This is supported by the fact that most of the participants, at the end of the study session, have reported that they could empathize more with the person depicted in the facial expression, when the contextual background recalled them of a painful situation experienced in first person in the past. This could mean that more automatic processes come into action when they have to "put themselves in the other's shoes", since past experiences can immediately activate without a conscious and explicit effort. Furthermore, as we have seen in Barrett's theory of constructed emotions (2017), these past experiences are coded into quite stable embodied and multisensorial representations, and they act as predictions when the individual perceives the actual emotional situation. We could think of these representations as having two nuances that can be integrated with each other. On one hand, they have a slow, explicit, and cognitive nature since they are stable and learnt over repeated experiences and they constitute top-down mechanisms. On the other hand, they can act in an automatic and unconscious way, hence activating more affective empathic processes. In other words, the nature of the representations could be considered as distinct from the way they can activate. Coming back to the present research results, it can be stated that, even in a not-completely visible painful situation, language could facilitate the tendency to empathize and to implement prosocial behaviours.

The difference between emotional painful and neutral contexts has been observed also when graphically exploring the conditional distributions of the PAS and empathy subjective ratings, as a function of the context. It has emerged that in the emotional context higher PAS responses are associated with higher empathy subjective ratings, and vice versa. On the other hand, in the neutral context it has been observed an opposite trend.

*The relationship between conscious perception and empathy ratings*

This last aspect can be reconnected to the third hypothesis, which concerns the relationship between empathy and consciousness. We have decided to analyze the possible situation where the PAS responses represent a predictor of the empathy ratings. It has been found that the PAS responses have a significant effect on the empathy distribution: higher PAS responses predict higher empathy ratings. In general, it can be stated that PAS ratings and empathy subjective ratings are positively associated: the increase in one of the two terms implies an increase in the other one.

This result could be interpreted theoretically in two ways: a higher conscious perception of the emotional stimulus leads to a higher tendency to understand empathically the other’s affective experience, independently of the contextual situation (i.e., congruent written sentences); a higher individual tendency to empathize with the other person’s affective states brings about an higher capacity to consciously access the emotional stimulus. Both views can be supported by the present research’s theoretical background. In fact, Barrett (2017) has proposed that both top-down and bottom-up processes are involved in the construction of an emotion, that could be extended to the other’s affective experience. Top-down processes involve predictions, that are considered as multimodal summaries that become more detailed as they stream to more developed and granular layers of the cortex. In this case it could be suggested that a higher recognition of the other’s affective states leads to a higher capacity to cognitively understand what the other is experiencing. Thus, it could be hypothesized that being able to clearly and consciously access to the other’s emotion facilitates the cognitive nuance of empathy, namely mentalizing. In fact, this latter derives from a further processing of the emotional context, which is constructed through multisensorial integration of diverse sources of stimuli (contextual, visual, or even imaginative) and a higher stimulus awareness. In other words, cognitive empathy could be thought to derive from a more elaborated processing of the emotion, which could involve a deeper conscious perception of it. In addition, as it has been seen in the second chapter, general research in psychology has suggested that there are different lower-level mechanisms which are fundamental for the development of other higher-level ones. In this specific context, it is known that one of the processes at the basis of empathy is the recognition of the emotion of the other person (Ward, 2017). In fact, it has been assumed that both recognition of the other emotion and empathic processes have a common basis: simulation mechanisms (Gallese, 2001; Goldman, 2006; Clark & Kiverstein, 2008; Preston & de Waal, 2002). According to the simulation theory, the recognition of an emotional facial expression and the attribution of a mental state to it necessitate to simulate or reproduce in first person the other' affective state. In addition, developmental and psychodynamic theories are in line with cognitive neuroscience ones, since it is known that our own and others' emotion recognition represents the foundation of other more complex abilities: the child has to learn, through the caregiver intersubjective ability, to identify his/her and others' affective states and, then, this achievement enables he/she to develop the capacity to understand and share the others' emotion (Saarni, Campos, Camras, & Witherington, 2006). More broadly, Saarni et al. (2006) has expressed this concept through the definition of emotional competency, which represents a series of progressively developed abilities, and it defines the adaptive nature of emotions. In fact, as it has also been studied by Barrett (2017), emotions are constructed in the service of allostasis and interoception, which are the basis for the organism's growth, survival, and reproduction. Thus, emotional competency unfolds gradually during the child development (Saarni et al. 2006): first, it encompasses the ability to identify and differentiate our own emotions, then to recognize the other's affective states (for example, through emotional facial expressions), afterwards to acquire an emotional lexicon, and finally the ability to empathize with the other's emotions and to regulate our own and the other's affective states.

Going further, it has been stated that also bottom-up processes intervene in the construction of emotions (Barrett, 2017), since they shape the so called "prediction error" (de Lange, et al., 2018): it represents unanticipated information and, contrary to predictions, it is accounted by a feedforward sweep which concludes in agranular and less developed structures of the cortex (Barrett, 2017). On the other hand, likewise predictions, it is implemented in the service of allostasis and interoception. The relevant aspect for this account of our results is that, only when top-down predictions and bottom-up prediction errors are integrated, the emotion is experienced consciously as an affect. In addition, this integration is recurrent and bi-directional: also in the case of the theories of consciousness, we have found that similar mechanisms are involved. In fact, Northoff and Lamme (2020) have suggested that an "inside-out" and recurrent movement (from the brain to the outer world and vice versa) allows for assigning a phenomenal nature to external stimuli. This hypothesis can, in turn, be reconnected to Evan Thompson's (2001) theory of consciousness and intersubjectivity: the author has suggested that consciousness is formed in the dynamic interrelation of the Self and the Other and, this intersubjective encounter, involves empathy. In other words, Thompson has proposed that empathic and embodied processes are at the basis of consciousness and of the mind. In fact, he has conceived cognition to be characterized by three aspects: embodiment, emergence, and self-other co-determination.

This means that the mind is embodied in the whole organism embedded in its environment, that it emerges in the interrelation of top-down and bottom-up processes, and from the dynamic co-determination of self and other. As it can be noted, the interaction between top-down and bottom-up processes is considered to come into action in the shaping of consciousness by various authors. Moreover, one of the central aspects of Thompson's theory (2001) concerns the form of intersubjectivity than enables consciousness: it is referred to as "open intersubjectivity", which is defined as an a priori structure to perceive the Other as an embodied subject, and its actual expression is empathy. In fact, it has been hypothesized that one's consciousness of oneself as an embodied individual placed in the world depends on empathy, in particular on the experience of oneself as an Other for the Alter-Ego. As it can be noted, the nuance of empathy that is suggested to be at the basis of consciousness is a more embodied, automatic, and unconscious one, and it comes before the more developed capacity of imaginative self-transposal. Thus, this form of intersubjectivity is more inherently human and it could be hypothesized to have a more "bottom-up" nature, since it does not derive from development, conscious reflection, or a further elaboration of the situation. To conclude, according to Thompson (2001) the relationship between empathy and consciousness, defined as the "core dyad", has two sides. The first one is relevant to this study, and it introduces the term "reciprocal empathy", which has been considered as the condition of possibility of consciousness: that is, viewing the Self and the Other as concretely co-determined. Reciprocal empathy is an a priori structure that makes the ego intrinsically structured to involve alterity. In line with this view, the embodied mind, which is the object of study of cognitive science, can be defined through reciprocal empathy and on the basis of the lived body (Merleau-Ponty, 1962).

In order to account for the last behavioural result, it is important to ponder on the nature of the supposed empathic basis of consciousness. Open intersubjectivity has been defined as a more automatic, embodied, implicit and a priori experience: it represents our empathic understand of the Other empathic grasp of oneself (Thompson, 2001). In addition to this component, "concrete intersubjectivity" is also necessary to establish our embodied experience of other subjects: that is, the concrete bodily experience of the other person. The fourth result of the present research regards the absence of a significant association between the QCAE responses and the PAS/Empathy ratings: the Questionnaire answers do not predict the PAS distribution and the empathy subjective ratings expressed during the experimental session phase. This outcome can bring out different suggestions that could converge on one possible account of the seeming contradictions of the results. Why does the Empathy subjective rating is associated with the PAS and the QCAE Empathy questionnaire responses are not? The current discovery could be explained by two main possible and integrable reasons. As it has been suggested before, the kind of empathy that is possibly involved in the foundation of consciousness is a more automatic and embodied one. It can be supposed that the Empathy subjective rating administered during the experimental session phase is closer to this definition than the QCAE questionnaire. In fact, the first one can be seen as a situational rating, that has to be answered quite faster and depending on the immediate grasping of the contextual situation associated with the painful facial expression. On the other hand, the QCAE is not directly related to the sentences and to the face stimuli of the experimental session phase. Moreover, it is administered in a separate session with no time pressure related to the necessity to complete all the phases of the experimental paradigm: this could have given the impression that it can be answered slower and with more reflection. To support this view, it can be noted that the QCAE responses and the Empathy subjective rating have not been found to correlate. Thus, they could be considered as two different measures and nuances of Empathy, at least in the way they have been included in the present study. In the context of the actual research, the empathy subjective rating could be considered as a more automatic, situational, and embodied measure of empathy, whereas the QCAE answers could be accounted as a more reflective, explicit and cognitive one (even if the questionnaire included, at a structural level, both affective and cognitive nuances). In addition, the participants were not aware of the content of the Questionnaire, since they have been informed that it was a general personality test. In other words, when empathy was part of a wider and more comprehensive paradigm and situation, it has been found to be related to conscious emotional perception. When empathy was measured in a "blind" way, as a separate factor, and not directly related to the painful context (sentence and facial expression), it has not been found to be associated to the level of awareness of the emotion. The authors of the questionnaire (Reniers et al., 2011) have affirmed that the QCAE aims to capture separately *individual* variations in cognitive perspective taking *tendencies* as well as differences in the types of emotional reactions *typically* experienced. Thus, and as separate from the experimental session, the answers to the QCAE could be viewed as an expression of individual *trait* empathy or, in other words, of the usual and typical *tendency* of the subjects to be empathetic in everyday life situations. So, only situational empathy seems to influence the conscious perception of the face in that very moment: as we have seen in the theoretical part of the thesis, there are different levels of consciousness (Dehaene et al., 2006) and empathy (Thompson 2001). In this case, the PAS (Overgaard et al., 2006) is measuring an explicit subjective situational level of the stimulus' awareness. Nonetheless, according to this research's authors, the idea that the conscious emotional perception influences empathic responses is more plausible and convincing than the opposite situation.

To conclude the discussion of the behavioural results, it has to be reminded that these implications are only theoretical and hypothetical suggestions but cannot be demonstrated with statistical and objective methods.

### 4.2.2. Discussion of the psychophysiological results

*Zygomaticus Major (ZM) and the Corrugator Supercilii (CS) activities*

The last result that has been found in the present research concerns the psychophysiological analysis. The aim of this latter has been to test a possible difference in the activity of two facial muscles between the painful emotional context and the neutral one. The two muscles whose activity has been registered are the Zygomaticus Major (ZM) and the Corrugator Supercilii (CS). The first one signals a positive affect, since it raises the lip, which usually results in a smile (Massaccesi et al., 2021). On the other hand, the second one is associated with negative affect, because it knits the brow into a frown, that is a sign of pain. The method of measurement is called Electromyography (EMG), and it has been recorded during the whole experimental phase. Afterwards, for the analysis, the signal has been extracted during the 4 seconds' sentence presentation, since the sentence is the operationalization of the contextual information. The purpose of this psychophysiological investigation has been to test if the facial muscles' activity represents a valid mapping of the differential affectivity of the context. In other words, we wanted to see if mimicry reflects the sentences' meaning at an emotional level: could it be considered a good measure to distinguish different linguistic and affective contexts? We have found that the CS activity is higher in the painful emotional context than in the neutral one, whereas the ZM activity does not differ between the two contexts. In particular, the activation of the muscles has been analyzed both during the whole 4 seconds' sentence's presentation and during the last 2 seconds. In fact, it has been observed that the activity of the CS muscle decreases in both contexts during the 4 epochs of the sentence presentation, but from the last 2 seconds it clearly differentiates: in the neutral context the reduction is more pronounced than in the emotional context, falling below the baseline level of activity and indicating a muscle's relaxation; otherwise, in the painful context the CS activity stays above the baseline level, signaling a muscle's contraction. Thus, we have decided to look at the last 2 seconds' sentence's presentation, and the difference of the CS activity between the two contexts has been confirmed. Finally, it can be stated that the CS facial muscle activity seems to be a good measure of the differences between and emotional painful and a neutral context. In particular, it is associated to painful contextual information. As it has been seen in the theoretical part of the present thesis, the facial mimicry signals simulation and embodiment processes linked to emotion recognition and empathy (Dimeberg et al., 2000). So, it can be supposed that a painful emotional context could facilitate sensori-motor simulation and embodiment mechanisms, with respect to a neutral context. As a whole, embodiment and sensori-motor simulation can be conceptualized like the body-related processes underlying our mind functioning and our experience of reality (Thompson, 2001). They support empathy, consciousness, the construction of emotions and any other perception of the world (Barrett, 2017; Azzalini et al., 2019). In fact, simulations and predictions are considered as embodied concepts that have the function of predictively regulating the body's internal environment (i.e., in the service of allostasis; Barrett, 2017). In addition, predictions are used as top-down representations that, with the interaction of bottom-up mechanisms, lead to the construction of perceptions (like emotions). As we have discussed for the previous results, it is unknown if an higher conscious perception of the painful facial expressions boosts empathic processes or vice versa, but it has been discovered an association between the two factors. Thus, it could be hypothesized that a painful context, mediated by mimicry and embodied empathy's processes, could increase the individual awareness of the painful faces (measured through the PAS scale). In fact, the context alone has not been found to modulate the PAS responses; on the other hand, there has been found a relationship between the empathy subjective rating and the PAS. So, if we take into consideration that empathy is conceptualized as an embodied and sensorimotor simulation of the other's affective experience (also mediated by the facial mimicry activity), it could be interesting to investigate if an higher CS activity could lead to higher PAS responses. This would be a possible experimental method to investigate the theoretical suggestion presented by Thompson (2001): that is, intersubjectivity and empathy are the foundation for consciousness.

## 4.3. LIMITS AND FUTURE DIRECTIONS

The aim of the present research has been to try to shape two main processes that are difficult to conceive in an objective and experimental way, namely consciousness and empathy. Thus, the first limit of this study (and similar ones) is the gap between theoretical conceptualizations and inner psychological processes on one hand, and the methods of measurements and experimental paradigms that can be implemented to investigate them on the other hand.

The first aspect of this discrepancy regards the calibration phase, that is the method of measurement of the threshold stimuli used in the experimental session phase. The absence of an influence of the contextual information on the PAS responses, that tended to be concentrated on low values for both neutral and painful sentences, could also be accounted by the fact that the resulting faces of the calibration phase were not so visible to the participants. In fact, the subjects have reported that they were not sure to have dealt with the calibration phase in the right way, and they have complained that the stimuli in the experimental session phase were very difficult to perceive. In future studies, it could be useful to modify the calibration method and its preparatory explanation in a way that the stimuli are actually seen clearly in 50% of the trials.

The second aspect of difficulty that we have encountered regards the definition of empathy and the implementation of a valid and reliable way to measure the construct. As we have seen in the theoretical part, it is known that there are different nuances of empathy: thus, if we want to measure the association between empathy and consciousness, it is necessary to define better what nuance of the two concepts we want to relate. In fact, in the present research has emerged a relationship between the empathy subjective rating and the PAS, but not between the QCAE responses and the PAS. When we drive conclusions and discuss the results, it is important to define what we are measuring in a more precise way. Thus, in future research it will be relevant to conceptualize in advance, before the analysis of the results, the different nuances of consciousness and empathy expressed by the methods of measurements that we are using.

A third more general consideration regards the number of participants. The final sample was composed by 39 participants, since the first 9 were excluded due to a modification of the calibration phase. Nevertheless, in order to acquire more precise results and a higher statistical power, in future studies it would be important to increase the sample dimensions. In fact, it has emerged that individual differences in the affective impact of the painful sentences have influenced the participants in the PAS and empathy subjective rating's judgements. With a higher number of subjects, these differences could be better averaged.

Another aspect that is important to highlight is that contextual information has been operationalized as written sentences, without a prosodic tone. In future studies, it will be relevant to investigate if prosodic expressions of pain, that include both semantic and auditory characteristics, could influence PAS responses. It could be that the absence of an effect of painful sentences on the conscious perception of painful facial expression is also due to the low affective power of reading sentences, compared to hearing meaningful expressions of pain. Affective painful prosody has been used in a previous study (Sessa et al., 2018): it has been found that this factor, interacting with the semantic meaning of the utterance, enhances the P3 ERP component, with respect to utterances pronounced with neutral prosody. In this case, the focus of the study was to investigate brain responses to others' pain, and not also behavioural measures of conscious perception. In future research, it would be important to relate this last aspect to painful and prosodic sentences, since it has been seen that multisensory integration is important in shaping our conscious perception of reality.

Finally, two main prospective directions could be taken into account with respect to psychophysiological correlates of emotions and empathic processes. The first one, that will be explored in our future studies, regards the measurement of skin conductance (SC) responses to painful and neutral sentences. The aim will be to observe if this anticipatory somatic feedback is a reliable modulator of the emotional impact of the painful sentence, since SC is a an automatic activation and indicates the level of the arousal response. We expect the SC to be higher during the presentation of emotional painful sentences than during the reading of neutral ones. The second aspect to examine in future studies would be to see if higher CS responses (i.e. embodiment and simulation processes) are associated with higher PAS responses, since we have seen that these facial muscles' activity modulates the affective meaning of the contextual information that has been presented.

Agresti A. (1989). Tutorial on modeling ordered categorical response data. *Psychological bulletin*, *105*(2), 290–301.

Agresti, A. (1999), Modelling ordered categorical data: recent advances and future challenges. Statist. Med., 18: 2191-2207.

Ainsworth, M. D. S. (1969). Maternal sensitivity scales. *Power*, *6*, 1379-1388.

Ainsworth, M. D. S. (1969a). Individual Differences in Strange-Situational Behaviour of One-Year-Olds.

Aitchison, L. and Lengyel, M. (2017) With or without you: predictive coding and Bayesian inference in the brain. *Curr. Opin. Neurobiol.* 46, 219–227

Alilović, J., Slagter, H. A., & van Gaal, S. (2021). Subjective visibility report is facilitated by conscious predictions only. Consciousness and cognition, 87, 103048.

Ammaniti, M., & Gallese, V. (2014). *The birth of intersubjectivity: Psychodynamics, neurobiology, and the self*. WW Norton & Company.

Aru, J., Bachmann, T., Singer, W., Melloni, L., 2012. Distilling the neural correlates ofconsciousness. Neurosci. Biobehav. Rev. 36, 737–746.

Atasoy, S., Roseman, L., Kaelen, M., Kringelbach, M.L., Deco, G., Carhart-Harris, R.L., 2017. Connectome-harmonic decomposition of human brain activity reveals dynamical repertoire re-organization under LSD. Sci. Rep. 7, 1–18.

Azzalini, D., Rebollo, I., Tallon-Baudry, C., 2019. Visceral signals shape brain dynamics and cognition. Trends Cogn. Sci. 23, 488–509.

Baars, B.J., Franklin, S., Ramsoy, T.Z., 2013. Global workspace dynamics: cortical “binding and propagation” enables conscious contents. Front. Psychol. 2013, 200.

Baron-Cohen, S. (1999). *The evolution of a theory of mind* (pp. 261-277). na.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders*, *34*(2), 163-175.

Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, *12*(1), 1-23.

Barsalou, L. W. (2008). Grounded cognition. *Annual review of psychology*, *59*(1), 617-645.

Beesley, P., & Tyrell, J. (2012). Transitional fields: Empathy and Affinity. *All Our Relations. Eds. Gerald McMaster and Catherine de Zegher. Sydney: The 18th Biennale of Sydney*, 379-381.

Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods*, *2*(2), 131.

Bion, W. R. (1962). The psycho-analytic study of thinking. *International journal of psycho-analysis*, *43*, 306-310.

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. Journal of Neuroscience, 37(40), 9603–9613.

Bora, D. J., & Dasgupta, R. (2020). Estimation of skin impedance models with experimental data and a proposed model for human skin impedance. *IET Systems Biology*, *14*(5), 230-240.

Bowlby, J. (1969). Attachment and loss v. 3 (Vol. 1).

Brodski, A., Paasch, G. F., Helbling, S., & Wibral, M. (2015). The faces of predictive coding. *Journal of Neuroscience, 35*(24), 8997–9006.

Brown, R., Lau, H., LeDoux, J.E., 2019. Understanding the higher-order approach to consciousness. Trends Cogn. Sci. 23, 754‑768.

Bruner, J. (1995). From joint attention to the meeting of minds: An introduction. *Joint attention: Its origins and role in development*, 1-14.

Buzsáki, G. (2006). *Rhythms of the brain.* Oxford University Press.

Carhart-Harris, R.L., 2018. The entropic brain - revisited. Neuropharmacology 142, 167–178.

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *The Quarterly Journal of Experimental Psychology Section A*, *58*(7), 1173-1197.

Castiello, U., Becchio, C., Zoia, S., Nelini, C., Sartori, L., Blason, L., ... & Gallese, V. (2010). Wired to be social: the ontogeny of human interaction. *PloS one*, *5*(10), e13199.

Chalk, M., Seitz, A. R., & Seriès, P. (2010). Rapidly learned stimulus expectations alter perception of motion. Journal of vision, 10(8), 2.

Chang, R., Baria, A. T., Flounders, M. W., & He, B. J. (2016). Unconsciously elicited perceptual prior. *Neuroscience of Consciousness, 2016*(1)

Cheng, Y., Lin, C. P., Liu, H. L., Hsu, Y. Y., Lim, K. E., Hung, D., & Decety, J. (2007). Expertise modulates the perception of pain in others. *Current Biology*, *17*(19), 1708-1713.

Christensen RHB (2019). "ordinal—Regression Models for Ordinal Data ." R package version 2019.12-10.

Christensen, R. H. B. (2019). ordinal - Regression Models for Ordinal Data. R package version 2019.12-10.

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., Bradley, D.C., Smith, M.A., Kohn, A., Movshon, J.A., Armstrong, K.M., Moore, T., Chang, S.W., Snyder, L.H., Lisberger, S.G., Priebe, N.J., Finn, I.M., Ferster, D., Ryu, S.I., Santhanam, G., Sahani, M., Shenoy, K.V., 2010. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. 13, 369‑378.

Clark, A. (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

Clark-Polner, E., Wager, T.D., Satpute, A.B., Barrett, L.F. (2016). Neural fingerprinting: Meta-analysis, variation and the search for brain-based essences in the science of emotion. In: Barrett, L.F., Lewis, M., Haviland-Jones, J.M., editors. The handbook of emotion, 4th edn. p. 146–165, New York: Guilford.

Cloherty, S. L., Hughes, N. J., Hietanen, M. A., Bhagavatula, P. S., Goodhill, G. J., & Ibbotson, M. R. (2016). Sensory experience modifies feature map relationships in visual cortex. eLife, 5

Consciousness in humans and non-human animals: Recent advances and future directions. Frontiers in Psychology, 4(OCT), 1–20.

Conte, S., Brenna, V., Ricciardelli, P., & Turati, C. (2018). The nature and emotional valence of a prime influences the processing of emotional faces in adults and children. *International Journal of Behavioral Development*, *42*(6), 554-562.

Craig, A.D. (2015). How Do You Feel?: an Interoceptive Moment with Your Neurobiological Self. Princeton: Princeton University Press.

Damasio, A. (1994). Descartes' error: Emotion, rationality and the human brain. *New York: Putnam*, *352*.

Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Houghton Mifflin Harcourt.

Damasio, A., 2010. Self comes to mind. Constructing the Conscious Brain, Self Comes to Mind: Constructing the Conscious Brain. Pantheon/Random House, New York, NY, US.

Damasio, A.R. (1999). The Feeling of What Happens: Body and Emotion in the Making of Consciousness. Houghton: Houghton Mifflin Harcourt.

D'Andola, M., Rebollo, B., Casali, A.G., Weinert, J.F., Pigorini, A., Villa, R., Massimini, M., Sanchez-Vives, M.V., 2018. Bistability, causality, and complexity in cortical networks: an in vitro perturbational study. Cereb. Cortex 28 (7), 2233‑2242.

Davis, M. H. (1980). Interpersonal reactivity index.

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception?. Trends in cognitive sciences, 22(9), 764–779.

De Vignemont, F., & Jacob, P. (2012). What is it like to feel another's pain?. *Philosophy of science*, *79*(2), 295-316.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and cognitive neuroscience reviews*, *3*(2), 71-100.

Dehaene, S. and Changeux, J.P. (2005) Ongoing spontaneous activity controls access to consciousness: a neuronal model for inattentional blindness. PLoS Biol. 3, e141

Dehaene, S., & Changeux, J. P. (2011). Experimental and theoretical approaches to conscious processing. Neuron, 70(2), 200–227.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing; A testable taxonomy. Trends in Cognitive Sciences, 10 (5), 204–211.

Dehaene, S., Changeux, J.P., Naccache, L., 2011. The global neuronal workspace model of conscious access: from neuronal architectures to clinical applications. Res. Perspectives in Neurosciences.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the national Academy of Sciences*, *95*(24), 14529-14534.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21.

Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. Frontiers in Human Neuroscience, 5 (December), 166

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological science*, *11*(1), 86-89.

Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13763–13768.

Eliades, S.J. and Wang, X. (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. Psychological Methods, 12(2), 121–138.

Engel, A.K., Singer, W., 2001. Temporal binding and the neural correlates of sensory awareness. Trends Cogn. Sci. (Regul. Ed.) 5, 16–25.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in human neuroscience*, *4*, 215.

Fingelkurts, A.A., & Neves, C.F., 2010. Natural world physical, brain operational, and mind phenomenal space-time. Phys. Life Rev. 7 (2), 195–249.

Finlay, B.L., Uchiyama, R. (2015). Developmental mechanisms channeling cortical evolution. Trends in Neurosciences, 38(2), 69–76.

Freud, S. (1940) An Outline of Psychoanalysis, W.W. Norton

Fridlund AJ, Cacioppo JT. Guidelines for Human Electromyographic Research. Psychophysiology. 1986;23:567-589.

Friston, K. (2005) A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 360, 815–836

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences, 1211–1221.

Gallagher, S. (1997). Mutual enlightenment: Recent phenomenology in cognitive science. *Journal of Consciousness Studies*, *4*(3), 195-214.

Gallagher, S., & Meltzoff, A. N. (1996). The earliest sense of self and others: Merleau-Ponty and recent developmental studies. *Philosophical psychology*, *9*(2), 211-233.

Gallese, V. (2001). The 'shared manifold' hypothesis. From mirror neurons to empathy. *Journal of consciousness studies*, *8*(5-6), 33-50.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, *2*(12), 493-501.

Gallese, V., & Wojciehowski, H. (2011). How stories make us feel: Toward an embodied narratology. *California Italian Studies*, *2*(1).

Gallese, Vittorio (2011). *Embodied Simulation Theory: Imagination and Narrative. Neuropsychoanalysis, 13(2), 196–200.*

García-Pérez, M. A. (2001). Yes-no staircases with fixed step sizes: psychometric properties and optimal setup. Optometry and Vision Science: Official Publication of the American Academy of Optometry, 78(1), 56–64.

Geangu, E., Benga, O., Stahl, D., & Striano, T. (2011). Individual differences in infants' emotional resonance to a peer in distress: Self–other awareness and emotion regulation. *Social Development*, *20*(3), 450-470.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press on Demand.

Gregory, R.L. (1980) Perceptions as Hypotheses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 290, 181–197

Happé, F. and Frith, U. (2006) The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25

Haxby, J. V., & Gobbini, M. I. (2011). *Distributed neural systems for face perception* (pp. 93-110). The Oxford Handbook of Face Perception.

Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron*, *68*(1), 149-160.

Hermes, D., Miller, K.J., Wandell, B.A., Winawer, J., 2015. Stimulus dependence of gamma oscillations in human visual cortex. Cereb. Cortex 25 (9), 2951–2959.

Hochstein, S. and Ahissar, M. (2002) View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Horton, P. C., Louy, J. W., & Coppolillo, H. P. (1974). Personality disorder and transitional relatedness. *Archives of general Psychiatry*, *30*(5), 618-622.

Huang, Z., Obara, N., Davis, H..(Hap), Pokorny, J., Northoff, G., 2016. The temporal structure of resting-state brain activity in the medial prefrontal cortex predicts selfconsciousness. Neuropsychologia 82, 161–170.

Husserl, E. (1973). *Zur Phänomenologie der Intersubjektivität: Texte aus dem Nachlass, Zweiter Teil, 1921-1928* (Vol. 14). The Hague: Martinus Nijhoff.

Hustvedt, S. (2011). Three emotional stories: reflections on memory, the imagination, narrative, and the self. *Neuropsychoanalysis*, *13*(2), 187-196.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *science*, *286*(5449), 2526-2528.

Ibáñez, A., Hurtado, E., Lobos, A., Escobar, J., Trujillo, N., Baez, S., ... & Decety, J. (2011). Subliminal presentation of other faces (but not own face) primes behavioral and evoked cortical processing of empathy for pain. *Brain research*, *1398*, 72-85.

Kagan, J. (1982). Canalization of early psychological development. *Pediatrics*, *70*(3), 474-483.

Koch, C., Massimini, M., Boly, M., Tononi, G., 2016. Neural correlates of consciousness: progress and problems. Nat. Rev. Neurosci. 17, 307‑321.

Lamme, V.A.F., 2006. Towards a true neural stance on consciousness. Trends Cogn. Sci. 10, 494–501.

Lamme, V.A.F., 2010. How neuroscience will change our view on consciousness. Cogn. Neurosci. 1, 204‑220.

Lamme, Victor A.F., Super, H., Spekreijse, H., 1998. Feedforward, horizontal, and feedback processing in the visual cortex. Curr. Opin. Neurobiol. 8, 529–535.

Lavenex, P. and Amaral, D.G. (2000) Hippocampal-neocortical interaction: a hierarchy of associativity. *Hippocampus* 10, 420–430

Lee, T.S. and Mumford, D. (2003) Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 20, 1434–1448

Leek, M. R. (2001). Adaptive procedures in psychophysical research. Perception & Psychophysics, 63(8), 1279–1292.

Litt, C. J. (1986). Theories of transitional object attachment: An overview. *International Journal of Behavioral Development*, *9*(3), 383-399.

Liu, L., Luo, H., 2019. Behavioral oscillation in global/local processing: global alpha oscillations mediate global precedence effect. J. Vis. 19 (May 1 (5)), 12.

Louie, A.. (2010). Robert Rosen's anticipatory systems. Foresight - The journal of future studies strategic thinking and policy. 12. 18-29.

Mack, A. and Rock, I. (1998) Inattentional Blindness, MIT Press

Mancia, M. (2006). Memoria implicita e inconscio precoce non rimosso: loro ruolo nel transfert e nel sogno. *Rivista di psicoanalisi*, *52*(3), 629-655.

Massaccesi, C., Korb, S., Skoluda, N., Nater, U. M., & Silani, G. (2021). Effects of appetitive and aversive motivational states on wanting and liking of interpersonal touch. *neuroscience*, *464*, 12-25.

Merleau-Ponty, M. (1962). *Phenomenology of perception* (Vol. 22). London.

Merletti, R., Botter, A., Troiano, A., Merlo, E., & Minetto, M. A. (2009). Technology and instrumentation for detection and conditioning of the surface electromyographic signal: state of the art. *Clinical biomechanics*, *24*(2), 122-134.

Metcalf, D. R., & Spitz, R. (1978). The transitional object: Critical developmental period and organizer of the psyche. *Between reality and fantasy: Transitional objects and phenomena*, 97-108.

Montague, P. R., Dolan, R. J., Friston, K. J., & Dayan, P. (2012). Computational psychiatry. Trends in cognitive sciences, 16(1), 72–80.

Moutoussis, K. and Zeki, S. (2002) The relationship between cortical activation and perception investigated with invisible stimuli. Proc. Natl. Acad. Sci. U. S. A. 99, 9527–9532

Mumford, D. (1992) On the computational architecture of the neocortex. *Biol. Cybern.* 66, 241–251

Niedenthal, P. M. (2007). Embodying emotion. *science*, *316*(5827), 1002-1005.

Northoff, G., & Lamme, V. (2020). Neural signs and mechanisms of consciousness: Is there a potential convergence of theories of consciousness in sight?. Neuroscience and biobehavioral reviews, 118, 568–587.

Northoff, G., 2013. What the brain's intrinsic activity can tell us about consciousness? Atri-dimensional view. Neurosci. Biobehav. Rev. 37 (May (4)), 726‑738.

Northoff, G., Huang, Z., 2017. How do the brain's time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC). Neurosci. Biobehav. Rev. 80, 630–645.

Northoff, G., Wainio-Theberge, S., Evers, K., 2019. Is temporo-spatial dynamics the "common currency" of brain and mind? In Quest of "Spatiotemporal Neuroscience" Phys. Life Rev.

Overgaard, M., Rote, J., Mouridsen, K., & Ramsøy, T. Z. (2006). Is conscious perception gradual or dichotomous? A comparison of report methodologies during avisual task. Consciousness and Cognition, 15(4), 700–708.

Park, H.-D., Blanke, O., 2019. Coupling inner and outer body for self-consciousness. Trends Cogn. Sci. 23, 377‑388.

Park, H.D., Tallon-Baudry, C., 2014. The neural subjective frame: from bodily signals to perceptual consciousness. Philos. Trans. Biol. Sci.

Pastore, M. (2015). Analisi dei dati in psicologia. *Il mulino*.

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science (New York, N.Y.)*, *357*(6351), 596–600.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramachandran, V. S. (2000). Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution.

Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of personality assessment*, *93*(1), 84-95.

Rigotti, M., Barak, O., Warden, M. R., Wang, X. J., Daw, N. D., Miller, E. K., & Fusi, S. (2013). The importance of mixed selectivity in complex cognitive tasks. *Nature*, *497*(7451), 585–590.

Rizzolatti, G., & Arbib, M. A. (1998). Language

Rosch, E., Varela, F., & Thompson, E. (1991). The embodied mind. Cognitive science and human experience.

Rosen, R. (2012). Anticipatory systems. In *Anticipatory systems* (pp. 313-370). Springer, New York, NY.

Saarni, C., Campos, J. J., Camras, L. A., & Witherington, D. (2006). Emotional development: Action, communication, and understanding.

Sameroff, A. (1975). Transactional models in early social relations. *Human development*, *18*(1-2), 65-79.

Sanchez-Vives, M.V., Barbero-Castillo, A., Perez-Zabalza, M., Reig, R., 2020. GABA*B receptor-modulation of thalamocortical dynamics and synaptic plasticity*. Neuroscience. 17

Sandberg, K., Timmermans, B., Overgaard, M., & Cleeremans, A. (2010). Measuring consciousness: Is one measure better than the other? Consciousness and Cognition,19(4), 1069–1078.

Seibt, B., Mühlberger, A., Likowski, K. U., & Weyers, P. (2015). Facial mimicry in its social setting. *Frontiers in psychology*, *6*, 1122.

Semi, Antonio Alberto, Tecnica del colloquio. Milano: Cortina, 1985

Seriès, P. and Seitz, A.R. (2013) Learning what to expect (in visual perception). *Front. Hum. Neurosci.* 7, 668

Sessa, P., Meconi, F., & Han, S. (2014). Double dissociation of neural responses supporting perceptual and cognitive components of social cognition: Evidence from processing of others' pain. *Scientific reports*, *4*(1), 1-8.

Sessa, P., Schiano Lomoriello, A., & Luria, R. (2018). Neural measures of the causal role of observers' facial mimicry on visual working memory for facial expressions. *Social cognitive and affective neuroscience*, *13*(12), 1281-1291.

Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: a double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, *132*(3), 617-627.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*(7075), 466-469.

Singer, T., Seymour, B., O'doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*(5661), 1157-1162.

Solms, M., 2019. The hard problem of consciousness and the free energy principle. Front. Psychol. 9.

Stein, E. (1964), *On the Problem of Empathy*, trans. Waltraut Stein (The Hague: Martinus Nijhoff).within our grasp. *Trends in neurosciences*, *21*(5), 188-194.

Sterling, P., & Laughlin, S. (2015). *Principles of neural design*. MIT press.

Stueber, K.. 2008. "Empathy." In *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. http://plato.stanford.edu/entries/empathy.

Summerfield, C. and De Lange, F.P. (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. Nature Reviews Neuroscience, 15(11), 745–756.

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2006). Mistaking a house for a face: Neural correlates of misperception in healthy humans. Cerebral Cortex, 16(4), 500–508.

Tambelli R. (a cura di), Manuale di psicopatologia dello sviluppo. Bologna: Il Mulino, 2017.

Tanabe, S., Huang, Z., Zhang, J., Chen, Y., Fogel, S., Doyon, J., Wu, J., Xu, J., Zhang, J., Qin, P., Wu, X., Mao, Y., Mashour, G.A., Hudetz, A.G., Northoff, G., 2020. Altered global brain signal during physiologic, pharmacologic, and pathologic states of unconsciousness in humans and rats. Anesthesiology. 132 (June (6)), 1392–1406.

Thompson, E. (2001). Empathy and consciousness. *Journal of consciousness studies*, *8*(5-6), 1-32.

Tononi, G., Boly, M., Massimini, M., Koch, C., 2016. Integrated information theory: from consciousness to its physical substrate. Nat. Rev. Neurosci. 17, 450–461.

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(1), 3-48.

VandenBos, G. R. (2007). *APA dictionary of psychology*. American Psychological Association.

Ward. (2017). *The student's guide to social neuroscience* (2. ed). Routledge.

Waytz, A., Zaki, J., & Mitchell, J. P. (2012). Response of dorsomedial prefrontal cortex predicts altruistic behavior. *Journal of Neuroscience*, *32*(22), 7646-7650.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current Directions in Psychological Science, 7(6), 177–88.

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655-664.

Winnicott, D. W. (1967). 1971. The mirror-role of mother and family in child development, Playing and Reality. London and New York.

Wood, A., Lupyan, G., Sherrin, S., & Niedenthal, P. (2016). Altering sensorimotor feedback disrupts visual discrimination of facial expressions. *Psychonomic Bulletin & Review*, *23*(4), 1150-1156.

Wood, A., Rychlowska, M., Korb, S., & Niedenthal, P. (2016). Fashioning the face: sensorimotor simulation contributes to facial expression recognition. *Trends in cognitive sciences*, *20*(3), 227-240.

Zahavi, D. (1997). Horizontal intentionality and transcendental intersubjectivity. *Tijdschrift voor filosofie*, 304-321.

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: progress, pitfalls and promise. *Nature neuroscience*, *15*(5), 675-680.

Zeki, S. (2003) The disunity of consciousness. Trends Cogn. Sci. 7, 214–218

# Sitography

www.themathworks.com