UNIVERSITÀ
DI PARMA

DIPARTIMENTO DI MEDICINA E CHIRURGIA

CORSO DI LAUREA MAGISTRALE IN PSICOBIOLOGIA E NEUROSCIENZE COGNITIVE

How architecture influences our perception of emotional body postures: Eye-tracker and electroencephalography studies in virtual reality

Relatore:

Chiar.mo Prof. Vittorio Gallese

Controrelatore:

Chiar.mo Dott. Giovanni Vecchiato

Laureanda:

Gaia Maria Galasso

ANNO ACCADEMICO 2020 - 2021

Abstract

Despite the amount of time that we spend within the built environment, how architecture influences our perception of others' affective states remains unexplored. To this aim, in two experiments we investigated the perception of emotional body postures after the experience of different virtual architectures through an adaptation aftereffect paradigm: subjects were projected in a virtual environment making a dynamic experience of a virtual promenade within the scene. Particularly, in experiment 1, we collected subjective scores and behavioural gaze responses, while in experiment 2 the electroencephalographic (EEG) activity of subjects was recorded. The form of the architecture significantly affected the subjective judgment of the avatar's body. Subjective ratings on avatars' bodily arousal were higher in the low arousing architecture. In addition, after the dynamic experience of virtual architecture, we found that subjects spent more time looking at the avatar's body posture when it was presented within the low arousing architecture. Specifically, participants attended more to the avatar's face after experiencing the low arousing architectural form dynamically. In the second experiment, we found two significant effects: the factor Form contributes in the early time window, compatible with a modulation of the P200; the factor Body contributes in the late one, compatible with a modulation of the Late Positive Potential (LPP). Particularly, a higher P200 amplitude emerged over centro-parietal electrodes for avatars presented within low arousing architectures compared to high arousing ones. This difference is maximum around the first second of observation. In addition, high arousing bodies produced an increment of the LPP amplitude compared with the other two conditions, middle and low. Overall, our findings show that architecture influences the avatar's bodily arousal perception. This evidence suggests that perception of the architectural environment can have different consequences for subjects' attention level towards emotionally relevant stimuli reflected

by behavioural gaze and P200 amplitudes. The findings increase our knowledge of how architecture around us influenced our perception; such findings could pave the way for creating human-centred spaces where people are facilitated in interaction with each other, thus making our everyday social life better.

Index

Introduction

The way we relate to other people depends on the emotional states that we perceive from them. Every social interaction happens within a context, especially in built environments, and the context plays a crucial role in the perception of others' affective states to the extent that it modifies and amplifies them. Considering the significant amount of time that human beings spend within enclosed architectural environments, it becomes crucial to understand the built environment's role in perceiving the other's affective states. The primary goal of the present thesis is to investigate the influence of architectural spaces in the perception of emotional body postures. To this aim, in the present chapter, architectural and bodily factors will be firstly analysed individually. I shed light on the emotional representation through body expressions and the neural basis of their perception. Subsequently, the chapter will focus on the role of context in influencing the perception of body expressions, specifically on architectural features. The reported findings allowed us to hypothesise that body postures and architecture could share some affective components and some perceptual mechanisms, possibly involving emotion, attention and sensorimotor networks. To this aim, an adaptation paradigm was designed to show that the arousing state generated by the dynamic experience of a virtual architecture biases the arousal perceived in the body posture of a virtual avatar.

1. Body postures and Emotions

Humans are sensitive to other people's emotional and gestural action signals, and from an early age, they are trained to recognize them (Bayet & Nelson, 2019; Geangu & Vuong, 2020). Social recognition consists of emotion recognition, action understanding, and interaction in general (Uithol & Gallese, 2015). There are many different modalities through which we express and perceive emotional information (de Gelder et al., 2015). People infer the emotional state of others from different information sources and, specifically, through verbal (i.e., voice and tone) and non-verbal components. Non-verbal components represent one of the main information channels in social communication (Mehrabian, 2017), including both facial and bodily expressions. The role of bodies and body expressions is increasingly investigated, as there are several pieces of evidence to demonstrate its importance in social interaction and social recognition (de Gelder, 2009; de Gelder et al., 2015; de Gelder & Hadjikhani, 2006). Anybody could agree that body language plays a fundamental role in the daily life of each one, as Mehrabian and Friar (1969) argued that changes in body posture reflect changes in the emotional states of people. In confirmation of this, Bindemann and colleagues have shown that when observing social interactions in natural scenes people spend 40% of the time looking at the body to identify a person and his mental state (Bindemann et al., 2010).

Body expressions help us understand the subject's state from static and dynamic features. In a static form, body expressions may be decoded by an anatomical-form description such as certain arm configurations and movement directions (Dael et al., 2012, 2013). For example, an upright posture with raised arms easily describes happiness (Coulson, 2004; Geangu & Vuong, 2020), while for sadness the head is often dropped, arms are brought close to the body (Pollux et al., 2019), and recognition is mainly concentrated on the trunk (Bachmann et al., 2020). For static body postures, these specific features are mostly located in the upper part of the body (i.e., arms, torso, and hands) (Geangu & Vuong, 2020). Yet, in a dynamic form, body expressions may be described by how gestures are performed in relation to time and space. For instance, high arousing emotions (e.g., happiness, anger) relates positively to more expansive spatial extension, high velocity of movement, and high energy (Pollick et al., 2001), in contrast with low arousing emotions (e.g., sadness, boredom) which are characterized by slow-paced movement and body closure (Bachmann et al., 2020; Dael et al., 2012). Overall, our skills for emotion recognition from bodily expression occur by bottom-up processing (Bachmann et al., 2020; Neri, 2009) and are enhanced by the following factors: postures, spatial extension, the velocity of movement and movement energy. Another important factor is the salience of perceived stimulus: when bodies are more salient, subjects look at them for more time (Bachmann et al., 2020) due to the salience increases selective visual attention. Hence, the combination of time and spatial factors, postures, and stimuli salience can be considered essential for body expressions recognition.

For many years the role of bodily expressions was underestimated compared to facial expressions. In fact, it has been assumed that body postures do not provide information about specific emotions but only about their intensity (Dael et al., 2013), or gross affect states, while main studies focused their attention on facial expressions considered universal and consistent. Moreover, humans are skilled at recognizing the face's expressions from an early age (Bayet & Nelson, 2019). However, through recent research, it has been found that the perception of bodily expressions is similar to that of facial expression (de Gelder, 2009; Van den Stock et al., 2007) and that, as they do with faces, infants are able to recognize bodily expressions at an early age (Geangu & Vuong, 2020; Heck et al., 2018). Indeed, both processes convey some information to reveal the emotions and intentions of the others (Bianchi-Berthouze et al., 2003; Kana & Travers, 2012), and occur rapidly and automatically (de Gelder & Hadjikhani, 2006). Moreover, some autonomic and physiological responses of recognition are elicited if we consider both faces and bodies (de Gelder, 2006, 2009): for example, the emotional contagion, i.e. the tendency to automatically mimic and synchronize our expressions with those of another person, observed by emotional faces is also present in body perceptions

(Blakemore & Frith, 2005; Tamietto & de Gelder, 2008). Indeed, when subjects observed body expressions with covered faces, their facial responses were consistent with the emotional valence expressed by the stimuli (Tamietto & de Gelder, 2008).

Despite these similarities, there are still some differences between face and body: unlike facial expressions, body representations convey information about others' actions and intentions closely linked to emotions (de Gelder et al., 2010). Furthermore, in certain circumstances, the human body can signal emotional states better than faces do, as in carrying information from a considerable distance (de Gelder, 2009; Enea & Iancu, 2016; Reed et al., 2006) when faces are not visible; or to discriminate between some emotions, e.g., fear and anger (Meeren et al., 2005). In the same way, when someone tries to hide a lie, body language is more difficult to control than facial expressions, so when face and body convey incongruent emotion, body expression is the key to emotion recognition (Van den Stock et al., 2007)

Behavioural studies have shown that the body can transmit both discrete emotion categories and affective dimensions (Coulson, 2004; Kleinsmith et al., 2006).

In this regard, the Basic Emotion Model, the most popular emotion theory (Paul Ekman, 1999), argues that there is a limited amount of universal emotions, or better, definite basic emotions (i.e., fear, disgust, anger, surprise, sadness, happiness) (Ellsworth & Scherer, 2003), that occurs in distinct events of human life. However, the categorization of discrete emotions does not represent all possible affective states. It may not reflect the complexity of the affective state conveyed by the other people because there is significant variability in the category of different emotions that we experience (Lindquist et al., 2012). For example, the disgust that occurs when watching others eating disgusting food involves a different brain state compared with the disgust that occurs when watching surgical operations. Also, discrete emotion categories are not mapped by distinct patterns of brain locales in the human brain (Lindquist et al., 2012). Research that included a large pool of emotions has shown that differences could be explained through some dimensions of interindividual states, such as arousal, valence, potency, attention, orientation, and approach-avoidance (Dael et al., 2013; Scherer, 2009). Among these, arousal and valence dimensions have the most replicable evidence for explaining emotional states, posing, as stated in the circumplex model (Russell, 2003) (Figure 1), as two independent and bipolar axes anchoring a circular structure, ranging from unpleasantness to pleasantness for valence[1] and from inactivated to activated for arousal[2].

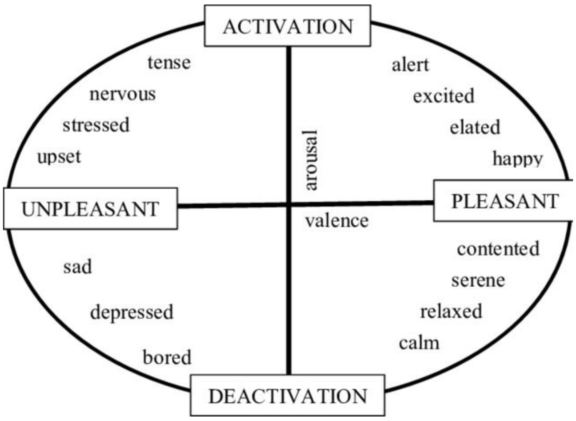

Figure 1|*The circumplex model of affect. The x-axis depicts the valence dimension (from unpleasant to pleasant); the y-axis represented the arousal dimension (from deactivation to activation). Inside the circle, there are some examples of discrete emotions at different arousal and valence levels. (adapted from* Posner et al., 2005).

Each emotional state reflects distinct arousal and valence levels; these two dimensions are also used by human observers when describing affective expressions (Kleinsmith &

---

[1] *Valence refers to the intrinsic pleasantness or goal conduciveness of the emotion-eliciting object or event* (Dael et al., 2013)

[2] *Arousal represents the degree of physiological excitation of the sympathetic nervous system* (Dael et al., 2013)

Bianchi-Berthouze, 2013). In fact, arousal and valence dimensions are considered the core affect and the basic internal representation of emotions (Scherer, 2009). These affective states allow an organism to know if something in the environment has motivational salience (Lindquist et al., 2012) and to have a dynamic view of emotions and display different emotional qualities (Scherer, 2009).

1.1 Neurophysiological bases of body perception

Since many behavioural studies have emphasized the centrality of body postures (e.g. Atkinson et al., 2004; Coulson, 2004; M. E. Kret & de Gelder, 2010), neuroimaging studies have investigated the activation of specific areas during the perception of emotions and in their recognition. The results show that the perceptual processing of bodies is realized by a specialized mechanism based on relations among the features of the stimulus rather than based on the analysis of single body features (Borhani et al., 2015). For example, when we perceive bodies presented upside-down, we have much more difficulty recognizing them than inverted objects. Evidence derived from neuroimaging studies reported that actions and emotions conveyed by postures, parts, or silhouettes of bodies activate specific brain areas (e.g., Borhani et al., 2015; de Gelder et al., 2015; Downing et al., 2001; Kana & Travers, 2012; Peelen & Downing, 2007). Particularly, one of such regions is the extrastriate body area (EBA), located at the posterior interior temporal sulcus (Downing et al., 2001). Moreover, further research has revealed that the fusiform body area (FBA) (Peelen & Downing, 2005), which partially overlaps with the face-sensitive area (FFA), is also involved in the perception of body postures. Therefore, the patterns of neural activity suggest a similarity between faces and bodies (de Gelder, 2006; Peelen & Downing, 2007). These specific areas are both activated during bodies' perception. Generally, these activity brain areas are related to sensory-motor and behavioural information (Peelen & Downing, 2007). This network of brain activation is more strongly activated when observing emotional bodies, rather than neutral ones (Dael et al., 2013).

The perception and recognition of emotional body posture is also regulated by a mechanism known as Embodied Simulation (Gallese, 2009; Rizzolatti & Sinigaglia, 2016). This can be defined as *the recognition that we are human beings whose minds, bodies, environment, and culture are interconnected at sundry levels* (Gallese et al., 2015, pp. 30-31). When we see emotional body postures, the perception might be linked to the simulation of the respective emotion. Our social interaction and perception of bodies become meaningful by means of such a simulation. In this sense, Embodied Simulation is interpreted as a non-conscious, pre-reflective functional mechanism that can take place during our interactions with others, being plastically modulated by context (Gallese, 2017). The neural basis of this process is localized in the Mirror Neurons System (MNS): a set of premotor neurons, initially discovered in the premotor cortex of the macaque monkey (F5) (Gallese, 2008), that fire both when an action is executed and when it is observed being performed by someone else (Gallese, 2009). Subsequently, neurons similar to MNS were identified in a sector of the posterior parietal cortex reciprocally connected with area F5 (Fogassi & Gallese, 2002; Gallese, 2008). After discovering MNS in monkeys several studies were conducted on humans through non-invasive methods such as transcranial magnetic stimulation (TMS), functional magnetic resonance imaging (fMRI), and electroencephalography (EEG), demonstrating the presence of a network of cortical areas activated by action observation in humans too (Rizzolatti & Craighero, 2004).The MNS appears to be the neural substrate of action understanding through Embodied Simulation (Gallese, 2009).

1.1.1 Electrophysiological studies

Event-related potential studies reported that the electrophysiological correlates of emotional stimuli, like faces and bodies, differentiate themselves based on the processing stadium. The perception of emotional bodies might be divided into three different steps (Luo et al., 2010). The first electrophysiological response allows automatic processing and encodes the physical features (Thierry et al., 2006). In the second stage, the structural encoding occurs, during which emotional and neutral stimuli are distinguished (Borhani et al., 2015). The third stage is responsive to motional and affective significance and emotional discrimination (Farkas et al., 2020).

Particularly, the first step identifies the stimulus and reflects initial spontaneous attentional capture, where the processing is very fast, automatic, and coarse. In fact, Thierry et al. (2006) reported that the latency of early components, i.e. P1 and N1, was significantly modulated by stimulus category: object responses were significantly delayed relative to faces and bodies (Thierry et al., 2006). In the successive elaboration steps, the stimuli are processed in more detail.

During the second stage, attentional resources are preserved and there is a partial sensorial encoding of the cues. For instance, response to human bodies shows a prominent negative deflection at occipito-temporal electrodes, the N170. This potential peak ranges between 150 and 230ms, at the same amplitude as the human faces response component (Meeren et al., 2005; Stekelenburg & de Gelder, 2004). More specifically other studies have found that the perception of the body elicited a specific negative component in an EBA, peaking at 190ms (N190; Borhani et al., 2015; Thierry et al., 2006). The emotional and motion components modulated these specific electrophysiology responses (Borhani et al., 2015). However, different studies reported conflicting results on whether electrophysiological responses are modulated by categorical emotions expressed by facials or body postures. This means that N170 modulation may be produced by individual discrete emotions (happiness, anger, sadness) and by body postures' different arousal levels. In fact, the amplitude of N170 is directly proportional to body posture arousal levels (Borhani et al., 2015). This result is interesting because it shows that body postures are able to convey not only discrete emotions but also affective states: for example, different levels of arousal or valence. In addition, for bodies with implied motion, the amplitude of N170 is higher than the one for static bodies. Possibly, this occurs because, evolutionally, perception of movement is a high adaptation value, enabling observes to extract motion-related information from static images where motion is implied (Verfaillie & Daems, 2002). Indeed, previous works have shown that body postures can convey different affective states, especially when animated with movements derived from emotional walking (Presti et al. 2022, currently under review, Bachmann et al., 2020; Borhani et al., 2015). Moreover, emotional stimuli perception elicits a fronto-central positive peak around 200 ms after the stimulus presentation. This component is usually identified as P200 in target stimulus search paradigms (Luck, 2014, pg. 79), and it has been shown that its amplitude is modified in response to emotional stimuli in a range between 150-275ms (Calvo et al., 2013; Feng et al., 2012; Paulmann et al., 2013). The literature shows that high-arousing pictures elicit a more positive-going ERP beginning at 200ms (Cuthbert et al., 2000). During the attentive processing of emotional stimuli, P200 amplitudes reflect enhanced attention to the cue (Delplanque et al., 2004; Paulmann et al., 2013). This component allows fast information discrimination to facilitate behavioural responses (Gerdes et al., 2013). For instance, Calvo and colleagues (2013) reported that the P200 amplitudes increase with emotional stimuli compared to neutral ones. This response is reported for different types of emotional stimuli, e.g. emotional scenes, faces, emotional prosody (Feng et al., 2012; Paulmann et al., 2013), and also in response to multimodal stimuli (i.e. emotional pictures and sound simultaneously; Gerdes et al., 2013). Enhanced P200 amplitudes to emotional stimuli would suggest that these capture automatic attention resources (Olofsson & Polich, 2007).

The third emotional processing step occurs late (between 300 and 900ms) and involves sustained attention and elaborative processing. In this time window, the typical EEG feature is the late positive potential (LPP) at centro-parietal sites. The LPP reflects facilitated attention to emotional stimuli, which modulates its amplitude according to different arousal levels (Cuthbert et al., 2000; Hajcak & Foti, 2020; MacNamara et al., 2022). For instance, Farkas and colleagues have shown that during body presentation, the LPP amplitude was modulated by the categories of the stimuli (Farkas et al., 2020). Overall, the affect-modulation of ERPs can be identified at both early and later processing stages, during which processing in the second and third stages is sensitive to attentional resource availability while processing in the first stage remains relatively independent.

1.2 The influence of context in the perception of emotional body posture

Besides body features, it is essential to emphasise that affective perceptions are not immune to the context contributing to understanding the observed emotion (Righart & de Gelder, 2008). Most studies investigated body posture perception in an isolated way, but we perceive these as part of the natural environment in daily life. Therefore, the social aspect is essential because different body expressions can be differently perceived depending on the social context. Yet, in real scenes, bodies and context interact (J. Wang et al., 2020).

Context can be encoded spontaneously, without the influence of the subject's consciousness, especially when categorizing emotions (Ngo & Isaacowitz, 2015). In confirmation of this, Aviezer and colleagues (2011) have shown that even if participants were asked to ignore the context or consider it irrelevant, they were still influenced when the incongruent context was present (Aviezer et al., 2011). Thus, visual context is not always meaningfully related to the target's emotion, indeed may be somewhat independent (Ngo & Isaacowitz, 2015). However, most background visual scenes are perceived to provide important contextual information for emotion recognition. The information provided by emotional cues is combined with the context (M. Kret et al., 2013) so that global information can be extracted rapidly. Therefore, context may facilitate social recognition: the more ambiguous the emotional expression, the greater the influence on the subjects' perception of the emotional stimuli (Xu et al., 2017).

Some studies began to investigate the effect of background scenes on emotion processing. For example, it was reported that facial expression recognition is significantly faster in congruent emotional scenes than within incongruent emotional ones (Righart & de Gelder, 2008), in which subjects had the worst performance (Ngo & Isaacowitz, 2015; Righart & de Gelder, 2008). In addition, Righart & de Gelder demonstrated that emotional scenes could affect the processes of facial expression at the early stages: N170 amplitudes were larger for some emotionally congruent conditions than they were for incongruent conditions (Righart & de Gelder, 2008). Xu and colleagues have not replicated such result; finding that a congruent effect could happen at the later stages of face processing. Indeed, there was no early modulation of N170 amplitudes for the congruent scenes, but modulation of the LPP amplitudes was larger for congruent scenes than for incongruent ones (Xu et al., 2017). Previous studies show that LPP is a positive component modulated according to the congruence of stimuli (Calbi et al., 2017; Li, 2021) and appears to be greater in response to emotional rather than neutral stimuli (Choi et al., 2014). It is further noted that LPP amplitudes reflect motivated attention, which is evoked process as approach-avoidance (Xu et al., 2017). For instance, Schupp et al. (2004) showed that angry faces intensify LPP amplitudes more than other facial expressions, indicating that more motivational attention was allocated to the stimuli that conveyed threatening information, which is important to the evolutionary level (Schupp et al., 2004).

Similar context effects have been found for bodies. Kret and colleagues (2013) investigated how body expressions (with visible or invisible faces) are processed when presented in a social-emotional context, either congruent or incongruent. The congruent emotional scenes influenced participants' attention and the last one was not specifically directed toward emotionally incongruent cues. This result was obtained in other studies (e.g., M. E. Kret & de Gelder, 2010), particularly those in which facial expressions were blurred, so that it could only be measured the influence level caused only by bodily expressions. Body expressions recognition was facilitated by a congruent context, particularly in the congruent fear condition (M. E. Kret & de Gelder, 2010). This result is not strange because, evolutionally, fear is an intensive emotion that requires a fast reaction. In addition, the interpretation of a neutral body posture is conditioned by the emotional background information (Van den Stock et al., 2014).

This interaction between body and scene perception was also reported at the brain level: a threatening background scene with the presence of a body compared to a neutral one showed increased activity in bilateral EBA (Van den Stock et al., 2014). This EBA activation occurs only for neutral bodies and not for the emotional body or only emotional scenes (Van den Stock et al., 2014). This means that EBA responds to threat signals displayed by the body as well as by the background scene, probably because the response to the body is biased in the direction of the scene due to perceptual bias (Van den Stock et al., 2014).

2. Architectural Environments

Every social interaction happens within a context, especially in built environments where we spend most of the time (ca. 90% of our life (Coburn et al., 2020)). For such reason, investigating behavioural and neurophysiological responses to body perception in different architectural environments is crucial. It is now common knowledge that systematic variations in architectural features lead to changes in behavioural outcomes (Colin Ellard, 2015) and modifications of mental states (Coburn et al., 2020), both in the short and long term. In addition, architectural environments can change and amplify perception experience (Pasqualini et al., 2018) and shape the affective states (e.g., arousal, valence) (Fich et al., 2014; Presti et al., 2021). Emotion is the pre-reflective response of a human organism to build architectural environments, which is why it is essential to understand how all environments have an inherent emotionality (Gallese et al., 2015, pp.30-31).

Architectures with high ceilings or open spaces were judged as more pleasant (Coburn et al., 2020; Vartanian et al., 2015), while low ceilings and enclosed spaces were associated with higher arousing states (Fich et al., 2014). On average, preferences for ceiling height peak around 3 metres (Coburn et al., 2020), since these increase perceptions of spaciousness.

Additionally, studies with 2D stimuli revealed that enclosed spaces generated a greater fear sensation than open spaces (Vartanian et al., 2015). Behavioural studies in virtual reality have shown that subjects prefer to experience the virtual promenade within wide spaces rather than enclosed ones, linking the latter environments to unpleasantness and high arousing states (Presti et al., 2021). Reducing the surrounding spaces leads to unpleasantness and high arousal states because it's perceived as a constraint (Gallese & D., Ruzzon, 2016; Presti et al., 2021). The size and placing of the windows appear to have relevance to the well-being of the inhabitants: access to outdoor views was found to reduce the stress level (Leather et al., 1998), and the entry of the light into the architecture is essential to promote circadian rhythms (Acosta et al., 2019). Placing the windows closer to the subject height produced a more pleasant experience and lower arousal values (Presti et al., 2021). These results also support the theory that humans prefer environments with a greater affordance of visual prospects (Coburn et al., 2020). The colour features investigated by Yildirim and colleagues show that warm walls are considered more arousing compared with cold walls, while cold environments evoke feelings of spaciousness and calm (Yildirim et al., 2011).

In summary, all this evidence confirms that architectural environments, and variation of physical features, influence our psychophysiological states (Lindal & Hartig, 2013) and produces a variation of the extra-personal space, especially when the subject can dynamically perceive it (Presti et al., 2021).

Therefore, it becomes relevant to understand human cerebral reactions produced by the perception of architecture. Investigating architecture perception with a neuroscientific approach means studying the interaction between the built environment and the central nervous system's activity.

2.1 Architectural perception and Embodied Simulation

Environments shape and amplify the perceptive experience of each of us in a similar way to art perceptions (Gallese et al., 2015). This happens through an unconscious projection of self on the perceived object, and neuroscience explains it by a specialized embodied mirror mechanism. Freedberg and Gallese (2007), hypothesized that when people observe and contemplate artistic images, their brains and bodies react by simulating the same actions and emotions depicted by the artist (Freedberg & Gallese, 2007). This process occurs through the joint activation of sensorimotor, emotional, and hedonic networks (Di Dio & Gallese, 2009). In any aesthetic experience, there are corporeal dimensions (Gallese et al., 2015), and Embodiment theories in architecture emphasize the importance of these dimensions. During the perception of the built environment, we simulate the forms and materials with our bodies (Mallgrave, 2013). However, more than any other art, architecture prompts the simulation of wanting to move within it. Therefore, in architectural perception, the visual and motor systems are co-dependant and play a central role (Djebbara, 2021). Indeed, space perception is based on activation of sensory and motor systems.

For such reason, there is an increasing interest in investigating the perception of architecture as a multi-sensory and embodied experience. In order to do this, in an electrophysiological study, Vecchiato et al. (2015) have shown that there was the suppression of μ-rhythm in rooms judged as pleasant and comfortable, highlighting the involvement of the fronto-parietal network in the perception of places (Vecchiato, Tieri, et al., 2015). Moreover, further EEG results report the activation of the frontal-midline theta over a brain network that includes frontal, orbitofrontal, and left temporal areas, with the activity of these regions particularly increased in the environment where it was perceived as a state of high presence (Vecchiato, Tieri, et al., 2015). These results may reflect positive emotional experiences associated with exploring the virtual environment. The perception of pleasant interiors activated sensorimotor regions, suggesting the involvement of Embodied Simulation mechanisms (Gallese, 2009), regulating the simulation of actions, emotions and corporeal sensations during the art and architecture perceptions (Freedberg & Gallese, 2007). Moreover, the architectural environment's perception could also affect emotional and motivational dimensions. In fact, as the artwork perceptions (Chatterjee & Vartanian, 2014; Lacey et al., 2011), this one activates orbitofrontal and prefrontal brain regions (Vecchiato, Jelic, et al., 2015). This might explain how variations in architectural features lead to behavioural outcomes.

3. Experimental hypothesis: architectural spaces influences the perception of body postures

The primary goal of the present study is to investigate how the perception of emotional body postures changes within different surrounding architectures. Architectural environments elicit a broad range of aesthetic experiences, from feeling comfortable to even shaping our perceptual experience (Coburn et al., 2020; Pasqualini et al., 2018). The hypothesis is that architectures and body postures share some perceptual mechanisms involving sensorimotor brain areas (Bonner & Epstein, 2017). In fact, the same structures involved in our own body experiences and perceptions contribute to understanding the world around them (Robinson & Pallasmaa, 2015). According to the spatial features, the architecture contributes to modulating the perception of body expressions. For such reason, a study exploring the relationship betweenbodies and architectures is essential to reach an exhaustive comprehension in everyday life scenarios.

3.1 The adaptation effect

An adaptation aftereffect paradigm was designed to highlight the link between architectural and body feature perception. In this kind of paradigm, the first stimulus is the adapter that generates a bias in the opposite direction when perceiving the following target stimulus (Gibson, 1933). First adaptation studies were about visual perception; subsequently, studies have focused on face and body features perception (Halovic et al., 2020; Kloth et al., 2010; Kovács et al., 2006; Webster & MacLeod, 2011), . It must be said that this has already been investigated in cross-category or modal stimuli (Kovács et al., 2006; X. Wang et al., 2017). For instance, Wang and colleagues investigated cross-modal adaptation by adapting subjects to the sound of laughter and then judging the face. The results show that hearing laughter biased the subsequent judgment of facial expressions (X. Wang et al., 2017). Another type of adaptation effect is the conceptual aftereffects (Halovic et al., 2020). This phenomenon is generated by some higher concepts shared by both stimuli (adapter and target stimulus) and can produce the adaptation even when the adapting stimulus is different to the target (e.g. face-body aftereffect). Halovic et al. (2020) showed as conceptual aftereffects are tolerant to changes between the emotional adapting and target stimuli (e.g. happiness vs sadness) if the underlying concept is perceivable in both stimuli (Halovic et al., 2020). For these reasons, we hypothesise that there may be such an adaptation effect between architecture and body posture if they can be conceptually associated with different arousing states.

3.2 Virtual reality

In order to conduct our experiment, we decided to use Virtual Reality because it enlarges the methodological toolbox of social and behavioural studies (Bönsch et al., 2018). In fact, a growing number of studies in psychology and neuroscience demonstrate that such difficulties can be surmounted using Immersive Virtual Reality (VR, Bohil et al., 2011; Sanchez-Vives & Slater, 2005). Three-dimensional spaces demand a higher level of surrounding perception and therefore difficult to investigate using flat screens. Nowadays, the advancements in VR technology allow increasing realism and a sense of immersive in experimental scenarios, despite 2D studies (Bohil et al., 2011; Slater et al., 2010; Vecchiato, Tieri, et al., 2015). Virtual environments are under the experimenter's full control, creating the possibility of dissociated stimuli, which can not be isolated in real-life scenarios. VR is an essential tool that allows the evaluation of the emotional perception of people during the progressive variation of architectural features within de-contextualized environments (Presti et al., 2021). This methodology allowed previous studies to focus on the architectural features, e.g., ceiling height, wall width and colour, and windows position, showing that these modifications of the contexts can affect psychological responses. Changes in these characteristics are evaluated with different levels of arousal and valance. Moreover, the perception of architecture is directly related to motion, so it is an unmissable feature of its perception (Presti et al., 2021). Virtual reality is a realistic way better to perceive the sense of movement in the spaces. Also, VR is a tool for studying brain activation during spatial navigation (Sanchez-Vives & Slater, 2005; Vecchiato, Jelic, et al., 2015).

3.3 Experimental studies

We investigated the judgment of emotion from body postures in architectural environments in two different experiments through the adaptation aftereffect paradigm. *Experiment 1* consists of a behavioural study in two conditions, i.e., dynamic and static. In the dynamic condition, participants, after the virtual promenade, judged the level of arousal in an avatar with different body postures. In the static condition, participants did not make any virtual promenade within the space; they were directly placed in a static position in the environment and judged the avatar arousing states. The hypothesis is that the arousing states generated by the dynamic experience of virtual architecture (i.e., the adapting stimuli) will bias in the opposite direction the arousal perceived in the body posture of a virtual avatar (i.e., the target stimulus). Therefore, the arousing state generated during the dynamic experience of a virtual architecture will bias in the opposite direction of the arousal perceived in the body posture of a virtual avatar. Thus, when the walk happens into the high arousal architectural environment, the perception of emotional body postures will result in lower arousal than in empty spaces.

We decided to use affective states, particularly the arousal component, because both the body posture and architecture are characterized and related mostly by the arousal dimension (Coburn et al., 2020; Karg et al., 2010; Kleinsmith et al., 2005, 2011; Presti et al., 2021).

Successively, in *Experiment 2*, we proposed the same paradigm in the dynamic condition, with the addition of electroencephalography (EEG). This has allowed us to analyse the neurophysiological basis of the perception of body postures in different architectural spaces. We investigated which ERPs modifications reflect the adaptation aftereffect to understand if the perceptive bias measured with behavioural responses also causes a change at the neurophysiological level. We expect a different response in P200 and LPP components depending on different levels of perceived arousal. The components of interest are the P200 and LPP because they show a modulation of the amplitude according to the perceived arousal (Farkas et al., 2020; Hietanen et al., 2014; Li, 2021; Stekelenburg & de Gelder, 2004).

EXPERIMENT 1

Behavioural Study

Materials and methods

Participants

Thirty-two participants (27.16 ± 4.16 years, 17 female) were enrolled for this study. Sample size was determined using a power analysis computed through the G*Power 3 software (Faul et al., 2007) considering the 2x2x3 within subject design with the "as in SPSS" option and setting the significance level (α) at 0.05, the desired power ($1- \beta$) to 0.95, the number of groups to 1 and the nonsphericity correction ε to 1. The value of the η_p^2 was set to 0.07 based on previous research showing how emotional perception is influenced by the context (Ngo & Isaacowitz, 2015). All participants were naïve to the purpose of the experiment and had normal, or corrected to normal, vision and no previous history of neurological disorders. The study was approved by the local ethical committee (Comitato Etico AVEN) and conducted according to the principles expressed in the Declaration of Helsinki and each participant provided written informed consent before participating in the experiment.

Stimuli

Architecture. Architectural stimuli were selected from a set of 54 virtual architectures, which were previously tested through a behavioural study (Presti et al., 2021). Specifically, architectures were designed as a combination of three consecutive nuclei where architectural elements such as the side walls distance, the ceiling and windows height could decrease, increase or remain constant between one nucleus and the next one. The height of the windows changed as a percentage of the ceiling height: 75% and 100%

of the ceiling height in the second and third nucleus respectively when increasing, 25% and 0% when decreasing, 50% when constant. Thus, for this study we considered two possible architectural forms corresponding to a high and a low level of arousal. In fact, in the previous study, architectures with decreasing sidewalls distance, increasing windows and ceiling height were found to generate the highest arousing scores on the subject's affective states. Conversely, architectures with a constant ceiling, increasing sidewalls distance and decreasing windows height, generate the lowest arousing effect. Hence, we used these two architectural configurations in a warm and cold colour tone. Moreover, each architecture was presented in two shades of colour, totalizing four architectures.

Panel A of figure 2 illustrates a view of the architectural stimuli from the entrance perspective.

In addition, we used an empty scene as the control condition. Specifically, it was represented by the custom scene of Unity 3D engine Software (2019.1.0f2), which had brown-coloured ground and a blue sky. In this space, there were no reference points. Thus, to give a landmark in such an environment, grey alternate vertical lines were located in the middle of the scene on the ground level.

Panel B of figure 2 shows a view of the empty space.

Emotional Body Posture. We selected male avatar's body postures conveying different levels of arousal from a set of emotional body postures previously tested through an online experiment (Presti et al, 2021). The ratings provided by subjects for each body posture of avatar were averaged and then sorted in a linear way. We distinguished three levels of body postures, conveying low, middle, and high arousal, respectively. Hence, we selected 30 body postures from these categories: the low arousing group included 10 body postures with the lowest score of perceived arousal, contrarily the high arousing group included 10 body postures with the highest score of perceived arousal. Finally, 10 body postures at the centre of the distribution constituted the stimuli of the middle arousing group. The arousal level was significantly different ($F(2,27) = 102.93$ $p < 0.001$) across all the three groups, as revealed by Bonferroni corrected pairwise comparison (low < middle, $p < 0.001$; middle < high, $p < 0.001$; low < high, $p < 0.001$).

Each body posture was presented within all the architectures.

Panel C of figure 2 illustrates an example of three avatars with a low, middle, and high level of arousal, respectively.

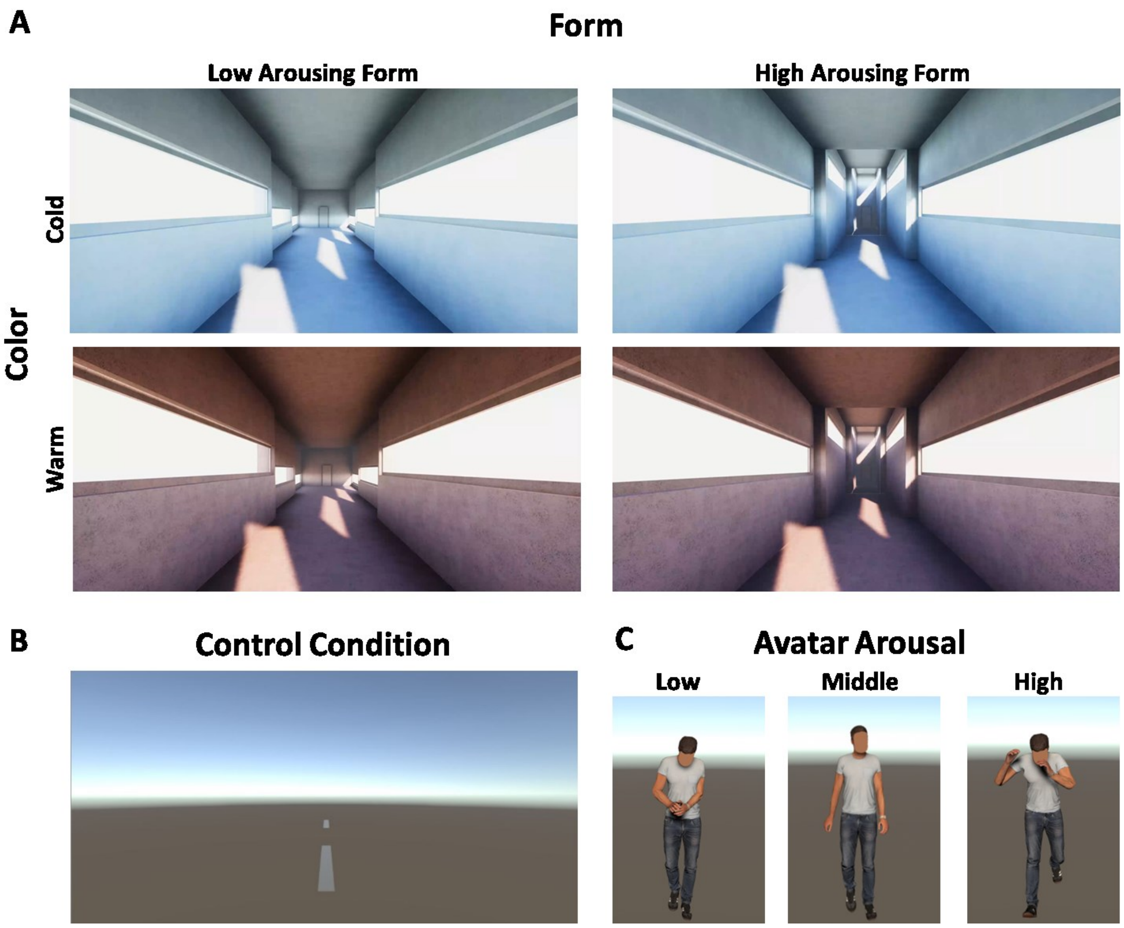

Figure 2 | *Stimuli examples.* ***Panel A*** *shows the architectural stimuli. Columns represent the low and high arousing forms, respectively. Rows shows the two possible colours of the architecture (cold and warm, respectively).* ***Panel B*** *shows the empty control environment.* ***Panel C*** *display examples of three avatar's body postures with low, middle, and high level of arousal, respectively*

Experimental Setup

We performed the experiment in a highly immersive virtual reality environment realized with the HTC Vive Pro Eye head-mounted display (HMD). This device is equipped with two AMOLED screens, with a resolution of 1440 x 1600 pixels per eye, a refresh rate of 90 Hz, and a field of view (FOV) of 110°. Furthermore, the HTC Vive Pro Eye includes the Tobii eye-tracking system, allowing the recording of eye data such as the eye openness and gaze origin and direction with an accuracy of 0.5° - 1.1° (within FOV 20°) https://www.vive.com/uk/product/vive-pro-eye/specs/. Unity was integrated into the

HMD via the Steam VR asset to control the experimental procedure and collect data. Eye data were collected with a sampling frequency of 90 Hz utilizing the SRanipal (v1.3.0.9) plugin. The experiment ran on a laptop equipped with Windows 10 Home (64-bit), Intel Core i7 -9750H, 32 GB RAM, and the NVIDIA GeForce RTX 2070 graphics card.

Experimental Procedure

The experiment consists of dynamic and static conditions. The movement of a first-person perspective camera generating a virtual promenade realized the dynamic condition. Instead, in the static condition subjects did not make any virtual promenade. The presentation order of the dynamic and static conditions of the experiment was randomized among subjects. After reading written instructions, the HMD was positioned on the participant's head and comfortably arranged, providing a clear view of the virtual environment. Afterward, the eye-tracking calibration was administered, which consisted in centring the HMD to the eye level, calibrating the interpupillary distance and finally in focusing the gaze on a central point that moved to four consecutive peripheral positions.

Dynamic Condition. The participants were projected in a virtual environment and experimental trial started with 500ms of static perception of this one, followed by a straight virtual promenade within the scene, lasting 12.5 s. Inside the architectures, the virtual promenade started at the beginning of the first nucleus and ended at the end of the second one. At the end of the virtual promenade, subjects experienced the environment statically for 500ms, and then the avatar appeared for 3s in the middle of the scene. Hence, they were asked to judge the arousal level expressed by the avatar's body posture answering the question: *"this person looks in a ... state"* ranging from "Deactivated" to "Activated" by mean of a visual analog scale (VAS). Subjects used the Vive controller to move the VAS cursor with a time limit of 6 s to provide their judgments. In this condition, we used 5 different scenarios, i.e. four different architectures and the control empty scene, with 30 different body postures for a total of 150 trials divided into 6 blocks. In the first and last one, we presented 15 body postures randomly within the control empty scene. Instead, the 30 body postures were randomly presented within the virtual architectures in each of the central blocks. Blocks were separated by a pause in which subjects were permitted to take the HMD off and have some rest. Subjects performed a new eye tracker calibration at the beginning of each block.

Static Condition. In the static condition, subjects did not make any virtual promenade within the space. Thus, they were directly projected at the end of the second architectural nucleus. At each trial, subjects experienced the space statically for 500ms, and then the avatar was presented in the middle of the scene for 3s. Then, subjects judged the avatar arousing state as in the dynamic condition. We adopted the same subdivision in blocks as for the dynamic condition.

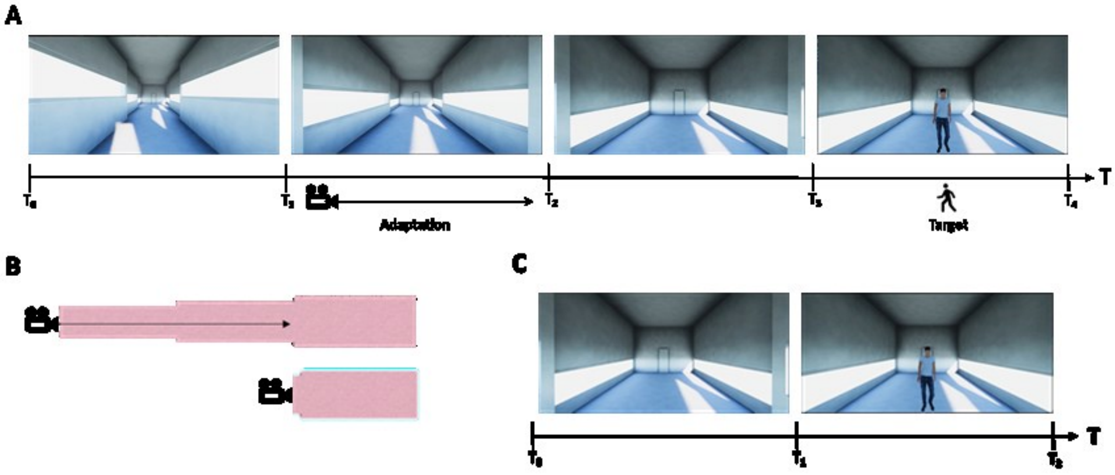

Figure 3| *Schematic representation of experimental trials.* ***Panel A*** *shows the dynamic condition: each experimental trial started with 500ms of static perception of the virtual environment from the entrance perspective (T0-T1), followed by a straight virtual promenade within the first two nuclei, lasting 12.5 s (T1-T2). Afterwards, subjects statically experienced the environment for 500ms (T2-T3) and then the avatar was presented for 3s (T3-T4).* ***Panel B*** *represents the difference between the dynamic (upper panel) and static condition (low panel), showing that in the dynamic condition the virtual camera moved from the entrance to the end of the second nucleus of the architecture. Instead, in the static condition, the camera remained stuck at the end of the second nucleus.* ***Panel C*** *shows the experimental trial of the static condition: subjects were placed at the end of the second nucleus and statically experienced the architecture for 500ms (T0-T1). Afterwards, the avatar was presented for 500ms (T1-T2).*

Data Preprocessing

A trial clean-up of *Experiment 1* was made before statistical analysis, which consisted in discarding trials with dips of attention. To consider a possible attention decrease, for each subject, we computed the blink rate of each trial and marked as bad trials the outliers of the relative distribution. We used peak detection to define a blink, setting the threshold value at 0.6 (1 stands for eye closed and 0 for eye opened) with a 300ms minimum distance between consecutive peaks (Caffier et al., 2003; Haq & Hasan, 2016). This procedure led to the rejection of the 3% of the trials, thus ensuring a blink rate significantly lower after the bad trials rejection (paired-sample t-test ($t(31) = 4.891$, $p < .001$)). No trial was discarded from the static condition of the experiment after following the procedure described above. Indeed, there is no significant difference between the blink rate of the static and dynamic condition after the bad trial rejection (paired-sample t-test (t(31) = -1.777, p = .085)).

To detect for possible outliers, subjective arousal ratings were z-scored, considering the mean and standard deviation of scores provided in all the experimental conditions. Three subjects resulted as outliers in the dynamic version of the experiment and two subjects in the static one. Thus, they were discarded from further analysis.

Arousal Ratings

Subjective ratings were z-scored considering the mean ($\bar{x}_e$) and standard deviation (σ_e) resulting from the scores provided in the empty scene.

$$x_z = \frac{x - \bar{x}_e}{\sigma_e} \qquad (1)$$

Such normalized scores were then analysed via two distinct 2x2x3 (Form x Colour x Body) repeated measures (rm) ANOVA, for both conditions of the experiment.

Gaze analysis

Data from one subject in the dynamic condition were discarded due to technical issues during the eye tracking recording, thus we analysed data from 28 subject in the dynamic condition and from 30 in the static one. The analysis was focused on subject’s gaze behaviour during the presentation of the avatar. Thus, we identified 4 different ROIs representing the face, the trunk, the arms and the legs, respectively (Kleinsmith & Semsar, 2019; Pollux et al., 2019b). For our analysis, we focused on the time duration of subject’s fixations over the defined ROIs. Fixations that lasted less than 200ms were discarded from the analysis (Negi & Mitra, 2020; Salthouse & Ellis, 1980; Salvucci & Goldberg, 2000). We did not consider the number of fixation due to the high positive correlation emerged with the fixation times (Pearson's linear correlation coefficient R = 0.68, $p < 0.001$).

Firstly, for both the dynamic and static condition of the experiment, we analysed the averaged fixation times. We considered the fixation times emerged in the empty control scene as baseline, thus fixation times in the architectures were z-scored considering those of the empty control scene as a baseline. Data were statistically analysed by means of two distinct rm ANOVA which within factors were ROI (face, trunk, arms, legs), Form (low arousing form, high arousing form), Colour (cold, warm) and Body (low, middle, high). Tukey's test was used for post-hoc comparisons among means.

Furthermore, to include in the analysis also important spatio – temporal dynamics (Chaby et al., 2017), we investigated the spatial distribution and time course of fixation. For both the experimental condition, i.e. the dynamic and static one, we computed the time spent looking at each ROI within time slices of 100ms each. Fixation times on avatars presented within the architectures were z-scored with respect of those relative to the empty scene. To graphically illustrate the results, we built 4 timelines (one per ROI), each composed by 30 consecutive time slices, using a graded colour-scale to codify the time spent looking at each ROI. Finally, a non-parametric analysis based on Montecarlo method (5000 iterations) was performed to compare fixation durations over time among different experimental conditions. We adopted a cluster correction method to control for multiple comparisons, considering that our data had a strongly correlated spatio-temporal structure. Finally a meta-permutation was performed, running the permutation test 20 times and averaging the p-values together (Cohen, 2014).

Data analysis was performed with Matlab 2018b (The Mathworks, Inc., Natick, MA, USA) and Statistica 7 (StatSoft Europe) software.

Experimental Results

Arousal ratings. Figure 4 shows the results of the rm ANOVA on arousal ratings that subjects gave in the dynamic and static condition of the experiment. The main effect Body resulted to be significant in both condition of the experiment: dynamic ($F(2,56) = 92.046$, $p < 0.001$, $\eta_p^2 = 0.767$) and static ($F(2,58) = 84.055$, $p < 0.001$, $\eta_p^2 = 0.743$). For the two experimental conditions, Bonferroni corrected pairwise comparisons revealed that arousal ratings were significantly different among the three levels of the avatar's bodily arousal (low < middle, $p < 0.001$; low < high, $p < 0.001$; middle < high, $p < 0.001$). The main effect Form was significant only in the dynamic condition ($F(1,28) = 5.864$, $p = 0.022$, $\eta_p^2 = 0.173$), revealing higher arousal ratings given within the architecture with the low arousing form. No significant effect was found for the factor Form in the static condition of the experiment ($F(1,29) = .469$, $p < 0.499$, $\eta_p^2 = 0.016$). The main factor Colour did not result in any significant effect for neither of the two experimental conditions: dynamic ($F(1,28) = 2.041$, $p = 0.164$, $\eta_p^2 = 0.068$) and static ($F(1,29) = 1.629$, $p = 0.212$, $\eta_p^2 = 0.053$). However, in the static condition, two-way interactions Colour x Form ($F(1,29) = 5.349$, $p = 0.028$, $\eta_p^2 = 0.156$) and Colour x Body ($F(2,58) = 6.025$, $p = 0.004$, $\eta_p^2 = 0.172$) resulted to be significant. Specifically, Bonferroni corrected pairwise comparisons showed that bodies with middle level of arousal received higher arousal ratings within the warm coloured architecture rather than in the cold coloured one (warm > cold, $p = 0.027$). The interaction Form x Body was not significant for both the condition of the experiment: dynamic ($F(2,56) = 1.599$, $p = 0.321$, $\eta_p^2 = 0.039$) and static ($F(2,58) = 2.030$, $p = 0.141$, $\eta_p^2 = 0.065$). However, we compared ratings among the two different forms considering the different arousal level of avatar. For the middle body level paired sample t-tests highlighted a significant difference (high arousing form < low arousing form, t(57) = 2.9, p = 0.005) but no significant effect in the static one (t(59) = 0.018 , p = 0.986). No significant effect emerged in the other two body level of arousal in both condition: low body level (dynamic: high arousing form < low arousing form, t(28) = 0.52, p = 0.602; static: high arousing form < low arousing form, t(29) =1.84, p =0.075); high body level (dynamic: high arousing form < low arousing form, t(28) = 1.04, p = 0.307; static: high arousing form < low arousing form, t(29) = -0.54, p = 0.595).

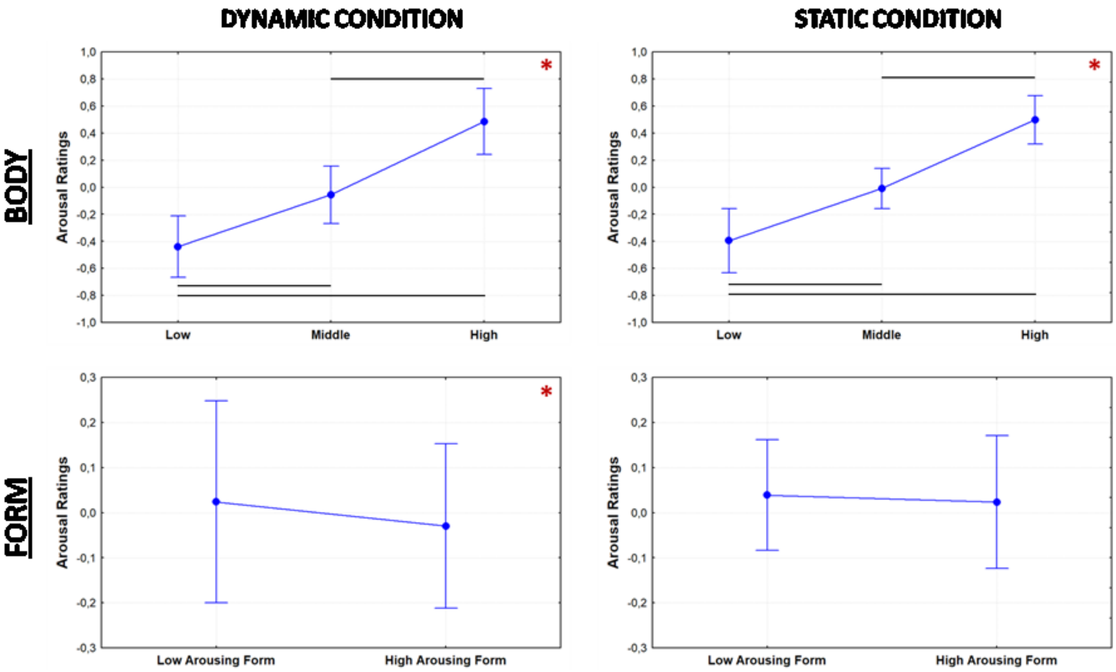

Figure 4* | *Arousal Ratings. *Results of the rm ANOVA on arousal ratings for the dynamic (left column) and static (right column) condition of the experiment. Upper and lower panels show the distribution of the arousal ratings for the main factors Avatar and Form, respectively. Data are presented with their mean and 95% confident interval. Red asterisks indicate significant main effects, while black lines show significant pairwise comparisons.*

Gaze analysis. Results of the two rm ANOVAs on averaged fixations times revealed a significant effect for the main effect Body in the dynamic (F(2,64) = 4.682, p = 0.013, η_p^2 = 0.147) as well as in the static condition (F(2,58) = 3.856, p = 0.027, η_p^2 = 0.117). HSD Tuckey corrected pairwise comparison revealed that subjects spent more time looking at the body of the avatar when it showed a low level of arousal (dynamic condition: low > high, $p = 0.014$; static condition: low > high, $p = 0.020$). The main effect ROI also resulted to be significant, but only in the dynamic condition ($F(3,81) = 6.692$, $p = 0.001$, $\eta_p^2 = 0.176$), not in the static one ($F(3,87) = 1.778$, $p = 0.157$, $\eta_p^2 = 0.058$). Specifically, in the dynamic condition, subjects spent significantly less time looking at the trunk of the avatar compared to the other ROIs (trunk < face, $p = 0.013$; trunk < arms, $p = 0.001$; trunk < legs, $p = 0.005$). We found a significant interaction Body x ROI in the dynamic ($F(6,162) = 35.468$, $p < 0.001$, $\eta_p^2 = 0.567$) as well as in the static ($F(6,174) = 35.880$, $p < 0.001$, $\eta_p^2 = 0.553$) condition. Specifically, HSD Tuckey corrected pairwise comparison revealed that subjects spent more time looking at the trunk when the body had a middle level of arousal rather than a low (dynamic and static condition: $p <0 .001$) or high level (dynamic and static condition $p < 0.001$). Also, the arms where fixed for a longer period of time whether the body had a high or a low level of arousal rather than a middle level (dynamic and static condition: high > middle, $p < 0.001$; low > middle, $p < 0.001$; static condition: high > low, $p = 0.002$). We found that subjects spent less time looking at the legs when the body had a high level of arousal (dynamic and static condition: high < low, $p < 0.001$; high < middle, $p < 0.001$). No significant differences emerged regarding the time spent looking to the face depending on the arousal level of the body. Figure 5, upper panels, shows the results of the rm ANOVAs on the averaged fixation times for the main factor Form. A significant effect was found in the dynamic condition ($F(1,27) = 5.069$, $p = 0.032$, $\eta_p^2 = 0.158$), showing that subjects spent more time looking at the avatar when it was presented within the low arousing architecture. No significant effect emerged in the static condition as regards the main factor Form ($F(1,29) = 2.327$, $p = 0.138$, $\eta_p^2 = 0.074$). Figure 5, lower panels, represents instead the two-way interaction Form x ROI for the two conditions of the experiment. The interaction was significant in the dynamic condition

(F(3,81) = 6.630, p < 0.001, ηp2 = 0.197), as well as in the static one (F(3,87) = 6.248, p < 0.001, ηp2 = 0.177). Specifically, as concerns the dynamic condition, HSD Tuckey corrected pairwise comparisons revealed that subjects spent more time looking at the face of the avatar when it was presented within the architecture with a low arousing form rather than in the high arousing one (low arousing form > high arousing form, p = 0.03). Instead, in the static condition, no significant comparisons emerged between fixation times on avatar's ROIs viewed within architectures with low vs. high arousing form. Also, the three way interaction Form x ROI x Body was significant in the dynamic condition (F(6,162) = 2.292, p = 0.037, η_p^2 = 0.117) but not in the static one (F(6,174) = 1.276, p = 0.271, η_p^2 = 0.042). The main effect Colour resulted to be significant (F(1,27) = 14.517, p = < 0.001, η_p^2 = 0.349) in the dynamic condition, in contrast to the static condition (F(1,29) = 3.677, p = 0.065, η_p^2 = 0.112). The two way interaction Colour x ROI was not significant in the dynamic condition (F(3,81) = 0.849, p = 0.470, η_p^2 = 0.030), meanwhile it was significant in the static one (F(3,87) = 7.542, p < 0.001, η_p^2 = 0.206). Specifically, HSD Tuckey corrected pairwise comparisons revealed that fixation times over the trunk were higher when the avatar was presented within the warm coloured architecture rather than in the cold one (warm > cold, p = 0.008).

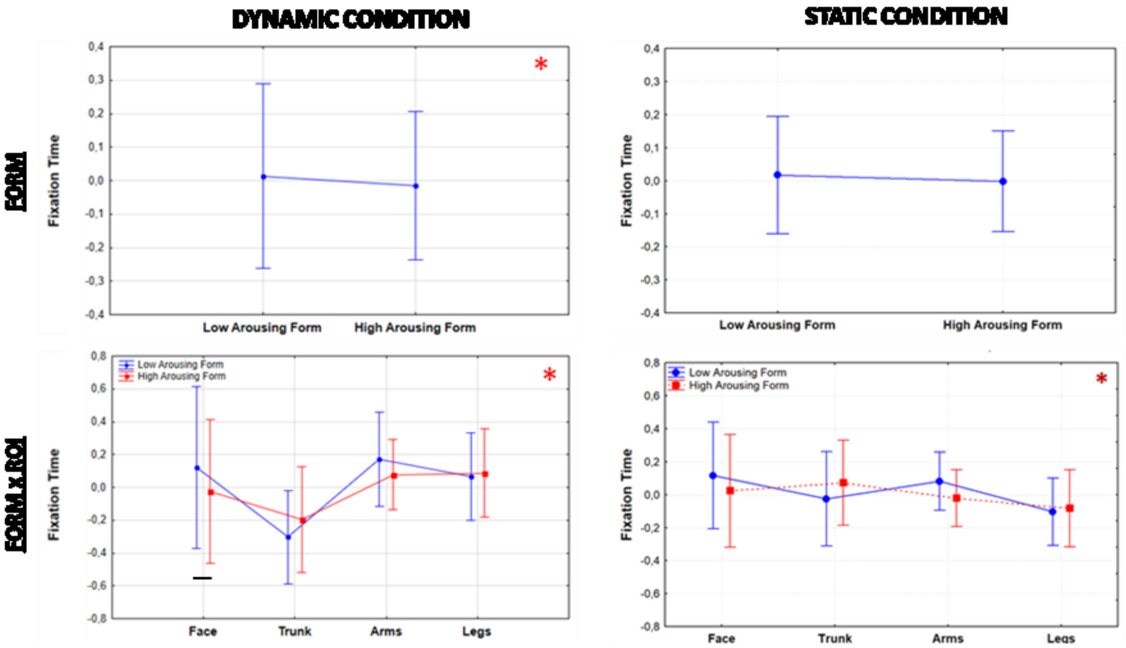

Figure* 5| *Averaged Fixation Times. *Results of the rm ANOVA on averaged fixation times for the dynamic (left column) and static (right column) condition of the experiment. Upper and lower panels show the distribution of the averaged fixation time for the main factors Form and for the interaction Form x ROI, respectively. Same colour code as Figure 3.*

Results of the dynamic analysis of the fixation times are illustrated in figure 6. The fixation time course over the considered ROIs was compared across different experimental conditions. Specifically, we deeply investigated those differences emerged from the rm ANOVA on the averaged fixation times, by finding the precise time period in which such differences occurred. Thus, for the dynamic condition of the experiment, we firstly investigated differences in the time course of fixation depending on the form of the architecture (low arousing vs. high arousing) in which the avatar appeared. Figure 6, panel A, shows that within the range of 600ms to 1300ms from the presentation of the avatar, subjects spent significantly more time looking at the face of the avatar within the low arousing architecture rather than in the high arousing one ($p = 0.002$). Then, we compared the time course of fixations on avatars observed within architectures with different forms considering only those avatars with the same level of arousal, thus we conducted three separate analyses (one per level of arousal, i.e. low, middle and high).

For avatars with a middle level of arousal, we found that between 1200ms and 1500ms from the avatar onset, subjects spent more time looking at the arms whether the avatar was presented within the low arousing architecture rather than in the high arousing one (p = 0.045). Similarly, between 900 and 1100ms, subjects spent more time staring at the face of the avatar when it was presented within the low arousing architecture (p = 0.045). For a high level of arousal expressed by the avatar's body posture, we found that between 600ms and 900ms from the avatar onset, the time spent staring at the face of the avatar was significantly higher when it was presented within the low arousing architecture rather than in the high arousing one (p = 0.048). Furthermore, we found that between 1100 - 1500ms and 2400 - 2800ms, subjects spent significantly more time looking at the arms when the avatar appeared within the low arousing architectures rather than within the high arousing one (interval 1100 - 1500ms: p = 0.046; interval 2400 - 2800ms: p = 0.046). Finally, no significant differences emerged comparing subjects gaze behaviour on avatars with low level of arousal observed within low or high arousing architecture.

The same analysis was conducted for the static condition of the experiment, revealing only a significant difference regarding the time spent looking at the trunk of avatars with middle level of arousal. Specifically, within the range of 0ms to 300ms from the presentation of the avatar, subjects spent significantly more time looking at the trunk of the avatar when it was presented within the low arousing architecture rather than in the high arousing one (p = 0.015).

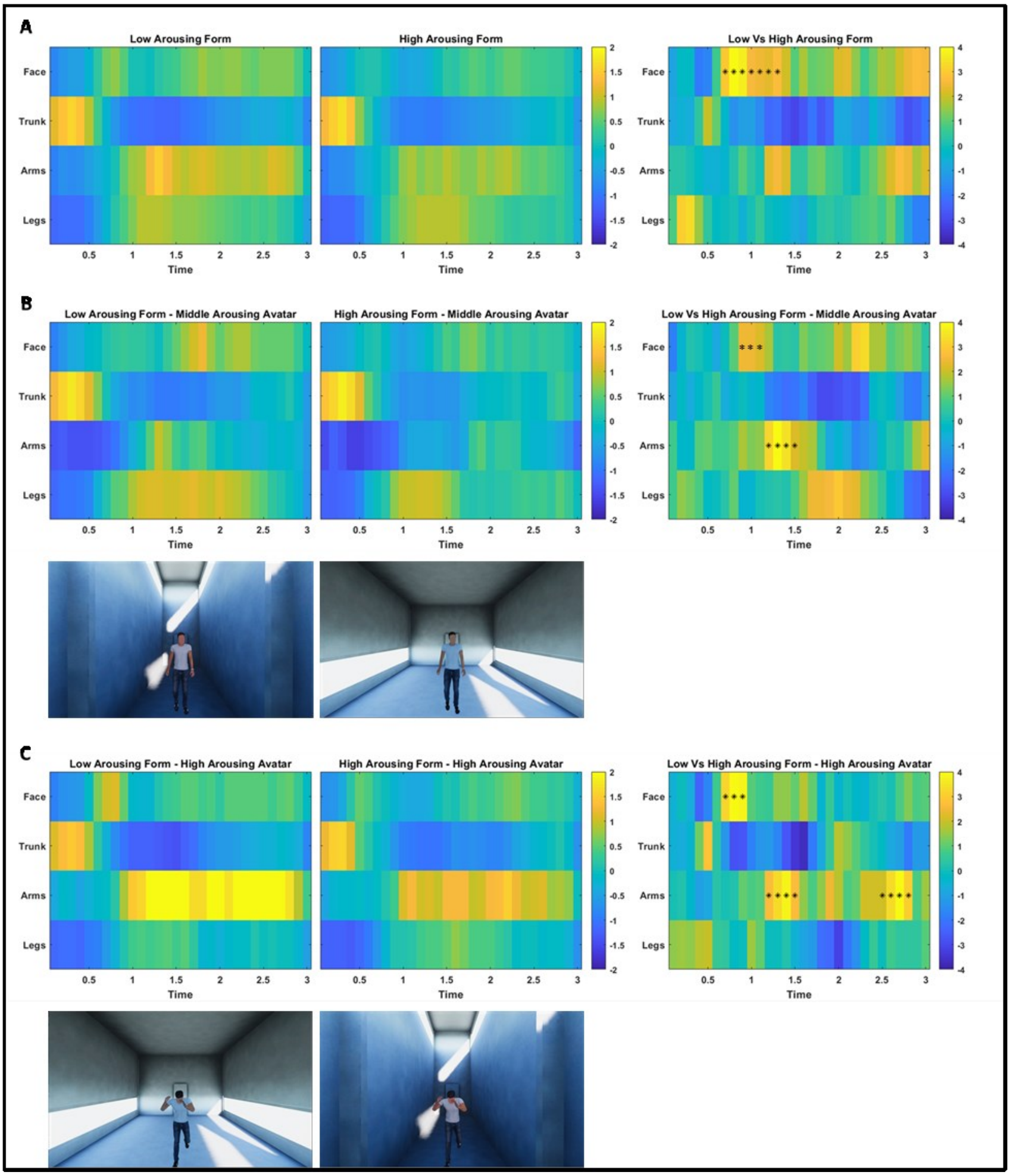

Figure 6| ***Fixation Times Dynamic Analysis. Panels A*** *show the time course of fixations on avatar's ROIs within the low (left panel) and high (middle panel) arousing architecture. The colour of each time bin codes the time spent staring at the ROI. In the right panel, a statistical comparison between the two conditions is presented: the colour of each time bin represents the t – statistic and black asterisks indicate significant comparisons (corrected for multiple comparisons using cluster-based statistics).* ***Panels B*** *represent the dynamic course of fixation times over avatars with a middle level of arousal within architectures with a low (left panel) and high (middle panel) arousing form. An avatar with a middle level of arousal within the relevant architecture is presented below each panel. In the right panel, a statistical comparison between the two condition is presented. Left and middle* ***panel C*** *represent fixation times over avatars with a high level of arousal when presented within the low and high arousing architecture, respectively. An avatar with a high level of arousal within the relevant architecture is presented below each panel. The statistical comparison is illustrated in the right panel.*

Discussion

We designed an adaptation aftereffect paradigm to investigate the relationship between the architectural experience and the perception of affective body posture. The study consisted of two conditions, i.e., dynamic and static. In the dynamic condition, subjects judged an avatar's arousal level after a virtual promenade in the architectural environment. In the static condition, subjects were directly placed at the end of the environment, in a static position, and judged the avatar’s arousal level.

In the dynamic condition, subjects’ arousal ratings were higher within the architecture with the low arousing form when compared with those assigned in the high arousal form. In the static condition, such a difference was not found. Furthermore, we found that the colour of the architecture did not modify the judgment given to the avatar. The result of the eye gaze data showed that subjects, in the dynamic experience of virtual architecture, spent more time looking at the body posture when it was presented within the low arousing architecture. In the static condition, the form does not influence the gaze responses of subjects. Particularly, the analysis of the fixation time shows that there are higher fixation times related to the observation of the avatar’s face after the dynamic experience of the low arousing architectural form. This difference is maximum around the first second of observation. As a control, we found that subjective ratings of arousal were coherent with the bodily arousal level of the avatar, both in the dynamic and static conditions.

In line with our hypothesis, the difference in the subjective scores provided in the dynamic and static conditions highlighted an adaptation aftereffect. The adaptation aftereffect is a bias on the perception of two consecutive stimuli, where the first stimulus is the adapter that influences in the opposite direction the perception of the following target stimulus

(Gibson, 1933). In recent years, such a paradigm has been exploited to investigate the interaction between different stimuli perceptions (Hedger et al., 2013). Indeed, different adapting and target stimuli, which share some conceptual meaning, are still able to generate an adaptation aftereffect (Halovic et al., 2020). In this regard, our findings demonstrate that a dynamic architectural experience modulates emotional bodies' perception. Specifically, such an architectural perception is able to generate different arousing states in subjects (Presti et al., 2021), thus influencing the subsequent perception of the arousal level of the avatar’s body. Hence, the perception of other’s affective state transcends the category of the stimuli itself, being also modified by the inner state of the subject. In particular, an adaptation aftereffect was generated by the combination of two different stimuli (Hedger et al., 2013; X. Wang et al., 2017) whose common ground was the arousal dimension: a low arousing state generated by the architectural experience (adapting stimulus) made subjects perceive a higher arousing state in the following perception of the avatar body posture (target stimulus). Considering that the architectural factors did not influence the arousal judgment of the emotional bodies in the static condition, we argue that such modulation occurs due to the dynamic architectural experience and not to a simple static architectural exposure. These results can show that the dynamic perception of architectural environments may rely on some similar perceptual mechanisms by which we process emotional body postures. In fact, the cortical motor network is not just a mere muscles controller but provide the motor representations content of the space, objects and others’ actions and emotions (Gallese et al., 2015). Specifically, the mirror mechanism points to a pre-reflective mechanism called embodied simulation: an unconscious and automatic projection of self on the perceived object or space around us influences our perceptive experience. (Gallese et al., 2015). This embodied simulation involves the way we perceive other’s affective body as well as the space around us, thus creating a bridge between our body and the surrounding space (Gallese, 2017; Gallese et al., 2015). Hence, we can argue that the perception of the architecture and emotional body posture relies on a shared sensorimotor network (Rizzolatti & Sinigaglia, 2016) thus laying the foundations for the generation of an adaptation aftereffect.

Moreover, we found that the different arousing forms influenced participants' responses to the middle level of arousal of the avatar more than the other categories. Indeed, as Hedger and colleagues showed, neutral stimuli are more likely to be affected by the adaptation aftereffect (Hedger et al., 2013). The colour factor has not produced significant differences during the dynamic experience of the architecture. Instead, in the static condition we observed that the colour modified the subject's perception of body postures more than the architecture's form. This is because during a static perception of the architectural environment, our judgement of other's affective states is more conditioned by a static, low–level feature such as colour. Conversely, the dynamic experience of the architecture is modulated by the variation of forms, which represent a main effect compared to the colour features.

The eye gaze analysis showed that in the dynamic experience of virtual architecture subjects focused their attention more on the avatar's body when it was presented within the low arousing architecture compared to the high arousing one. Specifically, subjects focused their attention on the avatar's face around the first second of the presentation within the low arousing architecture. This result could be due to the lowering of subjects' level of anxiety in such architecture, meaning that they felt comfortable looking at the face of the avatar (M. E. Kret et al., 2017) even if it was covered. In addition, the analysis of the fixation time showed that subjects spent more time looking at the arms within the low arousing architecture. Considering that the arms are a body component that subjects focus on when high arousal bodies are perceived (Geangu & Vuong, 2020), we can assume that such results could highlight an effect of adaptation aftereffect. Indeed, after the adaptation within the low arousing architecture, subjects judge a higher arousing state in the avatar body posture, thus attending more to body parts that are more descriptive of high arousing states. In the experience within the warm coloured architecture, the subjects' gaze was more focused on the avatar's trunk. The literature reports that the warm environment increases subjects' level of arousal (Yildirim et al., 2011); in this direction subjects in a state of a high level of arousal focused their attention on the trunk and avoid the face (M. E. Kret et al., 2017).

Overall, we can assume that the arousal states generated by the dynamic experience of the architecture shape our perception of affective body postures and that this can modify the subjective judgment of arousal and the participants' gaze behaviour. A possible explanation could be that these results occur due to an adaptation aftereffect. Another finding is that the emotional states generated during the architectural experience modify the attentional states, as suggested by differences in fixation times during the observation of the avatar.

Experiment 2 will investigate the neural correlates underpinning the perception of emotional body postures within different architectures to increase the knowledge of the psychophysiological mechanisms regulating such an adaptation paradigm.

EXPERIMENT 2

ERPs Study

Materials and methods

In this experiment, a different sample of participants was tested, asking to perceive the same stimuli of *Experiment 1* (see Stimuli Exp. 1 for details, Figure 2) with the same task (see Experimental Procedure Exp. 1) only in dynamic condition. We investigated the electrophysiological responses underpinning the perception of emotional body postures within different virtual architectures and evaluated the neural correlates of the adaptation aftereffect found in the previous behavioural experiment (Experiment 1). Moreover, we analysed how the information that comes from bodies is processed at the electrophysiological level. Particularly, we focused the analysis on specific event-related potentials (ERPs), such as the P200 and LPP.

Participants

Sixteen participants (26,7 ± 4,2 years, 9 female) were enrolled for this study. All participants were naïve to the purpose of the experiment and had normal, or corrected to normal, vision and no previous history of neurological disorders. The study was approved by the local ethical committee (Comitato Etico AVEN) and conducted according to the principles expressed in the Declaration of Helsinki and each participant provided written informed consent before participating in the experiment.

EEG Recording and Preprocessing

The electroencephalogram was continuously recorded using the 128-channels Geodesic EEG System (Electrical Geodesics Inc., Oregon) with the vertex as an online reference (Cz). Electrode impedances were kept below 50 kΩ, reference 10 kΩ, throughout the experiment. Continuous EEG were done through NetStation software (Electrical Geodesics, Inc., Eugene, OR, USA).

The EEG data of each participant has been imported into EEGLAB, a MATLAB toolbox for signal processing. The outermost belt of electrodes of the sensor net, prone to show residual muscular artefacts, was excluded and the original template was reduced from 129 (128 channel plus the reference) to 110 channels. Hence, we discarded 19 peripheral channels (E43, E48, E49, E56, E63, E68, E73, E81, E88, E94, E99, E107, E113, E119, E120, E125, E126, E127, E128) mainly located on the cheeks and the nape (Calbi et al., 2017).

Then, we have subsampled the data at a rate of 250 Hz, and the PREP pipeline, an EEGLAB plugin (Bigdely-Shamlo et al., 2015) for pre-processing of data, was performed. PREP pipeline allows for the identification of bad channels, their interpolation, removal of the line noise, and referencing to the average channel.

Considering that ICA decompositions are notably higher quality when data is high pass filtered above 1-2 Hz, data were band-pass filtered (2-100 Hz) (Klug & Gramann, 2021; Dimigen, 2020; Winkler et al., 2015), segmented in epochs (from -1500ms pre-stimulus before to 4000ms after the avatar presentation) and then we performed data inspection and bad trials rejection. In order to reduce the number of components resulting from the ICA, Principal Component Analysis (PCA) was computed on such data to identify the number of components that explain 99% of the variance, then the ICA was made on such components. To remove muscular, ocular, and cardiac artefacts and bad channels the Independent Component Analysis (ICA) was performed (Iriarte et al., 2003). Then, using ICLabel (Pion-Tonachini et al., 2019) a visual inspection and bad components selection has been made on the components identified by ICA.

Subsequently, EEG data obtained from the PREP pipeline were band-pass filtered (0,1-30 Hz), considering that this type of filtering is the most appropriate for the subsequent ERP analysis of EEG data (Luck, 2014), and segmented in epochs (from -1000ms pre-stimulus before to 1500ms after avatar presentation). Afterward, the resulting ICA weights were applied to this dataset (Dimigen, 2020; Djebbara et al., 2019; Klug & Gramann, 2021; Ullsperger & Debener, 2010), and finally, bad ICA components were rejected (15% of components were removed). A further visual inspection for bad trial removal was performed, leading to the rejection of 6 ± 5.7 trials on average.

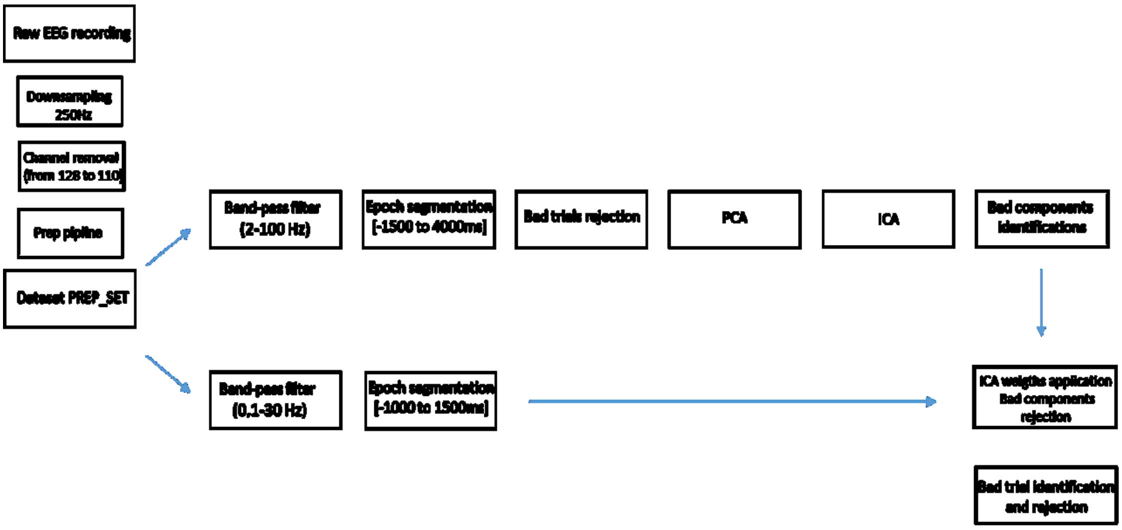

Figure 8| *EEG data preprocessing.*

Behavioural Data Analysis

Subjective arousal ratings and eye tracking data analyses were performed in the same way as *Experiment 1* (see Data Analysis Exp.1 for details). We applied the same procedure to detect attention decrease, for each subject, of each trial to mark as bad trials the outliers of the relative distribution. This procedure led to the rejection of 2% of the trials, thus ensuring a blink rate significantly lower after the bad trials rejection (paired-sample t-test (t(18) = 2.748, p 0.013)). To detect possible outliers, subjective arousal ratings were z-scored, considering the mean and standard deviation of scores provided in all the experimental conditions. Two subjects resulted as outliers in this experiment.

EEG Data Analysis

Analysis of EEG data was performed using the Factorial Mass Univariate Toolbox (FMUT; Fields & Kuperberg, 2020), an extension to Mass Univariate Toolbox (MUT; Groppe et al., 2011), compatible with the EEGLAB toolbox, so that we could explore ERP responses to stimuli on a temporal and spatial dimension. FMUT is a robust method of analysis that decreases type I and type II errors (Durston & Itier, 2021). Based on emotional stimulus processing levels (see section 1.1 Electrophysiological Studies for details), we analysed ERPs in two separate time windows: early (0 – 300ms) and late (300 – 900ms), thus ensuring that smaller effects around the ERPs of interest were not missed. Then, two separate factorial analyses were performed in such contiguous time windows with within-subjects factors body (low, middle, high), form (low arousing form, high arousing form), and colour (warm, cold). A permutation-based clustering correction was performed for correcting for multiple comparisons (significance threshold was set to 0.05). We set the number of permutation to 10000 and considered channel neighbours that were at a distance of ca. 3.5cm. Finally, pair-wise comparisons were conducted for the significant factors returned by the analysis, within the time interval of significant clusters. Specifically, levels were contrasted by means of a mass univariate analysis, that was made from temporal average of the significant time interval, finally correcting for multiple comparisons using cluster mass permutation tests.

Experimental Results

Behavioural Results

Arousal ratings. Figure 8 shows the results of the rm ANOVA on arousal ratings that subjects gave in the experiment. The main effect Body resulted to be significant (F(2,26) = 76,564, p < 0.001, ηp2 =0.854). Bonferroni corrected pairwise comparisons revealed that arousal ratings were significantly different among the three levels of the avatar's bodily arousal (low < middle, p < 0.001; low < high, p < 0.001; middle < high, p < 0.001). The main effect Form was significant (F(1,31) = 6,689, p = 0.02, ηp2 =0.339), revealing higher arousal ratings given within the architecture with the low arousing form. The main factor Colour did not result in any significant effect (F(1,13) =0,034, p =0.854, ηp2 = 0.002). The interaction Form x Body was not significant for the condition of the experiment: (F(1,13) = 0,034, p = 0.156, ηp2 = 0.133). However, we compared ratings among the two different forms considering the different arousal level of Body. For the middle avatar level paired sample t-tests highlighted a significant difference (high arousing form < low arousing form, t(3.016) =, p = 0.012). No significant effect emerged in the other two avatar level of arousal (low avatar level: high arousing form < low arousing form, t(13) = -0.39, p = 0.697; high avatar level: high arousing form < low arousing form, t(13) = 1.37, p = 0.193)

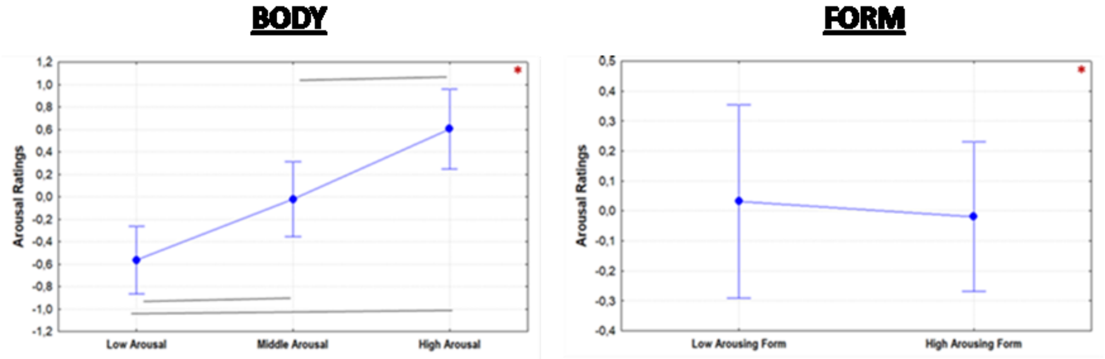

Figure 8* | *Arousal Ratings. *Results of the rm ANOVA on arousal ratings. The panels show the distribution of the arousal ratings for the main factors Avatar (left) and Form (right), respectively. Data distributions are presented with their mean and 95% confident interval. Red asterisks indicate significant main effects, while black lines show significant pairwise comparisons.*

ERP Data

The electrophysiological results of FMUT in the early time window (0 – 300ms) showed a main effect of Form (p= 0.04) spanning between 200-284ms across one significant cluster over central electrodes (Figure 9). There were no statistically significant clusters for the others factor of the analysis (Body p= 0.95; Colour p = 0.26). None of the interactions returned significant clusters (Body x Form p = 0.44, Body x Colour p = 0.47, Form x Colour p = 0.11, Body x Form x Colour p = 0.25).

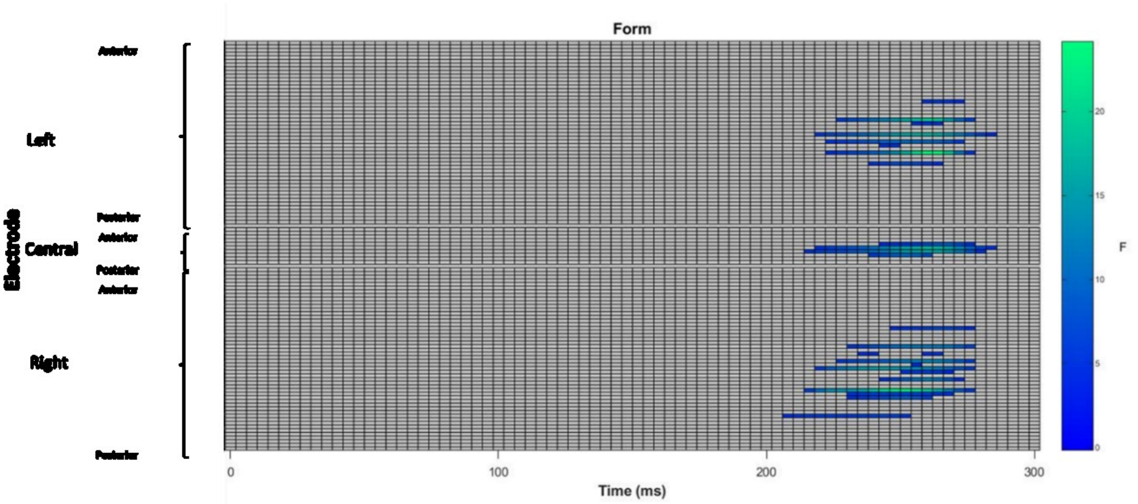

Figure 9| *Result of main effect of Form, spanning* ***200-284ms*** *across one significant cluster, from the factorial mass univariate toolbox. The x-axis depicts time in milliseconds and the y-axis depicts electrodes. The plot is split into three subgroups of electrodes from top to bottom: the left side of the scalp, the central end right side. In each subgroup, electrodes are plotted from most anterior to most posterior. The colour of the scale encodes the F statistic. Significant cluster electrodes are: E7, E13, E30, E31, E42, E53, E54 (left side); E6, E55, E72, Cz (central side); E78, E79, E80, E86, E91, E93, E104, E106, E122 (right side).*

The Cluster Mass Permutation test spanning significant time range (200 – 284ms) revealed a significant main effect between Low Arousing Form and High Arousing Form (p = 0.03). ERPs waveform showed that amplitudes for Low Arousing Form were higher than High Arousing Form across the fronto-central site (Figure 10).

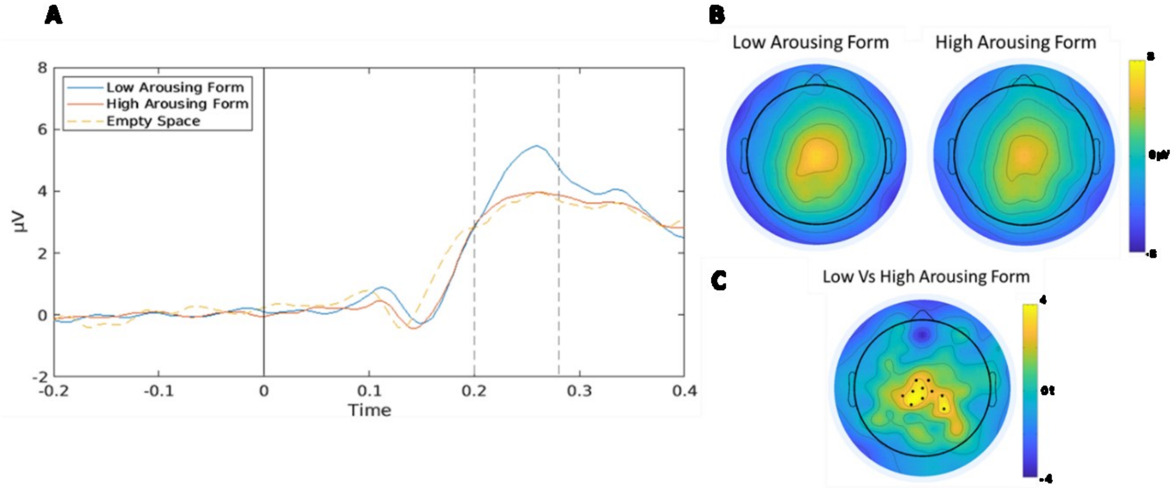

Figure 10| **Panel A** *Grand averaged ERPs from significant electrodes resulted from FMUT analysis, in response to low arousing form (blue line), high arousing form (red line) and empty space (yellow line). The range time within the dotted lines shows a statistical difference between the two waves.* **Panel B** *The topographic maps depict the distribution of voltage differences in response to Low and High arousing forms (200 – 284ms). The colour scale encodes the voltage on the scalp.* **Panel C** *The topographic map depicts the comparison within the two conditions. The colour scale encodes the t-score within the Low and High arousing forms. Significant electrodes are marked (E7, E31, E37, E54, E55, E80, E86, E87, E106, Cz.).*

In the late time window (300-900ms) the Factorial Mass Univariate analysis revealed a significant main effect of Body at one spatial–temporal cluster (p = 0.007), which ranged from 360ms to 900ms, located in central - occipital electrodes (Figure 11). There were no statistically significant clusters for the other factors of analysis (Form p = 0.14; Colour p = 0.10). None of the interactions returned significant clusters (Body x Form p = 0.70, Body x Colour p = 0.20, Form x Colour p = 0.07, Body x Form x Colour p = 0.19).

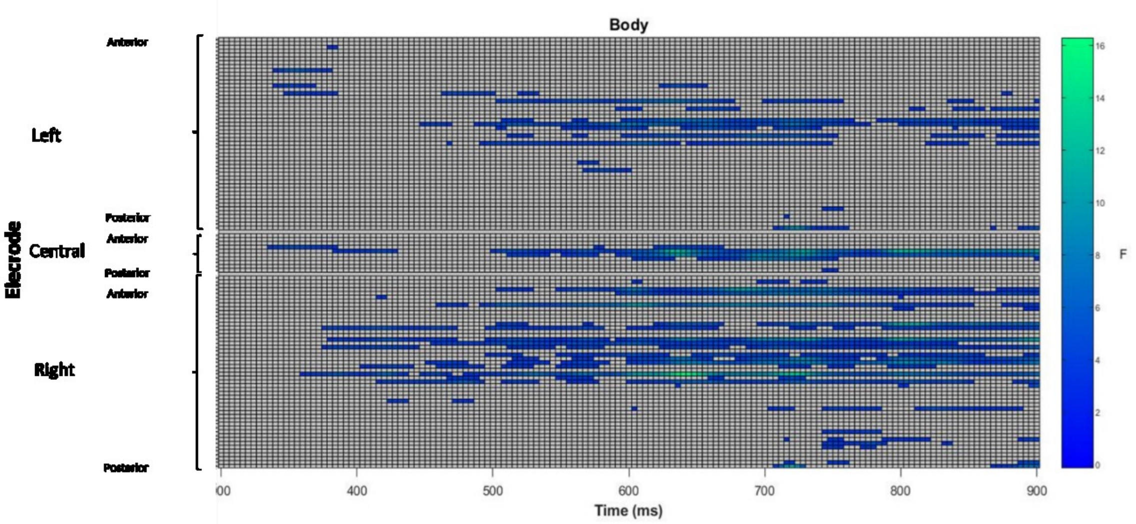

***Figure* 11|** *Result of main effect of Body, spanning 360-900ms across one significant cluster, from the factorial mass univariate toolbox. The x-axis depicts time in milliseconds and the y-axis depicts electrodes. The plot is split into three subgroups of electrodes from top to bottom: the left side of the scalp, the central end right side. In each subgroup, electrodes are plotted from most anterior to most posterior. The colour of the scale encodes the F statistic. Significant cluster electrodes are E7, E13, E28, E29, E31, E36, E37, E52, E54 (left side); E6, E16, E55, Cz (central side); E1, E2, E5, E80, E87, E91, E92, E96, E103, E104, E105, E106 ,E111, E112, E114, E115, E118 (right site).*

ERPs waveform showed that amplitudes for the avatar that had a high level of arousal were more positive compared to the other levels across the centro-parietal site (Figure 12).

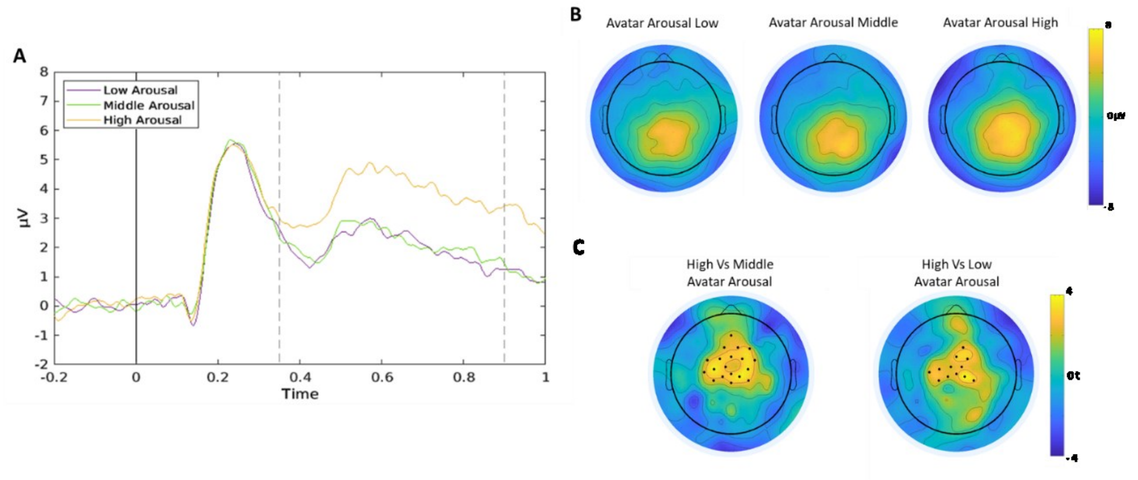

***Figure 12*| Panel A** *Grand averaged ERPs from significant electrodes resulted from the factorial mass univariate toolbox, in response to avatar arousal low (purple line), avatar arousal middle (green line) and avatar arousal high (yellow line). The time range within the dotted lines shows a statistical difference between the waves.* **Panel B** *The topographic maps depict the distribution of voltage differences in response to Low, Middle and High arousing avatars (360 – 900ms) The colour scale encodes the voltage on the scalp.* **Panel C** *The topographic map depicts the significant comparison within the conditions. The colour scale encoding the t-score within High arousal and Middle arousal avatar (significant electrodes: E5, E6, E7, E13, E29, E30, E31, E36, E37, E55, E80, E87, E105, E106, E111, E112, E118, Cz) and High arousal and Low arousal avatar (significant electrodes: E5, E6, E7, E30, E31, E36, E37, E80, E87, E105, E106, E112, Cz). Significant electrodes are marked.*

Specifically, the Cluster Mass Permutation test revealed a significant main effect between the avatar that had a high level of arousal and middle arousal avatar level (p = 0.006) and the avatar that had a high level of arousal and the low arousal avatar level (p = 0.02), but not between levels low arousal and middle arousal (p = 0.64).

Discussion

In *Experiment 2*, we found a significant effect on subjective scores and EEG features, confirming the behavioural results observed in *Experiment 1*. In particular, subjects' ratings showed higher arousal ratings given within the architecture with the low arousing form. These results highlight an adaptation aftereffect generated by the virtual walk within the architecture on the following perception of emotional body expressions. In addition, the EEG results showed two significant effects. The perception of the body postures within low arousing architectures was associated with an increase of an early cerebral component, compatible with the P200. The late time window of the ERP analysis showed a significant effect on the emotional body perception associable with the LPP. Overall, our findings shed light on how the virtual architectural experience modulates our perception of the other's emotional state and thus affecting our social behaviours.

We aimed to investigate the neural correlates underpinning the perception of emotional body postures within different virtual architectures. Therefore, the EEG data were divided into two different time windows of interest, early and late, based on affective stimuli processing levels. The significant early effect of the architecture is reflected in a modulation of the P200 amplitude at centro-parietal electrode-sides, while the significant late effect of the body is echoed in differently modulated LPP amplitude at centro-parietal electrode-sites.

The P200 has been considered an index of affective picture processing (Carretié et al., 2001, 2004). The literature shows that affective pictures elicit different amplitudes of this component (Calvo et al., 2013; Cuthbert et al., 2000). Particularly, high arousing stimuli elicited larger P200 amplitudes than low arousing stimuli (Feng et al., 2012; Paulmann et al., 2013). Our analysis revealed a significant effect of the architecture spanning 200-

284ms. Particularly, low arousing architecture elicited a significantly larger P200 amplitude during the observation of the body postures compared to high arousing architecture, which was equal to the empty space (control condition). The greater P200 amplitude during the experience within the low arousing form could reflect an effect of the adaptation aftereffect: after the perception of low arousing architectures, subjects perceive a higher arousing state in the avatar body posture, which is reflected in the higher amplitude of the P200. A further interpretation of this finding could relate to the attentional features of the early P200. Frequently, this component is identified in the attentional or target stimulus search paradigms (Luck, 2014, p. 79). During the attentive processing of the emotional stimuli, P200 is assumed to reflect enhanced attention to emotional stimuli so that they can be preferentially processed if need be, suggesting that emotional cues mobilize automatic attention resources (Delplanque et al., 2004; Paulmann et al., 2013). The underlying factor determining amplitude modulations is selective attention to objects within the affective image that is assumed to be of intrinsic relevance (Olofsson et al., 2008). Then, this component may reflect an important early mechanism that facilitates the processing of motivationally or emotionally relevant stimuli to support fast discrimination of the information to facilitate accurate behavioural responses (Gerdes et al., 2013). Our results could emphasize the connection between the dimensions of arousal and attention (Kuppens et al., 2013). The P200 amplitude increment during the perception of emotional body postures after the dynamic experience in low arousing architecture may reflect the preservation of attentional resources. Indeed, emotional arousal modifies attention and heightens sensitivity to environmental cues. Particularly, the perception of a low arousing cue broadens the attention and cognition, which may serve adaptive functions (Gable & Harmon-Jones, 2008), while high arousing stimuli harm attention performance by reducing the impact of the other information (Fenske & Eastwood, 2003).

LPP component is a positive potential showing at approximately 300ms after stimulus onset, and it appears to reflect the combined activity of dorsal and ventral extrastriate visual regions (Sabatinelli et al., 2013). The amplitude of this component is modified the arousal level of the stimulus (Hajcak & Foti, 2020; MacNamara et al., 2022). Our analysis revealed a significant effect on the body factor spanning 360-900ms. The results aligned with the literature and showed that LPP amplitudes were modulated for high and low arousing stimuli differently. Indeed, the perception of high arousing body postures increased LPP amplitude compared with the other two conditions, middle and low. Therefore, the LPP may represent a prolonged and associative conceptual analysis of the affective cue that is more closely associated with rated arousal. In line with the processing model of affective stimuli (Luo et al., 2010; MacNamara et al., 2022), the present finding confirms that the LPP component is a discriminant of affective stimuli. Therefore, we can say that larger LPP effects for high arousing body postures reflect processing of salient affective information, which might represent preferential processing of emotionally relevant stimuli, reflecting the engagement of the motivational system for more complete processing of the stimuli.

In short, these findings showed that the affect-modulation of visual ERPs during the perception of target stimuli in our adaptation paradigm can be identified at both early and later processing stages. In the early stage, we found a significant effect of the architecture that generates different affective states modulating our attentional resources, which are then employed when the subjects/participants perceive the emotional body expressions. This finding suggests that perception of the architectural environment can have very different consequences for attention, cognition, and behaviour. Then, the body stimuli are distinguished based on their affective level in the following processing stage, regardless of the factor architecture. Particularly, the LPP reflects increased cognitive processing of bodily stimuli, and its amplitude is sensitive to arousal differences.

Conclusion

The present thesis project aimed to provide new insights for studying the architectural environment's impact on social behaviour. Indeed, in our everyday lives, emotion perception occurs in complex situations where the surrounding architecture plays a crucial role in influencing our perception of others' affective states. Our experiments in virtual reality allowed us to recreate a highly realistic and full controlled environment to evaluate how we perceive others' affective states modulated by a dynamic architectural experience. Our findings showed that the arousal state generated by the dynamic experience of the architecture modulates the perception of emotional bodies. Behavioural gaze responses and EEG data aided us in understanding how this modulation is associated with different attentional levels towards affective stimuli. Indeed, our results emphasize that arousing architectures shape our attentional resources, thus highlighting that a dynamic architectural experience can broaden or decrease our attention to the other's affective states. We found that within architecture that generates a low arousing state, subjects tend to look more at the other and have a higher level of attention to the other people.

In the following steps of the project, we will increase the experimental sample and investigate the N170 component's role in the perception of emotional body posture and the modulation of the sensorimotor rhythm to increase knowledge of the neural correlates underlying the observed phenomenon. Subsequently, future experiments could integrate our findings and everyday elements, contributing to increasing the realism and explaining the impact of architecture on human life. In fact, our virtual architecture was de-contextualized. Conversely, in everyday life, people usually experience an architectural environment full of details. Also, to increase the realism, the virtual avatar may be improved, and interaction with him could be introduced.

Our experiments pave the way to investigate complex human perceptions underlying the perception of emotional body postures within architectural environments. In addition, such findings can help create better architecture for users' changing needs, allowing to improve social relationships. These human-centred spaces will facilitate social interaction, thus making our everyday social life better.

Bibliography

Acosta, I., Campano Laborda, M., Leslie, R., & Radetski, L. (2019). Daylighting design for healthy environments: Analysis of educational spaces for optimal circadian stimulus. *Solar Energy*, *193*, 584–596. https://doi.org/10.1016/j.solener.2019.10.004

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*(6), 717–746. https://doi.org/10.1068/p5096

Aviezer, H., Dudarev, V., Bentin, S., & Hassin, R. R. (2011). The Automaticity of Emotional Face-Context Integration. *Emotion (Washington, D.C.)*, *11*(6), 1406–1414. https://doi.org/10.1037/a0023578

Bachmann, J., Zabicki, A., Munzert, J., & Krüger, B. (2020). Emotional expressivity of the observer mediates recognition of affective states from human body movements. *Cognition & Emotion*, *34*(7), 1370–1381. https://doi.org/10.1080/02699931.2020.1747990

Barrett, L. F. (1998). Discrete Emotions or Dimensions? The Role of Valence Focus and Arousal Focus. *Cognition and Emotion*, *12*(4), 579–599. https://doi.org/10.1080/026999398379574

Bayet, L., & Nelson, C. A. (2019). The Perception of Facial Emotion in Typical and Atypical Development. In V. LoBue, K. Pérez-Edgar, & K. A. Buss (A c. Di), *Handbook of Emotional Development* (pagg. 105–138). Springer International Publishing. https://doi.org/10.1007/978-3-030-17332-6_6

Bianchi-Berthouze, N., Cairns, P., Cox, A., Jennett, C., & Kim, W. W. (s.d.). *On posture as a modality for expressing and recognizing emotions*. 7.

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, A. M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(6), 1477–1485. https://doi.org/10.1037/a0019057

Blakemore, S.-J., & Frith, C. (2005). The role of motor contagion in the prediction of action. *Neuropsychologia*, *43*(2), 260–267. https://doi.org/10.1016/j.neuropsychologia.2004.11.012

Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews. Neuroscience*, *12*(12), 752–762. https://doi.org/10.1038/nrn3122

Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *114*(18), 4793–4798. https://doi.org/10.1073/pnas.1618228114

Bönsch, A., Radke, S., Overath, H., Asché, L. M., Wendt, J., Vierjahn, T., Habel, U., & Kuhlen, T. W. (2018). Social VR: How Personal Space is Affected by Virtual Agents' Emotions. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 199–206. https://doi.org/10.1109/VR.2018.8446480

Borhani, K., Làdavas, E., Maier, M. E., Avenanti, A., & Bertini, C. (2015). Emotional and movement-related body postures modulate visual processing. *Social Cognitive and Affective Neuroscience*, *10*(8), 1092–1101. https://doi.org/10.1093/scan/nsu167

Caffier, P. P., Erdmann, U., & Ullsperger, P. (2003). Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, *89*(3), 319–325. https://doi.org/10.1007/s00421-003-0807-5

Calbi, M., Angelini, M., Gallese, V., & Umiltà, M. A. (2017). "Embodied Body Language": An electrical neuroimaging study with emotional faces and bodies. *Scientific Reports*, *7*(1), 6875. https://doi.org/10.1038/s41598-017-07262-0

Calvo, M. G., Marrero, H., & Beltrán, D. (2013). When does the brain distinguish between genuine and ambiguous smiles? An ERP study. *Brain and Cognition*, *81*(2), 237–246. https://doi.org/10.1016/j.bandc.2012.10.009

Carretié, L., Hinojosa, J. A., Martín-Loeches, M., Mercado, F., & Tapia, M. (2004). Automatic attention to emotional stimuli: Neural correlates. *Human Brain Mapping*, *22*(4), 290–299. https://doi.org/10.1002/hbm.20037

Carretié, L., Martín-Loeches, M., Hinojosa, J. A., & Mercado, F. (2001). Emotion and attention interaction studied through event-related potentials. *Journal of Cognitive Neuroscience*, *13*(8), 1109–1128. https://doi.org/10.1162/089892901753294400

Chaby, L., Hupont, I., Avril, M., Luherne-du Boullay, V., & Chetouani, M. (2017). Gaze Behavior Consistency among Older and Younger Adults When Looking at Emotional Faces. *Frontiers in Psychology*, *8*, 548. https://doi.org/10.3389/fpsyg.2017.00548

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. *Trends in Cognitive Sciences*, *18*(7), 370–375. https://doi.org/10.1016/j.tics.2014.03.003

Choi, D., Nishimura, T., Motoi, M., Egashira, Y., Matsumoto, R., & Watanuki, S. (2014). Effect of empathy trait on attention to various facial expressions: Evidence from N170 and late positive potential (LPP). *Journal of Physiological Anthropology*, *33*(1), 18. https://doi.org/10.1186/1880-6805-33-18

Cinzia, D. D., & Vittorio, G. (2009). Neuroaesthetics: A review. *Current Opinion in Neurobiology*, *19*(6), 682–687. https://doi.org/10.1016/j.conb.2009.09.001

Coburn, A., Vartanian, O., Kenett, Y. N., Nadal, M., Hartung, F., Hayn-Leichsenring, G., Navarrete, G., González-Mora, J. L., & Chatterjee, A. (2020). Psychological and neural responses to architectural interiors. *Cortex*, *126*, 217–241. https://doi.org/10.1016/j.cortex.2020.01.009

Cohen, M. X. (2014). *Analyzing Neural Time Series Data: Theory and Practice*. MIT Press.

Colin Ellard. (2015). *Places of the Heart*. Bellevue Literary Press.

Coulson, M. (2004). Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. *Journal of Nonverbal Behavior*, *28*(2), 117–139. https://doi.org/10.1023/B:JONB.0000023655.25550.be

Cuthbert, B. N., Schupp, H. T., Bradley, M. M., Birbaumer, N., & Lang, P. J. (2000). Brain potentials in affective picture processing: Covariation with autonomic arousal and affective report. *Biological Psychology*, *52*(2), 95–111. https://doi.org/10.1016/s0301-0511(99)00044-7

Dael, N., Goudbeek, M., & Scherer, K. (2013). Perceived Gesture Dynamics in Nonverbal Expression of Emotion. *Perception*, *42*, 642–657. https://doi.org/10.1068/p7364

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion (Washington, D.C.)*, *12*(5), 1085–1101. https://doi.org/10.1037/a0025737

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242–249. https://doi.org/10.1038/nrn1872

de Gelder, B. (2009). Why bodies? Twelve reasons for including bodily expressions in affective neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3475–3484. https://doi.org/10.1098/rstb.2009.0190

de Gelder, B., de Borst, A. W., & Watson, R. (2015). The perception of emotion in body expressions. *Wiley Interdisciplinary Reviews. Cognitive Science*, *6*(2), 149–158. https://doi.org/10.1002/wcs.1335

de Gelder, B., & Hadjikhani, N. (2006). Non-conscious recognition of emotional body language. *Neuroreport*, *17*(6), 583–586. https://doi.org/10.1097/00001756-200604240-00006

de Gelder, B., Van den Stock, J., Meeren, H. K. M., Sinke, C. B. A., Kret, M. E., & Tamietto, M. (2010). Standing up for the body. Recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. *Neuroscience and Biobehavioral Reviews*, *34*(4), 513–527. https://doi.org/10.1016/j.neubiorev.2009.10.008

Delplanque, S., Lavoie, M. E., Hot, P., Silvert, L., & Sequeira, H. (2004). Modulation of cognitive processing by emotional valence studied through event-related potentials in humans. *Neuroscience Letters*, *356*(1), 1–4. https://doi.org/10.1016/j.neulet.2003.10.014

Dimigen, O. (2020). Optimizing the ICA-based removal of ocular EEG artifacts from free viewing experiments. *NeuroImage*, *207*, 116117. https://doi.org/10.1016/j.neuroimage.2019.116117

Djebbara, Z. (2021, febbraio 2). The brain dynamics of architectural affordances during transition (2021). *Zak's Blog*. https://zakdjebbara.com/2021/02/02/the-brain-dynamics-of-architectural-affordances-during-transition-2021/

Djebbara, Z., Fich, L. B., Petrini, L., & Gramann, K. (2019). Sensorimotor brain dynamics reflect architectural affordances. *Proceedings of the National Academy of*

Sciences of the United States of America, *116*(29), 14769–14778. https://doi.org/10.1073/pnas.1900648116

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science (New York, N.Y.)*, *293*(5539), 2470–2473. https://doi.org/10.1126/science.1063414

Durston, A. J., & Itier, R. J. (2021). The early processing of fearful and happy facial expressions is independent of task demands – Support from mass univariate analyses. *Brain Research*, *1765*, 147505. https://doi.org/10.1016/j.brainres.2021.147505

Ellsworth, P. C., & Scherer, K. R. (2003). Appraisal processes in emotion. In *Handbook of affective sciences* (pagg. 572–595). Oxford University Press.

Enea, V., & Iancu, S. (2016). Processing emotional body expressions: State-of-the-art. *Social Neuroscience*, *11*(5), 495–506. https://doi.org/10.1080/17470919.2015.1114020

Farkas, A. H., Oliver, K. I., & Sabatinelli, D. (2020). Emotional and feature-based modulation of the early posterior negativity. *Psychophysiology*, *57*(2), e13484. https://doi.org/10.1111/psyp.13484

Feng, C., Wang, L., Liu, C., Zhu, X., Dai, R., Mai, X., & Luo, Y.-J. (2012). The Time Course of the Influence of Valence and Arousal on the Implicit Processing of Affective Pictures. *PLOS ONE*, *7*(1), e29668. https://doi.org/10.1371/journal.pone.0029668

Fenske, M. J., & Eastwood, J. D. (2003). Modulation of Focused Attention by Faces Expressing Emotion: Evidence From Flanker Tasks. *Emotion*, *3*(4), 327–343. https://doi.org/10.1037/1528-3542.3.4.327

Fich, L. B., Jönsson, P., Kirkegaard, P. H., Wallergård, M., Garde, A. H., & Hansen, Å. (2014). Can architectural design alter the physiological reaction to psychosocial stress? A virtual TSST experiment. *Physiology & Behavior*, *135*, 91–97. https://doi.org/10.1016/j.physbeh.2014.05.034

Fields, E. C., & Kuperberg, G. R. (2020). Having your cake and eating it too: Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, *57*(2), e13468. https://doi.org/10.1111/psyp.13468

Fogassi, L., & Gallese, V. (2002). The neural correlates of action understanding in non-human primates. In *Mirror neurons and the evolution of brain and language* (pagg. 13–35). John Benjamins Publishing Company. https://doi.org/10.1075/aicr.42.03fog

Freedberg, D., & Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Sciences*, *11*(5), 197–203. https://doi.org/10.1016/j.tics.2007.02.003

Gable, P. A., & Harmon-Jones, E. (2008). Approach-Motivated Positive Affect Reduces Breadth of Attention. *Psychological Science*, *19*(5), 476–482. https://doi.org/10.1111/j.1467-9280.2008.02112.x

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience*, *3*(3–4), 317–333. https://doi.org/10.1080/17470910701563608

Gallese, V. (2009). Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification. *Psychoanalytic Dialogues*, *19*(5), 519–536. https://doi.org/10.1080/10481880903231910

Gallese, V. (2017). Visions of the body. Embodied simulation and aesthetic experience. *Aisthesis. Pratiche, Linguaggi e Saperi Dell'estetico*, *10*(1), 41–50. https://doi.org/10.13128/Aisthesis-20902

Gallese, V., & D., Ruzzon. (2016). *Tuned Architecture. Harmony, Experience and Architecture.* https://www.overvieweditore.com/prodotto/tuned-architecture/.

Gallese, V., Pallasmaa, J., Mallgrave, H., & Robinson, S. (2015). *Architecture and Empathy*.

Geangu, E., & Vuong, Q. C. (2020). Look up to the body: An eye-tracking investigation of 7-months-old infants' visual exploration of emotional body expressions. *Infant Behavior and Development*, *60*, 101473. https://doi.org/10.1016/j.infbeh.2020.101473

Gerdes, A., Wieser, M., Bublatzky, F., Kusay, A., Plichta, M., & Alpers, G. (2013). Emotional sounds modulate early neural processing of emotional pictures. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/article/10.3389/fpsyg.2013.00741

Gibson, J. J. (1933). Adaptation, after-effect and contrast in the perception of curved lines. *Journal of Experimental Psychology*, *16*(1), 1–31. https://doi.org/10.1037/h0074626

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Hajcak, G., & Foti, D. (2020). Significance?... Significance! Empirical, methodological, and theoretical connections between the late positive potential and P300 as neural responses to stimulus significance: An integrative review. *Psychophysiology*, *57*(7), e13570. https://doi.org/10.1111/psyp.13570

Halovic, S., Kroos, C., & Stevens, C. (2020). Adaptation aftereffects influence the perception of specific emotions from walking gait. *Acta Psychologica*, *204*, 103026. https://doi.org/10.1016/j.actpsy.2020.103026

Haq, Z. A., & Hasan, Z. (2016). Eye-blink rate detection for fatigue determination. *2016 1st India International Conference on Information Processing (IICIP)*, 1–5. https://doi.org/10.1109/IICIP.2016.7975348

Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior & Development*, *50*, 42–51. https://doi.org/10.1016/j.infbeh.2017.10.007

Hedger, S. C., Nusbaum, H. C., Lescop, O., Wallisch, P., & Hoeckner, B. (2013). Music can elicit a visual motion aftereffect. *Attention, Perception, & Psychophysics*, *75*(5), 1039–1047. https://doi.org/10.3758/s13414-013-0443-z

Hietanen, J. K., Kirjavainen, I., & Nummenmaa, L. (2014). Additive effects of affective arousal and top-down attention on the event-related brain responses to human bodies. *Biological Psychology*, *103*, 167–175. https://doi.org/10.1016/j.biopsycho.2014.09.003

Kana, R. K., & Travers, B. G. (2012). Neural substrates of interpreting actions and emotions from body postures. *Social Cognitive and Affective Neuroscience*, *7*(4), 446–456. https://doi.org/10.1093/scan/nsr022

Karg, M., Kühnlenz, K., & Buss, M. (2010). Recognition of Affect Based on Gait Patterns. *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, *40*, 1050–1061. https://doi.org/10.1109/TSMCB.2010.2044040

Kleinsmith, A., & Bianchi-Berthouze, N. (2013). Affective Body Expression Perception and Recognition: A Survey. *IEEE Transactions on Affective Computing*, *4*(1), 15–33. https://doi.org/10.1109/T-AFFC.2012.16

Kleinsmith, A., Bianchi-Berthouze, N., & Steed, A. (2011). Automatic Recognition of Non-Acted Affective Postures. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *41*(4), 1027–1038. https://doi.org/10.1109/TSMCB.2010.2103557

Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2005). Grounding Affective Dimensions into Posture Features. In J. Tao, T. Tan, & R. W. Picard (A c. Di), *Affective Computing and Intelligent Interaction* (pagg. 263–270). Springer. https://doi.org/10.1007/11573548_34

Kleinsmith, A., De Silva, P. R., & Bianchi-Berthouze, N. (2006). Cross-cultural differences in recognizing affect from body posture. *Interacting with Computers*, *18*(6), 1371–1389. https://doi.org/10.1016/j.intcom.2006.04.003

Kleinsmith, A., & Semsar, A. (2019). Perception of Emotion in Body Expressions from Gaze Behavior. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3290607.3313062

Kloth, N., Schweinberger, S. R., & Kovács, G. (2010). Neural correlates of generic versus gender-specific face adaptation. *Journal of Cognitive Neuroscience*, *22*(10), 2345–2356. https://doi.org/10.1162/jocn.2009.21329

Klug, M., & Gramann, K. (2021). Identifying key factors for improving ICA-based decomposition of EEG data in mobile and stationary experiments. *European Journal of Neuroscience*, *54*(12), 8406–8420. https://doi.org/10.1111/ejn.14992

Kovács, G., Zimmer, M., Bankó, E., Harza, I., Antal, A., & Vidnyánszky, Z. (2006). Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cerebral Cortex (New York, N.Y.: 1991)*, *16*(5), 742–753. https://doi.org/10.1093/cercor/bhj020

Kret, M. E., & de Gelder, B. (2010). Social context influences recognition of bodily expressions. *Experimental Brain Research*, *203*(1), 169–180. https://doi.org/10.1007/s00221-010-2220-8

Kret, M. E., Stekelenburg, J. J., de Gelder, B., & Roelofs, K. (2017). From face to hand: Attentional bias towards expressive hands in social anxiety. *Biological Psychology*, *122*, 42–50. https://doi.org/10.1016/j.biopsycho.2015.11.016

Kret, M., Roelofs, K., Stekelenburg, J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, *7*, 810. https://doi.org/10.3389/fnhum.2013.00810

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological Bulletin*, *139*(4), 917–940. https://doi.org/10.1037/a0030811

Lacey, S., Hagtvedt, H., Patrick, V. M., Anderson, A., Stilla, R., Deshpande, G., Hu, X., Sato, J. R., Reddy, S., & Sathian, K. (2011). Art for reward's sake: Visual art recruits the ventral striatum. *NeuroImage*, *55*(1), 420–433. https://doi.org/10.1016/j.neuroimage.2010.11.027

Leather, P., Pyrgas, M., Beale, D., & Lawrence, C. (1998). Windows in the workplace: Sunlight, view, and occupational stress. *Environment and Behavior*, *30*(6), 739–762. https://doi.org/10.1177/001391659803000601

Li, X. (2021). Recognition Characteristics of Facial and Bodily Expressions: Evidence From ERPs. *Frontiers in Psychology*, *12*, 2595. https://doi.org/10.3389/fpsyg.2021.680959

Lindal, P. J., & Hartig, T. (2013). Architectural variation, building height, and the restorative quality of urban residential streetscapes. *Journal of Environmental Psychology*, *33*, 26–36. https://doi.org/10.1016/j.jenvp.2012.09.003

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, *35*(3), 121–143. https://doi.org/10.1017/S0140525X11000446

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique, second edition*. MIT Press.

Luo, W., Feng, W., He, W., Wang, N.-Y., & Luo, Y.-J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–1867. https://doi.org/10.1016/j.neuroimage.2009.09.018

MacNamara, A., Joyner, K., & Klawohn, J. (2022). Event-related potential studies of emotion regulation: A review of recent progress and future directions. *International Journal of Psychophysiology*, *176*, 73–88. https://doi.org/10.1016/j.ijpsycho.2022.03.008

Mallgrave, H. (2013). *Architecture and Embodiment: The Implications of the New Sciences and Humanities for Design* (1°). Routledge.

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16518–16523. https://doi.org/10.1073/pnas.0507650102

Mehrabian, A. (2017). *Nonverbal Communication*. Routledge.

Negi, S., & Mitra, R. (2020). Fixation duration and the learning process: An eye tracking study with subtitled videos. *Journal of Eye Movement Research*, *13*(6), 10.16910/jemr.13.6.1. https://doi.org/10.16910/jemr.13.6.1

Neri, P. (2009). Wholes and subparts in visual processing of human agency. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1658), 861–869. https://doi.org/10.1098/rspb.2008.1363

Ngo, N., & Isaacowitz, D. M. (2015). Use of context in emotion perception: The role of top-down control, cue type, and perceiver's age. *Emotion (Washington, D.C.)*, *15*(3), 292–302. https://doi.org/10.1037/emo0000062

Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological psychology*, *77*(3), 247–265. https://doi.org/10.1016/j.biopsycho.2007.11.006

Olofsson, J. K., & Polich, J. (2007). Affective visual event-related potentials: Arousal, repetition, and time-on-task. *Biological Psychology*, *75*(1), 101–108. https://doi.org/10.1016/j.biopsycho.2006.12.006

Pasqualini, I., Blefari, M. L., Tadi, T., Serino, A., & Blanke, O. (2018). The Architectonic Experience of Body and Space in Augmented Interiors. *Frontiers in Psychology*, *9*, 375. https://doi.org/10.3389/fpsyg.2018.00375

Paul Ekman. (1999). Basic Emotion. In *Handbook of Cognition and Emotion* (pag. 868). John Wiley & Sons.

Paulmann, S., Bleichner, M., & Kotz, S. (2013). Valence, arousal, and task effects in emotional prosody processing. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/article/10.3389/fpsyg.2013.00345

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, *93*(1), 603–608. https://doi.org/10.1152/jn.00513.2004

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*(8), 636–648. https://doi.org/10.1038/nrn2195

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51-61. https://doi.org/10.1016/s0010-0277(01)00147-0

Pollux, P. M. J., Craddock, M., & Guo, K. (2019a). Gaze patterns in viewing static and dynamic body expressions. *Acta Psychologica*, *198*, 102862. https://doi.org/10.1016/j.actpsy.2019.05.014

Pollux, P. M. J., Craddock, M., & Guo, K. (2019b). Gaze patterns in viewing static and dynamic body expressions. *Acta Psychologica*, *198*, 102862. https://doi.org/10.1016/j.actpsy.2019.05.014

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734. https://doi.org/10.1017/S0954579405050340

Presti, P., Ruzzon, D., Avanzini, P., Caruana, F., Rizzolatti, G., & Vecchiato, G. (2021). *Dynamic experience of architectural forms affects arousal and valence perception in virtual environments*. https://doi.org/10.21203/rs.3.rs-910384/v1

Reed, C. L., Stone, V. E., Grubb, J. D., & McGoldrick, J. E. (2006). Turning configural processing upside down: Part and whole body postures. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(1), 73–87. https://doi.org/10.1037/0096-1523.32.1.73

Righart, R., & de Gelder, B. (2008). Rapid influence of emotional scenes on encoding of facial expressions: An ERP study. *Social cognitive and affective neuroscience*, *3*(3), 270–278. https://doi.org/10.1093/scan/nsn021

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*(1), 169–192. https://doi.org/10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews. Neuroscience*, *11*(4), 264–274. https://doi.org/10.1038/nrn2805

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: A basic principle of brain function. *Nature Reviews Neuroscience*, *17*(12), 757–765. https://doi.org/10.1038/nrn.2016.135

Robinson, S., & Pallasmaa, J. (A c. Di). (2015). *Mind in Architecture: Neuroscience, Embodiment, and the Future of Design*. MIT Press.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Sabatinelli, D., Keil, A., Frank, D. W., & Lang, P. J. (2013). Emotional perception: Correspondence of early and late event-related potentials with cortical and subcortical functional MRI. *Biological Psychology*, *92*(3), 513–519. https://doi.org/10.1016/j.biopsycho.2012.04.005

Salthouse, T. A., & Ellis, C. L. (1980). Determinants of Eye-Fixation Duration. *The American Journal of Psychology*, *93*(2), 207–234. https://doi.org/10.2307/1422228

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the 2000 symposium on Eye tracking research & applications*, 71–78. https://doi.org/10.1145/355017.355028

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339. https://doi.org/10.1038/nrn1651

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion*, *23*(7), 1307–1351. https://doi.org/10.1080/02699930902928969

Schupp, H. T., Ohman, A., Junghöfer, M., Weike, A. I., Stockburger, J., & Hamm, A. O. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion (Washington, D.C.)*, *4*(2), 189–200. https://doi.org/10.1037/1528-3542.4.2.189

Slater, M., Spanlang, B., & Corominas, D. (2010). Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Transactions on Graphics (TOG*, 92.

Stekelenburg, J. J., & de Gelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *Neuroreport*, *15*(5), 777–780. https://doi.org/10.1097/00001756-200404090-00007

Tamietto, M., & de Gelder, B. (2008). Emotional contagion for unseen bodily expressions: Evidence from facial EMG. *2008 8th IEEE International Conference on Automatic Face Gesture Recognition*, 1–5. https://doi.org/10.1109/AFGR.2008.4813317

Thierry, G., Pegna, A. J., Dodds, C., Roberts, M., Basan, S., & Downing, P. (2006). An event-related potential component sensitive to images of the human body. *NeuroImage*, *32*(2), 871–879. https://doi.org/10.1016/j.neuroimage.2006.03.060

Uithol, S., & Gallese, V. (2015). The role of the body in social cognition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *6*(5), 453–460. https://doi.org/10.1002/wcs.1357

Ullsperger, M., & Debener, S. (2010). *Simultaneous EEG and FMRI: Recording, Analysis, and Application*. Oxford University Press, USA.

Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion (Washington, D.C.)*, *7*(3), 487–494. https://doi.org/10.1037/1528-3542.7.3.487

Van den Stock, J., Vandenbulcke, M., Sinke, C. B. A., & de Gelder, B. (2014). Affective scenes influence fear perception of individual body expressions. *Human Brain Mapping*, *35*(2), 492–502. https://doi.org/10.1002/hbm.22195

Vartanian, O., Navarrete, G., Chatterjee, A., Fich, L. B., Gonzalez-Mora, J. L., Leder, H., Modroño, C., Nadal, M., Rostrup, N., & Skov, M. (2015). Architectural design and the brain: Effects of ceiling height and perceived enclosure on beauty judgments and approach-avoidance decisions. *Journal of Environmental Psychology*, *41*, 10–18. https://doi.org/10.1016/j.jenvp.2014.11.006

Vecchiato, G., Jelic, A., Tieri, G., Maglione, A. G., De Matteis, F., & Babiloni, F. (2015). Neurophysiological correlates of embodiment and motivational factors during the perception of virtual architectural environments. *Cognitive Processing*, *16 Suppl 1*, 425–429. https://doi.org/10.1007/s10339-015-0725-6

Vecchiato, G., Tieri, G., Jelic, A., De Matteis, F., Maglione, A. G., & Babiloni, F. (2015). Electroencephalographic correlates of sensorimotor integration and embodiment during the appreciation of virtual architectural environments. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01944

Verfaillie, K., & Daems, A. (2002). Representing and anticipating human actions in vision. *Visual Cognition*, *9*(1–2), 217–232. https://doi.org/10.1080/13506280143000403

Wang, J., Run, Y., & Shi, H. (2020). Emotional state representation and detection method of users in library space based on body posture recognition. *Digital Library Perspectives*, *36*(2), 113–125. https://doi.org/10.1108/DLP-11-2019-0041

Wang, X., Guo, X., Chen, L., Liu, Y., Goldberg, M., & Xu, H. (2017). Auditory to Visual Cross-Modal Adaptation for Emotion: Psychophysical and Neural Correlates. *Cerebral cortex*. https://doi.org/10.1093/cercor/bhv321

Webster, M. A., & MacLeod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1571), 1702–1725. https://doi.org/10.1098/rstb.2010.0360

Winkler, I., Debener, S., Müller, K.-R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 4101–4105. https://doi.org/10.1109/EMBC.2015.7319296

Xu, Q., Yang, Y., Tan, Q., & Zhang, L. (2017). Facial Expressions in Context: Electrophysiological Correlates of the Emotional Congruency of Facial Expressions and Background Scenes. *Frontiers in Psychology*, *8*, 2175. https://doi.org/10.3389/fpsyg.2017.02175

Yildirim, K., Hidayetoglu, M. L., & Capanoglu, A. (2011). Effects of Interior Colors on Mood and Preference: Comparisons of Two Living Rooms. *Perceptual and Motor Skills*, *112*(2), 509–524. https://doi.org/10.2466/24.27.PMS.112.2.509-524