Citation: Abad-Hernando, S. (2023). Neurocognitive mechanisms of body representation & action prediction. (Unpublished Doctoral thesis, City, University of London)

Neurocognitive mechanisms of body representation & action prediction

Sonia Abad-Hernando

A thesis submitted for the degree of PhD in Psychology

City, University of London

Department of Psychology

Cognitive Neuroscience Research Unit

October 2023

Table of Contents

List of Figures

List of Tables

Acknowledgements

Thanks to my first supervisor Bea for her kindness, for giving me the opportunity to come to London in the first place as a research assistant through an Erasmus programme, for taking the odds to believe in me and accept me as part of the team, encouraging and supporting me through this long & bumpy road of doing a part-time and self-funded PhD. Thanks to my unofficial mentor Rami for his kindness, wisdom, availability, and interest on discussing with me any new ideas. To Bea & Rami, I am grateful for everything you have helped me learn both on science and life!

Thanks to Tina, who not only gave me valuable feedback on my first EEG experiment, but also gave me the opportunity to work in other projects, gaining knowledge and experience beyond my PhD. Thanks to the Psychology Department, where colleagues had always been very friendly and helpful in difficult times, and even resulted my main “sponsor” through all the teaching and marking I was trusted with. Thanks to the University and everyone that makes it the way it is, for giving me the opportunity to grow through this experience.

Thanks to my mother who always believed in me and supported my decision to start the London adventure. Thanks friends & family, who have supported me in so many ways, listened to my “PhD things” for 8 years and kept asking what it is about. Thanks to my closest ones, who have been incredibly helpful over the last crazy stretch. And last but not least, special thanks to those who also became fierce guinea pigs, always ready to pilot any experiment, and gave me the most scrutinized feedback a scientist could dream with!

Research Outputs

Research outputs in preparation from this thesis:

Chapter 3 in preparation for publication

- **Abad-Hernando S**, Galvez-Pol A, Forster B, Calvo-Merino, B. S Working memory for bodies and top-down modulation: an EEG study. *In preparation for PLOS one*

Chapter 5 in preparation for publication

- **Abad-Hernando S**, Forster B, Calvo-Merino, B. The advantage of action prediction. *In preparation for Neuroscience & Biobehavioural Reviews*

Chapter 6 and Chapter 7 in preparation for publication

- **Abad-Hernando S**, Forster B, Calvo-Merino, B when action prediction and action fluency meet aesthetic value. *In preparation for Journal of Experimental Psychology: Human Perception and Performance*

Other research outputs not in this thesis:

- Joly-Mascheroni R, **Abad-Hernando S**, Forster B, Calvo-Merino B. Embodiment and Multisensory Perception of Synchronicity: Biological Features Modulate Visual and Tactile Multisensory Interaction in Simultaneity Judgements. Multisens Res. 2021 Feb 1:1-18. doi: 10.1163/22134808-bja10020. Epub ahead of print. PMID: 33535162.
- Fernandez-Cotarelo A, **Abad-Hernando S**, Gil de Montes L, Korta K, Calvo-Merino B. The Dance Aesthetic Fluency Scale: Design and Validation of a Measure of Dance Knowledge. Submitted to *Empirical Studies of the Arts.*
- Hur YJ, **Abad-Hernando S**, Joly-Mascheroni R, Trupp M and Calvo-Merino B. Preference for social content in painting and the contributions of individual differences. Submitted to *British Journal of Psychology.*
- Forster B & **Abad-Hernando S**. In your skin? Somatosensory cortex is purposely recruited to situate but not simulate observed touch. Neuroimage. 2024 Apr 1; 289:120561. doi: 10.1016/j.neuroimage.2024.120561.

Declaration

I grant powers of discretion to the Director of Library Services to copy this thesis in whole or in part without further reference to the author. This permission covers only single page copies made for study purposes, subject to normal conditions of acknowledgement.

Abbreviations

1PP: First Person Perspective

3PP: Third Person Perspective

AON: Action Observation Network

AP: Action Prediction

ASD: Autism Spectrum Disorder

BG: Basal Ganglia

BM: Biological Motion

CDA: Contralateral Delayed Activity

CSE: Corticospinal Excitability

dPMC: dorsal Premotor Cortex

EBA: Extrastriate Body Area

EEG: Electroencephalography

ERP: Event Related Potential

FBA: Fusiform Body Area

fMRI: Functional Magnetic Resonance Imaging

Gold-DSI: Goldsmiths Dance Sophistication Index

HEOGL: Left Horizontal Electrooculogram

HEOGR: Right Horizontal Electrooculogram

IFG: Inferior Frontal Gyrus

IFM: Internal Forward Model

IPS: Intraparietal Sulcus

IRI: Interpersonal Reactivity Index

M1: Primary Motor Cortex

MAIA-2: Multidimensional Assessment of Interoceptive Awareness-2

MEG: Magnetoencephalography

MNS: Mirror Neuron System

PC: Predictive Coding

PE: Prediction Error

PET: Positron Emission Tomography

PFC: Prefrontal Cortex

PLD: Point Light Display

PM: Premotor Cortex

PPC: Posterior Parietal Cortex

S1: Primary Somatosensory Cortex

SAS: Supervisory Attentional System

SCx: Somatosensory Cortex

SDT: Signal Detection Theory

SEP: Somatosensory Evoked Potential

SMA: Supplementary Motor Area

STS: Superior Temporal Sulcus

SW: Slow Wave

TAS-20: Toronto Alexithymia Scale-20

Theory of Event Coding (TEC

TMS: Transcranial Magnetic Stimulation

ToM: Theory of Mind

TPJ: Temporoparietal Junction

V1: Primary Visual Cortex

VEOG: Vertical Electrooculogram

VEP: Visual Evoked Potential

VHF: Visual Hemifield

WM: Working Memory

VWM: Visual Working Memory

COVID-19 Impact Statement.

This statement is provided for the aid and benefit of future readers to summarise the impact of the COVID-19 pandemic on the scope, methodology, and research activity associated with this thesis. The academic standards for a research degree awarded by City, University of London and for which this thesis is submitted remain the same regardless of this context.

Neurocognitive mechanisms of body representation & action prediction

When the pandemic started there was a TMS study (chapter 4) pending ethics approval and a planned EEG study. Those experiments were meant to follow-up previous work, close each line of research (both sections of this thesis) and could no longer be implemented. The consecutive lockdowns and restrictions made it impossible to collect neurophysiological data in the lab for 2 years. Within all the uncertainty it was decided to change the research question and methodology on section B and run an online experiment (chapter 7) instead. The research aims and the scope of the section B of this thesis changed considerably. To compensate for the empirical loss, chapter 5 was reshaped from a humble literature review into a big new proposal for a fresh integrative perspective in action prediction, bringing different research areas, theories, and empirical evidence together.

Date of statement: 31/10/2023

Abstract

This thesis aims to provide a comprehensive and integrative understanding of how our brain constructs reality by unifying perspectives on attention, perception, and working memory. By investigating how we process biological information and the impact of movement, it embraces embodiment theory to highlight the role of physical presence in action comprehension. Shared neural substrates underpinning body and movement perception support both motor and social functions, with attention acting as a regulatory mechanism that integrates sensory input and cognitive expectations into a coherent experience. Working memory, engaging common neural pathways in perception and memory, plays a crucial role in maintaining high-fidelity body representations. An EEG experiment was implemented to explore how top-down attentional mechanisms modulate the maintenance of body-related images in working memory. This study reveals that attentional modulation enhances persistent activity in relevant sensory areas during memory tasks, suggesting a significant role for top-down control in maintaining high-fidelity body representations. Additionally, the thesis investigates action prediction as a multidimensional process informed by sensorimotor experiences, social knowledge, and contextual cues. It emphasizes prediction error as a key mechanism in the continuous adjustment of cognitive models, calling for an integrative perspective to fully understand its complexity. Lastly, a series of behavioural studies reveal that perceptual-motor fluency significantly facilitates action prediction, underscoring the interplay between perceptual and motor systems. Furthermore, research into aesthetic preferences indicates that individual differences in aesthetic appreciation can predict performance in action prediction tasks, highlighting the subjective aspects of perceptual-motor integration.

Structure of the Thesis

This thesis aims to provide a comprehensive exploration of how the human brain constructs reality through the integration of attention, perception, and working memory, emphasizing the embodied nature of cognition. The main body of work is organised under two separate sections, overarched by a general introduction and discussion.

Chapter 1 introduces the general framework of the thesis, focusing on the theoretical and neural foundations of action representation, body and movement perception, and the role of attentional mechanisms.

Section A: Working Memory (WM) and Top-Down Modulations

This section explores the role of working memory in maintaining body representations and how top-down attentional mechanisms modulate these representations.
Chapter 2 discusses sensory recruitment models in visual working memory (VWM) and their application to body-related information. It sets the background research looking at the interplay between external and internal attentional mechanisms in working memory paradigms. *Chapter 3* presents an EEG experiment designed to elucidate how top-down attentional mechanisms influence the maintenance of body-related images in working memory for bodies. This chapter discusses the findings that highlight the enhancement of sensory areas during memory tasks through attentional modulation. *Chapter 4* is a proposal to explore motor encoding of body-related information using Transcranial Magnetic Stimulation (TMS), aiming to further understand the neural mechanisms underlying action representation.

Section B: Action Prediction & Movement Fluency

This section focuses on the predictive nature of action observation, emphasizing the integration of sensorimotor experience, movement perception and individual differences. *Chapter 5* discusses the theoretical aspects of action prediction, emphasizing the role of sensorimotor experiences, social knowledge, and contextual cues. Based on an extensive literature review on the neural underpinnings of action prediction, aims to integrate predictive frameworks such as Predictive Coding (PC) and examines prediction error as a mechanism for adjusting cognitive models. It explores the biological cues and sensorimotor representations crucial for action prediction. *Chapter 6* presents a behavioural study investigating how perceptual-motor fluency facilitates action prediction. It discusses the implications of these findings for understanding the interaction between perceptual and motor systems in action prediction tasks. *Chapter 7* examines individual differences in aesthetic appreciation and their relationship to performance in action prediction tasks. This chapter explores the subjective aspects of perceptual-motor integration and their impact on predictive abilities.

Chapter 8 Synthesizes the findings across the thesis, emphasizing the interconnected roles of embodiment, attention, working memory, and predictive processes in constructing our perceptual reality. It discusses implications from the empirical and theoretical work, and future directions for research.

Chapter 1: General Framework of the Thesis

1.1. Action representation: embodiment (cognitive models Vs. neural representation)

The way actions are represented in the human brain has long been a topic of interest and debate among researchers. The ideomotor principle, originally articulated by William James in his seminal work "Principles of Psychology" (1890), is a foundational concept in the understanding of action representation, particularly within the cognitive psychological framework. According to James, every representation of a movement awakens to some degree the actual movement itself. His principle emphasizes the bidirectional link between action and perception: the idea that actions are represented in terms of their anticipated perceptual consequences (James, 1890). Action representation bridges the realms of cognitive models and neural underpinnings, with the concept of embodiment playing a pivotal role. This chapter delves into how actions are embodied, contrasting cognitive models and neural representations.

1.1.1. Embodiment and Its Importance

Embodiment posits that our bodily experiences play a fundamental role in shaping the way we think, perceive, and act. The body is not just a passive entity but actively contributes to our cognitive processes. The representation of actions, therefore, is intrinsically tied to our bodily experiences and capabilities. The concept of embodiment has emerged as a pivotal paradigm in cognitive psychology, neuroscience, and philosophy of mind, challenging traditional computational and representational models of cognition. Embodiment posits that cognitive processes are deeply rooted in the body's interactions with the world, implying that the mind cannot be fully understood without considering its physical substrate and environmental context. The theoretical importance of embodiment lies in its transformative impact on how cognition is understood, offering a robust alternative to Cartesian dualism and traditional computational models. By emphasizing the interdependence of mind, body, and environment, embodiment disrupts the notion of the mind as an isolated computational unit, proposing instead a holistic, integrated view of cognitive processes (Varela, Thompson, & Rosch, 1991). This perspective facilitates a re-evaluation of cognitive load, positing that the environment itself can function as an extension of the cognitive system, thereby enabling cognitive offloading (Clark, 1999). Moreover, embodiment offers a framework for grounding abstract concepts in sensory-motor experiences, thereby providing a physiological and experiential foundation for higher-order cognitive functions such as language comprehension and problem-solving (Barsalou, 2008). In sum, the theoretical significance of embodiment resides in its capacity to offer a more integrated and empirically grounded understanding of cognition, challenging and enriching traditional paradigms.

The empirical significance of embodiment manifests in various domains of cognitive science, providing concrete evidence for its theoretical underpinnings. For example, research on the action-perception cycle has validated the embodied viewpoint by illustrating that perceptual processes are not merely receptive but actively engage with and shape the environment, thereby strengthening the argument that cognition is deeply interlinked with action (Prinz, 1997). Further credence comes from the domain of social cognition, where the discovery of mirror neurons has offered a neurobiological mechanism to support the idea that understanding others’ actions and emotions is anchored in one’s own embodied experiences (Gallese & Goldman, 1998). Moreover, studies in psycholinguistics, such as those focused on sentence comprehension, have demonstrated that language processing often involves mental simulation of sensory and motor experiences, thereby bolstering the embodied framework (Zwaan & Taylor, 2006). Collectively, these empirical findings not only substantiate the theoretical claims of embodiment but also offer a methodological pathway for multidisciplinary research, affirming the embodied perspective as an empirically grounded and versatile approach in cognitive science.

1.1.2. Cognitive Models of Action Representation.

Cognitive models primarily focus on the mental structures and processes that facilitate action representation. Cognitive models of action representation aim to elucidate how actions are mentally coded, planned, and executed. These models address various cognitive processes, including attention, memory, and decision-making, and they often integrate insights from psychology, neuroscience, and computational modelling.

Ideomotor theories propose a fundamental integration between action and perception by positing that actions are mentally represented through their sensory consequences. This concept traces its roots back to William James in the late 19th century, who argued that merely thinking about the effects of a particular action could suffice to initiate that action (James, 1890). Modern adaptations of this theory, such as the *Theory of Event Coding (TEC)*, further elaborate on this by suggesting a common representational medium for both perceptual events and actions (Hommel et al., 2001). In this framework, actions are not merely coded as isolated motor commands but are instead linked bidirectionally with their anticipated sensory outcomes. This linkage allows for a more adaptive and flexible behaviour, as imagining or perceiving a desired outcome can automatically activate the corresponding motor program required to achieve that outcome. Thus, ideomotor theories offer a robust cognitive model for understanding how actions are planned and executed, emphasizing the reciprocity between sensory experiences and motor activities, and they have been supported by a wide range of empirical studies spanning cognitive psychology and neuroscience (Hommel et al., 2001).

Motor Schema Theory, originally proposed by Richard A. Schmidt in 1975, postulates that actions are represented in the form of generalized motor programs or schemas. These schemas serve as abstract templates that guide motor behaviour by encapsulating the invariant features of an action, such as the sequence and relative timing of movements, while allowing for flexibility through adjustable parameters like force and direction (Schmidt, 1975). According to the theory, executing a particular movement involves activating a relevant motor schema and specifying its open parameters based on the current task and environmental context. This enables the production of a wide range of specific movements using a limited set of generalized representations, thus offering an efficient mechanism for motor control. The theory has been influential in the fields of motor learning and control, and it has been supported by empirical research demonstrating that individuals can adapt generalized motor schemas to novel tasks and conditions. Its strength lies in its ability to account for both the stability and flexibility observed in human motor behaviour, providing a unifying framework for understanding how actions are mentally represented and executed (Schmidt, 1975).

The Affordance Competition Hypothesis is a model that aims to explain how actions are selected and executed in complex, dynamic environments. Developed by Paul Cisek, this hypothesis posits that multiple action possibilities, or "affordances," are simultaneously processed and represented by the nervous system whenever an individual encounters objects in the environment. These affordances inherently provide various potential actions that can be performed with regard to the object— for instance, a cup on a table affords gripping, lifting, or pushing (Cisek, 2007). According to the hypothesis, these affordances compete for selection based on a variety of factors, including current task goals, attentional focus, and prior experience. Essentially, the action that is most consistent with the individual's goals and the current contextual factors will "win" this competition and be executed. This competition is thought to occur at multiple neural levels, involving both cortical and subcortical structures, and enables rapid and flexible responses to changing environmental conditions (Cisek, 2007; Cisek & Kalaska, 2010). The Affordance Competition Hypothesis offers an integrated framework that accommodates both the reactive and planned aspects of action. It accounts for how immediate perceptual cues and longer-term goals can jointly influence action selection. Moreover, this hypothesis has received empirical support from neurophysiological studies, which show that multiple potential actions are represented in parallel in motor-related areas of the brain and that these representations are modulated by various factors such as attention and task demands (Cisek & Kalaska, 2005).

Internal Models, deeply rooted in principles from control theory, propose that the brain employs computational frameworks to manage interactions with the surrounding environment. These models primarily consist of forward and inverse components, each serving a distinct function in action planning and control. Forward models predict the sensory consequences of motor commands, acting as a simulator that allows the system to anticipate the outcomes of potential actions before they are executed. This predictive ability is crucial for tasks that require rapid adjustments, as it circumvents the latency inherent in sensory feedback loops (Wolpert & Ghahramani, 2000). Inverse models, on the other hand, are tasked with generating the motor commands necessary to achieve specific sensory outcomes. In other words, given a desired state or goal, the inverse model determines the set of motor commands that will realize this state. This is particularly useful for actions that require high precision, such as reaching for an object or articulating speech (Wolpert & Kawato, 1998). Both types of internal models interact cohesively, enabling a comprehensive and adaptable action control system. For example, a forward model could be used to simulate the outcomes of several possible actions generated by an inverse model, allowing for an informed selection based on predicted consequences. Importantly, these internal models are not static; they are updated and refined through experience and sensory feedback, ensuring an adaptive interaction with a dynamic environment (Kawato, 1999). The concept of internal models has gained substantial empirical support from neuroscience research, particularly studies that use neuroimaging techniques to identify brain regions associated with predictive coding and motor planning. Such research highlights the role of areas like the cerebellum in the instantiation and updating of internal models (Imamizu et al., 2000).

The Two-Visual-System Hypothesis, proposed by Milner and Goodale in 1995, posits that the brain has two specialized pathways for processing visual information, each serving distinct functions related to action representation. The ventral stream, often termed the "what" pathway, processes object characteristics such as shape, colour, and texture, thereby contributing to object identification and recognition. This stream primarily extends from the primary visual cortex to the inferior temporal cortex and is crucial for tasks that require conscious perception and recognition (Milner & Goodale, 1995; Goodale & Milner, 1992). Conversely, the dorsal stream, also known as the "where" or "how" pathway, is responsible for guiding actions in space. It extends from the primary visual cortex to the posterior parietal cortex and is involved in spatial processing, motion detection, and action planning. Unlike the ventral stream, which facilitates conscious recognition, the dorsal stream operates largely outside of conscious awareness to guide real-time actions, such as grasping or navigating (Milner & Goodale, 1995). The Two-Visual-System Hypothesis offers a framework to understand how the brain efficiently segregates and integrates visual information for different functional needs. According to this model, the ventral stream is optimized for creating a stable perceptual representation of the world, while the dorsal stream is specialized for transforming visual information into coordinated motor responses. This functional segregation allows for the parallel processing of visual information, enabling complex interactions with the environment (Goodale & Westwood, 2004). Empirical evidence for the Two-Visual-System Hypothesis has come from a variety of sources, including neuroimaging studies, lesion studies, and psychophysical experiments. Such research has demonstrated, for example, that damage to the ventral stream impairs object recognition but leaves action-guiding abilities largely intact, and vice versa for damage to the dorsal stream (Milner et al., 1991; Goodale et al., 1991).

The notion of a third visual pathway emerged as researchers found evidence for neural circuits that seem to be specialized for functions not easily categorized under the "what" or "where" dichotomies. For instance, some studies have suggested that there may be a separate pathway that deals with the "why" question, meaning it processes information related to the intentions or goals associated with observed actions (Mishkin & Ungerleider, 1982; Rizzolatti et al., 2001). Other postulations include a pathway for "non-conscious" vision, or one specifically dedicated to processing social cues (Allison, Puce, & McCarthy, 2000). Recently, Pitcher and Ungerleider (2021) proposed a third visual pathway that would be responsible for processing dynamic social cues and encompasses various higher socio-cognitive functions, involving tasks such as recognizing facial expressions, discriminating eye gaze, integrating audiovisual speech, and interpreting the actions and behaviours of other biological organisms (Pitcher & Ungerleider, 2021).

1.1.3. Neural Representation of Action.

Our understanding of how actions are represented in the brain has expanded dramatically with advances in neuroimaging. This section aims to give an overview of various neuroanatomical structures which had been shown to be involved in action representation. The history of identifying neural correlates for actions dates back to the late 19th and early 20th centuries, with seminal work carried out by researchers such as David Ferrier and Wilder Penfield. Ferrier, in the 1870s, used electrical stimulation on the cerebral cortices of animals to map out motor functions, providing one of the earliest insights into how actions may be neurally represented (Ferrier, 1876). However, these early investigations gained considerably more detail and rigor with the work of Penfield in the mid-20th century. Wilder Penfield, a neurosurgeon, performed a series of groundbreaking experiments where he electrically stimulated the cerebral cortices of conscious human patients undergoing surgery for epilepsy. Through this, Penfield created detailed cortical maps, including those of the primary motor cortex, demonstrating the localization of function for various motor actions (Penfield & Boldrey, 1937). These studies laid down the foundational understanding that specific regions of the cortex are linked to particular motor outputs, effectively acting as an early form of neural representation for actions.

The discovery of mirror neurons in the 1990s provided further complexity to the understanding of how actions are represented in the brain, not just in the context of execution but also in observation and understanding (Di Pellegrino et al., 1992; Rizzolatti & Craighero, 2004). The premotor cortex (PM), particularly the ventral part, serves as a home for mirror neurons, which are essential for understanding and imitating actions (Rizzolatti & Craighero, 2004). This area works closely with the inferior parietal lobule, another region rich in mirror neurons, to integrate sensory and motor information for action representation. Adjacent to the premotor cortex lies the primary motor cortex (M1), responsible for the execution of voluntary movements. It is the final cortical station that translates action plans into motor commands that orchestrate muscle contractions (Toga & Mazziotta, 2002). This cortex often communicates with the basal ganglia, which are crucial for action selection and initiation (Graybiel, 2008). The basal ganglia (BG) work in a loop with various cortical regions to fine-tune motor control and are considered pivotal in shaping both voluntary and habitual actions.

The supplementary motor area (SMA) adds another layer to action representation by being actively involved in the planning and coordination of complex movements. This includes generating sequences of actions and may involve functions ranging from simple motor preparation to higher-order tasks like planning (Nachev et al., 2008). The posterior parietal cortex is another important region that engages in action planning by transforming sensory input, particularly visual and somatosensory information, into action plans. This transformation process often involves collaboration with the frontal motor areas like the SMA and the premotor cortex (Andersen & Cui, 2009). Furthermore, the cerebellum, long associated merely with motor control, has come to be understood as instrumental in action prediction. It plays a role in maintaining internal models that anticipate the sensory outcomes of actions (Wolpert et al., 1998). The prefrontal cortex, particularly the dorsolateral region, also weighs in on action representation by engaging in higher-order planning and decision-making processes that ultimately inform motor outputs (Badre & D'Esposito, 2009).

Basal Ganglia and Cerebellum play crucial roles in motor control, procedural memory, and coordination of movements. The involvement of both structures in procedural memory has been emphasized in multiple studies. Doyon et al. (2009) suggest that both Basal Ganglia and Cerebellum are engaged during the learning of motor sequences and contribute to different aspects of motor skill acquisition. The Basal Ganglia are implicated in a variety of functions including motor control, executive functions, and behaviours. Regarding motor control, they are crucial for both initiating voluntary movements and inhibiting involuntary ones. A review by Graybiel (2008) discussed the Basal Ganglia's role in habit formation and procedural learning, which are crucial for efficient motor control. Moreover, a study by Klaus et al. (2019) delved into the mechanisms by which the Basal Ganglia contribute to movement selection and initiation. The Cerebellum has been traditionally associated with the fine-tuning of motor actions and is believed to store internal models that predict the sensory consequences of motor commands (Wolpert, Miall, & Kawato, 1998). Recent research by Sokolov et al. (2017) extends this role to include contributions to cognitive functions such as working memory, attention, and emotional regulation, suggesting a more comprehensive role than just motor coordination. Therefore, both the Basal Ganglia and the Cerebellum are vital for the efficient execution and coordination of movements, the storage of procedural memories, and ensuring smooth motor output. Advanced neuroimaging techniques and computational models continue to elaborate on the distinct yet complementary roles these structures play in neural circuitry underlying motor control. Lastly, the dual-stream hypothesis identifies the dorsal and ventral visual streams as pivotal in action representation. The dorsal stream is oriented toward spatial awareness and the real-time guidance of actions, while the ventral stream is geared toward object identification and recognition (Milner & Goodale, 1995).

The recent *third visual pathway* model proposed by Pitcher and Ungerleider (2020) extends on the lateral surface of the brain, from the primary visual cortex (V1) to the middle temporal area (MT/V5) and to the posterior/anterior superior temporal sulcus (p/aSTS). This third visual pathway takes into account two decades of work in the domain of visual social perception, specialized in neural computations of dynamic cues necessary for social perception (Pitcher & Ungerleider, 2020). These structures often work in concert, interconnected through complex neural networks, to enable the diverse range of human actions. Understanding how these areas interact to represent actions is an ongoing area of research, employing a variety of methodologies from neuroimaging to lesion studies.

Action Observation Network (AON). The development of functional magnetic resonance imaging (fMRI) techniques further facilitated the identification of a network of brain regions, now referred to as the Action Observation Network (AON), which are activated during action observation and are thought to serve similar functions as the mirror neuron system in macaques (Rizzolatti & Craighero, 2004; Rizzolatti & Sinigaglia, 2010). The Action Observation Network (AON) is a neural system implicated in the perception and understanding of actions performed by others (Rizzolatti & Sinigaglia, 2010; Molenberghs, Cunnington, & Mattingley, 2012). This network encompasses multiple brain regions, such as the inferior frontal gyrus (IFG) (Rizzolatti & Craighero, 2004), the posterior parietal cortex (PPC) (Caspers et al., 2010), and the superior temporal sulcus (STS) (Allison, Puce, & McCarthy, 2000), among others. The AON is posited to serve as a neural basis for action understanding, imitation, and possibly even for the comprehension of others' intentions and emotions (Rizzolatti & Sinigaglia, 2010; Caspers, Zilles, Laird, & Eickhoff, 2010). A key feature of the AON is its role in activating mirror neurons, which are neurons that fire both when an individual performs an action and when they observe the same or similar action performed by another (Rizzolatti & Craighero, 2004). The AON is thought to be integral to the learning of new motor skills through imitation (Cook et al., 2014). It is also implicated in social cognitive functions such as empathy and theory of mind, which involves attributing mental states to oneself and others (Keysers & Gazzola, 2014). Moreover, disruptions in the AON have been observed in conditions like autism spectrum disorder (ASD), which is characterized by impairments in social interaction and communication (Hamilton, 2013). Advancements in neuroimaging have allowed for the precise localization and functional mapping of the AON. Techniques such as functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG) have been employed to study how the AON is activated during action observation and how it interacts with other brain networks (Molenberghs, Cunnington, & Mattingley, 2012).

Despite the influential role attributed to the mirror neuron system (MNS), several researchers have raised critical perspectives challenging its centrality in action representation and embodiment. Hickok (2009) identified several problems with the mirror neuron theory, arguing that empirical data from both monkey studies and human neuroimaging do not conclusively support the necessity of mirror neurons for action understanding. Hickok presents eight specific issues, including the lack of direct evidence for mirror neuron involvement in understanding actions, discrepancies in data interpretation, and the need for alternative explanations that do not rely on mirror neuron activity (Hickok, 2009). Mikulan et al. (2015) propose a distributed network view of language and action processing, suggesting that mirror neurons might facilitate understanding without being directly responsible for action semantics. This perspective emphasizes a more holistic approach, where mirror neurons are part of a broader neural network that includes multiple regions and pathways. Such a network could support action understanding through integrative and distributed processes rather than a single, specialized system. Lingnau, Gesierich, and Caramazza (2009) highlight the variability in mirror neuron responses depending on the context and the specific actions being observed or executed. They argue that this variability challenges the notion of mirror neurons having a fixed role in action representation. Instead, their activation might be more context-dependent and influenced by other cognitive processes, suggesting that mirror neuron activity alone is insufficient to account for the complexity of action understanding.

In summary, while mirror neurons have been proposed as a crucial element in the neural representation of actions, several critical perspectives challenge this view. These critiques emphasize the need for more nuanced and comprehensive models that consider the distributed and context-dependent nature of neural processes involved in action understanding. Action representation in the brain is not localized to a single region or group of neurons, but involves a network of interconnected neuroanatomical structures. Each contributes uniquely to the different facets of action understanding, planning, selection, and execution.

1.1.4. Embodiment: Bridging the Gap.

The concept of embodiment is founded on the notion that cognition is deeply rooted in the body's interactions with the world. Several models and theories have been proposed to explain the role of embodiment in cognition, including action representation. Some early models could be considered a form of proto-embodiment theories. An example would be the *Extended Mind Hypothesis*, a philosophical perspective suggesting that cognitive processing is not confined to the brain but extends into the body and the environment, making the entire system a cognitive unit (Clark & Chalmers, 1998). Another example is *Sensorimotor Contingency Theory,* which suggests that perception is constituted by an individual's mastery of sensorimotor contingencies, the laws that govern the sensory changes brought about by motor actions (O'Regan & Noë, 2001). Another good example is the *Action-Specific Perception Model,* which proposed that perception is tailored for action, so that what we perceive in the environment is directly influenced by our capacity to act upon it (Proffitt, 2006).

In the early 1990's, mirror neurons changed everything. The discovery was initially made in macaque monkeys while the researchers were studying the neural mechanisms of motor action. They observed that specific neurons in the ventral premotor cortex (area F5) fired not only when the monkey performed a specific action, such as grasping an object, but also when the monkey observed another individual (either another monkey or a human) performing the same action (Di Pellegrino et al., 1992; Rizzolatti et al., 1996). There was direct evidence of neurons in motor areas firing when observing actions performed by another, suggesting a neurobiological basis for embodiment (Rizzolatti & Craighero, 2004). *Mirror Neuron System Theories* leveraged neurophysiological findings about mirror neurons to argue that the neural substrates for action, perception, and cognition are inherently linked.

Embodied Cognition Theory proposed that cognitive processes are intrinsically linked to and grounded in bodily experiences, a departure from traditional views of cognition as abstract and detached from the physical world (Barsalou, 1999; Wilson, 2002). According to this perspective, the very nature of an individual's bodily interactions with the environment plays a pivotal role in shaping perceptual experiences and cognitive functions (Glenberg, 1997). The theory extends its influence on a range of cognitive phenomena, from conceptual understanding to memory processes. For instance, Barsalou (1999) suggests that cognitive representations are not merely abstract symbols but are perceptual simulations that engage the same neural systems used in sensory and motor processes. This theory is supported by neuroscientific evidence showing activation in sensory and motor regions during cognitive tasks (Gallese & Lakoff, 2005). The concept also relates to language comprehension, as words and phrases can activate sensory-motor regions of the brain, implicating a deep-seated relationship between language and embodiment (Pulvermüller, 2013). Furthermore, embodied cognition has implications for memory, positing that memories are richer and more easily retrievable when associated with sensory experiences (Versace et al., 2014). Overall, Embodied Cognition Theory has brought a paradigm shift by integrating the body into the cognitive process, demonstrating that cognition is not just a function of the mind but is fundamentally shaped by bodily interactions with the world.

Grounded Cognition serves as an extension and refinement of Embodied Cognition, incorporating additional elements that emphasize the rootedness of cognitive processes in bodily experiences (Barsalou, 2008; Pezzulo et al., 2012). While Embodied Cognition posits that cognitive processes are deeply tied to bodily interactions with the environment, Grounded Cognition goes further by highlighting the roles of simulations, bodily states, and situated action in cognitive function (Barsalou, 2008; Glenberg & Gallese, 2012). Simulations refer to the mental reconstructions of sensory and motor experiences, and these serve as a fundamental mechanism for various cognitive tasks such as object recognition, problem-solving, and even language comprehension (Barsalou, 2008; Zwaan, 2014). Bodily states are viewed not merely as byproducts of cognitive processes but as integral elements that inform emotional and cognitive evaluations (Niedenthal, 2007). For instance, physiological changes, such as heart rate or muscular tension, can shape and inform cognitive appraisals and decision-making (Bechara et al., 1997). Situated action refers to the contextual dependency of cognitive processes, proposing that cognition is not an isolated, decontextualized mechanism but is directly shaped by its environmental and social contexts (Clark, 1997; Robbins & Aydede, 2009). Grounded Cognition argues that knowledge is not an abstract construct stored in the mind, but is instead grounded in the sensorimotor systems, thereby uniting cognition with action and perception (Barsalou, 2008).

Embodiment is a broad area of study that suggests that motor resonance participates in several higher order processes. However, the exact role played by specific brain structures and/or actual mirror neurons in these processes varies greatly across theories and authors. Critics of the Mirror Neuron System (MNS) theories have raised substantial concerns regarding the oversimplification of the role of mirror neurons in embodiment and cognitive processes. While early interpretations of mirror neurons suggested a direct link between action understanding and motor simulation (Rizzolatti & Sinigaglia, 2010), subsequent research has highlighted the complexity of the MNS and its interactions with other neural systems (Hickok, 2009). Furthermore, the specificity of mirror neurons has been questioned, as studies have demonstrated variability in their response properties and functional roles across different contexts and tasks (Kilner et al., 2007).

It has been argued that attributing cognitive functions solely to mirror neurons neglects the broader neural networks and cognitive processes involved in embodiment (Grafton & Hamilton, 2007). For instance, while mirror neurons may play a role in action understanding, they do not fully account for the rich perceptual experiences and cognitive processes involved in embodied cognition (Hommel et al., 2009). Additionally, the assumption that mirror neurons directly translate observed actions into motor representations overlooks the influence of higher-order cognitive processes, such as intention understanding and contextual knowledge, in action perception (Jacob & Jeannerod, 2005).

While the discovery of MNs has significantly advanced the field of cognitive neuroscience, critical perspectives urge caution against overreliance on MNs as the sole explanation for embodied cognition. These critiques highlight the importance of integrating MN research with broader cognitive and neural frameworks to develop a more comprehensive understanding of action representation and embodiment. This integrative approach ensures a balanced view that recognizes the contributions of both MNs and other neural systems in cognitive processes.

Challenges and Future Directions. While great strides have been made in understanding action representation, challenges persist. For instance, how specific are these neural representations? Do they cater to general categories of actions or specific, nuanced movements? Furthermore, the debate between localized vs. distributed representations of action continues to be a contentious topic.

In conclusion, it could be said that action representation embodies the intricate interplay between cognitive models and neural mechanisms. The embodiment perspective offers a unique lens, suggesting that our very physicality deeply influences the ways in which we mentally represent and understand actions.

1.2. Body perception & movement perception

The human body, with its intricate dance of muscles, bones, and neurons, is a marvel in motion. This section explores two interconnected realms: how we perceive our own bodies (body perception) and how we interpret the movement of bodies, both our own and others' (movement perception). Together, these domains shed light on the profound ways in which humans navigate and understand the world.

1.2.1. Neural representation

The history of identifying brain areas representing a topographic map of our body dates back to the late 19th and early 20th centuries, with seminal work carried out by researchers such as David Ferrier and Wilder Penfield. Ferrier, in the 1870s, used electrical stimulation on the cerebral cortices of animals to map out motor functions, providing one of the earliest insights into how actions may be neurally represented (Ferrier, 1876). However, these early investigations gained considerably more detail and rigor with the work of Penfield in the mid-20th century.

Wilder Penfield, a neurosurgeon, performed a series of groundbreaking experiments where he electrically stimulated the cerebral cortices of conscious human patients undergoing surgery for epilepsy. Through this, Penfield created detailed cortical maps, including those of the primary motor cortex, demonstrating the localization of function for various motor actions (Penfield & Boldrey, 1937). These studies laid down the foundational understanding that specific regions of the cortex are linked to particular motor outputs, effectively acting as an early form of neural representation for actions. Since 1937's study "Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation", we have been looking at an incomplete picture, the sensorimotor representation of the "man" instead of human. Although, in 2013, Paula M. Di Noto, Leorra Newman, Shelley Wall and Gillian Einstein published an article addressing how limited this representation still is, and clustered evidence pointing towards the "Hermunculus" (Di Noto et al., 2013). This study brings neurophysiological reports since Penfield, together with more recent neuroimaging studies, proposing two main candidates for the representation of female somatosensory cortex (see figure XX, adapted from Di Noto et al., 2013).

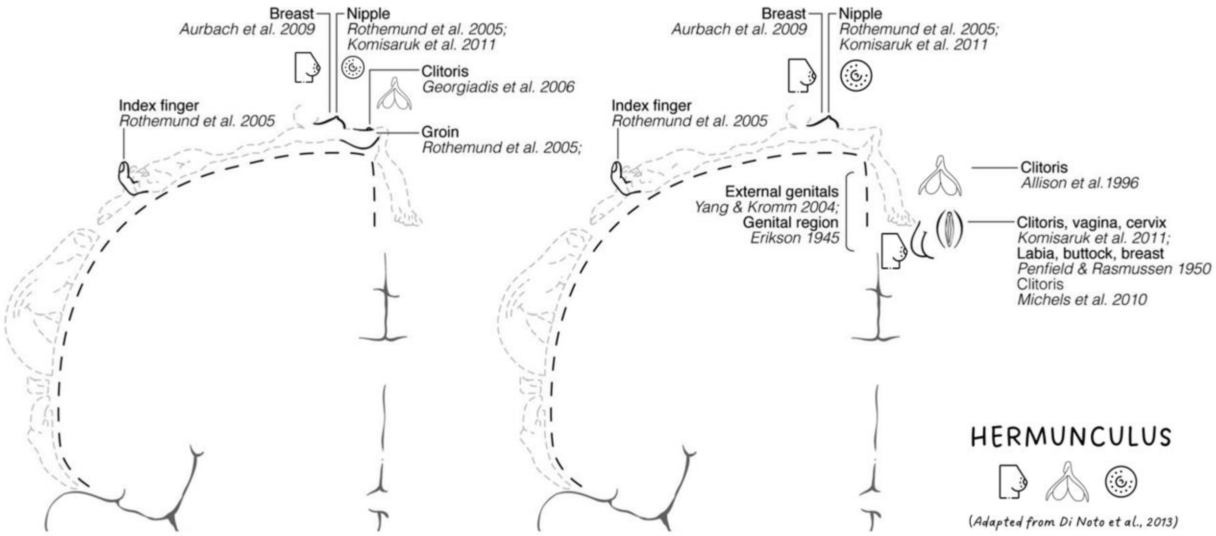


Figure XX (Fig. 0). Hermunculus: a more inclusive body representation in the brain. Two versions of the somatosensory homunculus were elaborated by Di Noto and colleagues (2013), each based on human female mapping research and showing varying placements for the genital region. For contextual understanding, Penfield's homunculus was outlined in grey dashes, highlighting the uncharted areas of the female somatosensory cortex (adapted from Di Noto et al., 2013).

1.2.2. Body Perception: An Overview

Body perception goes beyond mere physical awareness; it's the mental representation and understanding of one's own body. It refers to the multifaceted cognitive and neural processes through which individuals become aware of and interpret bodily signals, position, and appearance. At the heart of body perception is the concept of body schema, an organized model of the body in space that assists in motor control and action planning (Gallagher, 2005). This schema is not static but continuously updated through sensorimotor experiences. Body perception is shaped through various sensory modalities including proprioception, the sense of body position; interoception, the sense of internal bodily states; and exteroception, the sense of external stimuli like touch and temperature (Craig, 2002). These sensory signals are integrated in regions like the somatosensory cortex, posterior parietal cortex, and insular cortex (Ehrsson, 2012).

Body perception is also closely related to the phenomenon of embodiment, where cognitive functions are influenced and shaped by the body's interactions with the environment (Barsalou, 1999). The theory of embodied cognition suggests that body perception is not just a passive process but also actively shapes cognitive functions, such as attention, memory, and even social cognition (Niedenthal et al., 2005). In clinical contexts, disruptions in body perception are observed in a range of conditions such as anorexia nervosa, body dysmorphic disorder, and certain neuropsychological syndromes like somatoparaphrenia and hemispatial neglect (Feusner et al., 2010; Vallar & Ronchi, 2009). Technological advances, like neuroimaging, have further enriched our understanding of the neural substrates underlying body perception, with the insular cortex, for example, being implicated in interoceptive awareness (Craig, 2009).

Proprioception: The Sixth Sense.

Proprioception refers to the sensory modality that enables individuals to perceive the position and movement of their body parts in space, independent of visual input. Originating from Latin roots "proprius," meaning "one's own," and "capio," meaning "to take or grasp," proprioception is fundamental to a host of activities that require coordination and spatial awareness (Sherrington, 1907). This form of perception is mainly facilitated through specialized sensory receptors known as proprioceptors, which are located in muscles, tendons, and joints. Proprioceptors provide continuous feedback to the central nervous system about changes in muscle length, tension, and joint angles (Proske & Gandevia, 2012). These signals are integrated within the central nervous system, particularly in the cerebellum, posterior parietal cortex, and the somatosensory cortex, to form a coherent representation of body position (Wolpert, Goodbody, & Husain, 1998).

Proprioception is crucial for a range of functions, from basic ones like posture control and locomotion to more complex actions like coordinated movements and skilled tasks. It also plays a critical role in the planning and control of movement, serving as a core feedback mechanism for motor learning and adaptation (Shadmehr & Krakauer, 2008). Deficits in proprioception can arise due to neurological conditions such as stroke, peripheral neuropathies, or musculoskeletal disorders and can significantly impact motor skills and daily functioning (Rothwell et al., 1982).

1.2.3. Movement Perception

The ability to discern movement is critical, whether for interpreting a dancer's performance or predicting a predator's trajectory. Movement perception encompasses the recognition, interpretation, and prediction of motion in both animate and inanimate objects, including one's own body. This perception is integral to a wide range of activities, from basic navigation to complex social interactions. Visually, movement perception is largely mediated by specialized motion-sensitive cells found in the primary visual cortex (area V1) and other specialized areas like V5/MT (middle temporal area) (Born & Bradley, 2005). The phenomenon is not merely a bottom-up process triggered by retinal inputs but involves top-down influences like attention, expectation, and prior knowledge (Treue, 2003). For example, the interpretation of biological motion, such as the movements of other humans, is influenced by both low-level visual cues and higher-level cognitive processes, including theory of mind and social cognition (Blake & Shiffrar, 2007).

In the field of multisensory integration, movement perception is often a product of converging information from various sensory modalities. Besides visual information, auditory, vestibular, and proprioceptive inputs contribute to the accurate perception and interpretation of motion (Angelaki & Cullen, 2008; Butler et al., 2011). This multisensory integration takes place in specialized brain regions such as the superior colliculus and posterior parietal cortex (Stein & Stanford, 2008). Clinically, impairments in movement perception can manifest in several conditions, including akinetopsia, a rare neurological disorder where an individual loses the ability to perceive motion (Zeki, 1991).

Biological motion refers to the specialized perception and interpretation of motion cues that originate from biological entities, such as humans and other animals. It involves the ability to recognize and understand complex movements like walking, running, or dancing based on often sparse visual information (Johansson, 1973). These cues may consist of point-light displays (PLD), where only the joints or other key features of the moving entity are visible, and yet observers can readily discern the type of movement and attribute it to a biological source (Blake & Shiffrar, 2007). The study of biological motion is not merely an academic exercise but serves a host of adaptive functions. It plays a vital role in social cognition, enabling individuals to identify actions, intentions, and emotions of others (Heberlein, Adolphs, Tranel, & Damasio, 2004). Moreover, the ability to perceive biological motion has been found to be present from early infancy, suggesting that it may be innate to some extent (Fox & McDaniel, 1982). Neuroscientific investigations have pinpointed several brain regions involved in the perception of biological motion, including the superior temporal sulcus (STS), the fusiform body area (FBA), and the extrastriate body area (EBA) (Grossman et al., 2000; Downing et al., 2001). The STS, in particular, has been highlighted for its role in interpreting the 'animacy' and 'intentionality' of observed movements (Allison, Puce, & McCarthy, 2000). Clinical research suggests that impairments in biological motion perception can be symptomatic of various conditions, such as autism spectrum disorders, where individuals may struggle to correctly interpret social cues from body movements (Kaiser & Shiffrar, 2009). In conclusion, humans are extremely good at recognizing movements specific to living organisms. The PLDs of human walkers, despite their simplicity, are effortlessly recognized, highlighting our innate sensitivity to biological motion.

Interconnection Between Body and Movement Perception. Both these domains are deeply intertwined. A dancer's understanding of their own body influences their interpretation of another's dance. Similarly, observing others can refine one's own motor skills. From a theoretical perspective, body perception provides the baseline, spatially and temporally, upon which movement is understood (Blake & Shiffrar, 2007). That is, a coherent sense of one's body—its size, orientation, and position—forms the context for perceiving and interpreting both self-generated and externally observed movement (Ehrsson, 2007).

In terms of empirical significance, research in cognitive neuroscience has shown that several brain areas participate in both body and movement perception. These include the posterior parietal cortex (PPC), superior temporal sulcus (STS), and extrastriate body area (EBA), among others (Downing et al., 2001; Grossman et al., 2000). For instance, the STS plays a role in recognizing biological motion, while also contributing to the understanding of body-related social cues (Allison, Puce, & McCarthy, 2000). Meanwhile, the EBA is involved in the visual processing of body parts and is activated during the observation of human motion (Astafiev et al., 2004). Importantly, these neural systems do not work in isolation but are interconnected. For example, the perception of body orientation can influence the perceived direction of movement (Lenggenhager et al., 2007). Further, the phenomenon of "motor resonance" suggests that observing an action activates similar neural circuits as performing that action, indicating a link between action representation and body perception (Rizzolatti et al., 2001).

In summary, body and movement perception are mutually influential and share overlapping neural substrates. Their interconnection serves adaptive functions, facilitating not only motor coordination but also social interaction and communication. Given the centrality of these perceptual processes in human cognition, their integrated study remains a crucial area of research in cognitive psychology and neuroscience, highlighted by embodiment perspectives.

1.3. Attention: bottom-up and top-down modulations

Human perception is a dance between the raw data our senses provide and the interpretations our brain imposes. The act of selecting specific information for further processing is commonly known as attention. This chapter will introduce the main models of attention relevant to actions and explore the intricate interplay between bottom-up sensory processes, where information flows from our sensory organs to higher brain regions, and top-down modulation, where higher cognitive functions influence the processing of sensory information.

Attentional models serve as frameworks to understand how organisms allocate their resources to process specific pieces of information from their environment. Attention models have evolved significantly from early Broadbent's Filter Model (1958) and Kahneman's (1973) Capacity Model, to nuanced theories that take into account multiple cognitive and neurological factors. The seminal 1986 work "Attention to Action: Willed and

Automatic Control of Behavior" by Donald A. Norman and Tim Shallice introduced a two-tiered framework for understanding attention and action control. The model delineated two systems: Contention Scheduling and the Supervisory Attentional System (SAS). Contention Scheduling operates at a lower cognitive level to manage routine, automatic tasks using a set of competing "schemas", which are activated (via a "trigger data base") by relevant sensory input. Conversely, the SAS, influenced by Shallice's earlier work on executive functions (Shallice, 1982), operates at a higher cognitive level, coming into play during novel or conflict circumstances requiring deliberate attention and goal-directed action. This model had the unique contribution of highlighting the dual nature of attention, incorporating both automatic (bottom-up) and goal-directed (top-down) processes. This model has also inspired empirical research aiming to locate the neural substrates of the SAS, typically identifying regions within the prefrontal cortex as critical nodes (Miller & Cohen, 2001). In summary, Norman and Shallice's 1986 model provides a robust framework for understanding the automatic and controlled aspects of attention and action.

1.3.1. Interplay between Bottom-up and Top-down Processes

Since Norman & Shallice's model, attention has been often divided into two major types: stimulus-driven (also known as exogenous or bottom-up) and goal-directed (endogenous or top-down) attention. Stimulus-driven attention is activated by conspicuous features of external stimuli, leading to automatic processing of that information. On the other hand, goal-directed attention originates from internal cognitive processes, where information is intentionally sought based on individual goals or preferences (Connor et al., 2004; Corbetta & Shulman, 2002; Itti & Koch, 2001). While attention can be directed toward various sensory modalities through either stimulus-driven or goal-directed mechanisms, research has primarily focused on attention within the visual system (for a review see Katsuki & Constantinidis, 2014).

Early research in psychophysics indicated that in searches dictated by stimulus-driven (bottom-up) factors, targets readily "pop out" when they possess distinct features like colour or orientation compared to the background. These targets can be processed simultaneously, obviating the need to scan each element individually (Duncan & Humphreys, 1989; Treisman & Gelade, 1980). Conversely, when a search is driven by goal-directed (top-down) factors and there is little feature-based distinction between target and background, a deliberate, element-by-element scrutiny is required to identify the target (Wolfe & Horowitz, 2004). While these two modes of attention are often discussed as separate entities and are thought to engage different neural circuits and anatomical structures, they typically interact in real-world experiences, as highlighted in models of visual search (Wolfe, 2010). Studies involving patients with cortical lesions have further suggested specialized roles for subregions of the parietal cortex in these different attentional systems, but also underscored their interconnectivity (Shomstein, 2012; Shomstein et al., 2010).

Theoretical models often posit that attention is allocated to the most prominent stimulus, location, or feature generating the strongest neural response (Desimone & Duncan, 1995; Koch & Ullman, 1985; Wolfe, 1994). This idea often manifests as a "saliency map", where each feature type is processed in its own map, and these are subsequently integrated into a global saliency map. Attention is supposed to be oriented to the area on the map with the highest activation (Koch & Ullman, 1985). However, attention is not solely determined by inherent stimulus saliency; it's also influenced by its relevance to ongoing objectives, as emphasized in early visual search models (Wolfe, 1994). This has led to the proposal of a "priority map" that combines both stimulus-driven and goal-directed factors (Bisley & Goldberg, 2010; Serences & Yantis, 2006). Even when considered separately, both modes of attention often activate the same network of parietal and prefrontal cortical areas, lending support to the concept of a priority map influenced by both stimulus-driven and goal-directed factors (Katsuki & Constantinidis, 2012b).

1.3.2. Neurological Insights

Bottom-up attention in the visual domain initiates with elementary visual processing that takes place along the neural pathways of the visual cortex. Originating from the primary visual cortex (V1), feed-forward neural signals move upwards to various cortical regions, splitting into two primary visual streams: the ventral pathway, which focuses on object identification and feature recognition, and the dorsal pathway, concerned with spatial orientation and motion perception (Ungerleider & Haxby, 1994; Ungerleider & Mishkin, 1982). These cortical streams are arranged in a hierarchical fashion, where receptive fields expand and functional attributes become increasingly intricate as neural inputs travel from the early to the later stages of these pathways (Felleman & Van Essen, 1991). Consequently, bottom-up information capable of directing attention flows from the visual cortex toward the prefrontal cortex (PFC). In contrast to the data-driven nature of bottom-up processing, top-down modulation is conceptually driven. By directing our attention, we can enhance the processing of specific sensory information. It is a hypothesis shaping the way we will process and interpret the upcoming data. The primary impact of top-down attention is an enhancement of neural activity for the specific location, feature, or object that is behaviourally pertinent, while concurrently supressing neuronal responses to stimuli considered irrelevant. Such modulatory effects on neuronal responses have been reported in both the ventral and dorsal visual streams (Noudoost et al., 2010). As one ascends the cortical hierarchy, the degree of top-down influence on neuronal firing rates generally escalates (Cook & Maunsell, 2002; Luck et al., 1997). While primary sensory areas deal predominantly with bottom-up information, association areas in the cortex are crucial for top-down modulation. For instance, Gilbert and Li (2013) discuss how primary sensory areas are the initial recipients of bottom-up sensory inputs. In contrast, association areas, particularly in the prefrontal cortex and parietal lobe, are key players in top-down modulation of sensory information.

In a research experiment that employed neurophysiological recordings within the prefrontal cortex (PFC) and the posterior parietal cortex (PPC) during a visual search task, the temporal patterns of neuronal activity were examined (Buschman & Miller, 2007). The study revealed that neurons in the PPC exhibited shorter latencies when representing salient or "pop-out" stimuli compared to neurons in the PFC. This suggested that visual saliency is primarily represented in the PPC, which then transmits this information to the PFC. The experimental design by Buschman and Miller (2007) provided subjects with information about the target identity prior to initiating the search phase. With this top-down cue in place, lower cortical regions could more efficiently filter the attributes of incoming stimuli, thus expediting the representation of the target. This mechanism is at the heart of top-down control over attention (Chelazzi et al., 2001; Herrington & Assad, 2009; Luck et al., 1997).

Other neurophysiological studies have indicated a role for neuronal synchronization in the mechanics of attention (Bichot et al., 2005; Fries et al., 2001; Saalmann et al., 2007). Enhanced synchronization, particularly in the gamma (35- to 70-Hz) frequency range, has been documented in neurons selectively responsive to stimuli that are the focus of attention. This has been noted in tasks that require both spatial and object-based attention (Bichot et al., 2005; Fries et al., 2001; Taylor et al., 2005). Synchronization in the gamma band is theorized to serve as a means to tag information related to an attended stimulus, thereby differentiating it from stimuli that are not the focus of attention. γ-band synchronization has been proposed as a mechanism to label information about an attended stimulus and distinguish it from unattended stimuli (Katsuki & Constantinidis, 2013).

1.3.3. The Influence of Experience and Learning

A trained musician or a sommelier may process sensory data differently due to their training, showcasing the influence of experience on both bottom-up and top-down processes. Feedback loops allow higher brain regions to send signals back to primary sensory areas, modulating perception. A classical study integrating this dynamic is the one by Kilner et al. (2007), where it is discussed the mirror neuron system, indicating that both top-down and bottom-up processes might be at play in predicting others' actions based on observed motor acts. Since then, there has been increasing interest in integrated models that combine both bottom-up and top-down influences for more comprehensive explanations of action prediction and understanding (Csibra, 2008). Studies have also started to examine how top-down and bottom-up factors dynamically interact during action prediction. For example, research involving predictive coding suggests that top-down expectations can modulate the sensitivity of bottom-up sensory processing pathways during action observation (Adams et al., 2013).

Bottom-up sensory processes and top-down modulation together create the rich tapestry of human perception. Each moment of perception is a dialogue between the world's raw data and the brain's interpretations. Neither works in isolation. What we expect to see, hear, or feel can profoundly shape our perceptual experiences. Previous knowledge, cultural background, and current cognitive states can all adjust the way we interpret sensory input. Consider gestalt principles or ambiguous stimuli like the "face-vase" illusion. While the sensory data provides basic shapes and lines, maybe even biologically salient features, top-down processes can guide the interpretation of these into recognizable patterns and figures. Images can be interpreted in multiple ways depending on top-down influence. Our past experiences, training and expertise can shape how we process information. The brain's ability to reorganize itself based on experiences plays a role in how sensory processing evolves over time (changing with experience). Understanding this balance is key to unravelling the mysteries of how and what we perceive, interpret, and understand from our surroundings.

1.4. Measurements & Methods

The work elaborated on this thesis, with more emphasis on the experimental chapters, has been done from a quantitative approach. Quantitative methods focus on the collection and analysis of numerical data, the use of controlled experimental designs, and the application of statistical methods to understand and interpret psychological phenomena. By using quantitative methodology, the current work aims to provide objective, generalizable, and replicable findings that contribute to the theoretical and empirical understanding of the field.

1.4.1. Behavioural methods

Accuracy is often employed as a dependent variable to evaluate the performance of participants in experimental tasks. It is calculated as the ratio of the number of correct responses to the total number of trials, usually expressed as a percentage or proportion. This measure serves as a straightforward way to assess the efficacy with which subjects can complete a given task—be it perceptual, cognitive, or motor in nature (Stanislaw & Todorov, 1999).

Reaction Times (RTs) serve as a key dependent variable to assess the speed of cognitive and motor processes in response to specific stimuli. Typically recorded in milliseconds (ms), reaction times offer a high-resolution measure of the time elapsed between the presentation of a stimulus and the participant's subsequent behavioural response, such as pressing a button or vocalizing a word. The utility of RTs lies in their ability to provide insights into the temporal dynamics of cognitive operations.

Sensitivity or discriminability (d'). In Signal Detection Theory (SDT), d' (d-prime) represents sensitivity or discriminability, quantifying the ability of an observer to differentiate between signal events (e.g., a stimulus being present) and noise events (e.g., a stimulus being absent). d' is calculated as the difference between the z-scores of the hit rate and the false alarm rate, and it serves as a measure of discriminability—how well an individual can distinguish between signal and noise. Specifically, $d' = Z(\text{hit rate}) - Z(\text{false alarm rate})$, where Z represents the z-score transformation. A higher d′ value indicates better performance in distinguishing signal from noise, meaning greater sensitivity to the stimuli (Macmillan & Creelman, 2004).

Response bias (criterion c). Criterion (c) quantifies an observer's response bias and reflects the threshold at which an observer decides that a stimulus is present or absent. Criterion c is a measure that reflects an individual's response bias, or their tendency to say 'yes' or 'no' irrespective of their discriminability. It is calculated as the negative average of the z-scores for the hit rate and the false alarm rate: $c = -[Z(\text{hit rate}) + Z(\text{false alarm rate})]/2$. A c value of zero indicates no bias, whereas positive or negative values suggest a bias towards responding 'no' or 'yes,' respectively. Criterion (c) provides essential insights into an individual's perceptual decision-making biases (Stanislaw & Todorov, 1999).

1.4.2. Psychometric methods

A variety of questionnaires were used to measure individual differences. Below a description and brief overview of all the psychometric tools used in Chapter 7.

Multidimensional Assessment of Interoceptive Awareness-2 (MAIA-2) is a psychometric tool designed to assess various facets of interoceptive awareness, which is the conscious perception of internal bodily sensations. The questionnaire is comprehensive and multidimensional, capturing different dimensions of interoceptive awareness such as noticing, emotional awareness, self-regulation, and body-listening. The MAIA-2 aims to provide an in-depth understanding of how individuals differ in their awareness and interpretation of bodily signals, which can be crucial for understanding individual differences in emotional regulation, stress response, and even susceptibility to various mental health conditions (Mehling et al., 2012).

The Goldsmiths Dance Sophistication Index (Gold-DSI) is a specialized psychometric questionnaire aimed at measuring an individual's level of sophistication and expertise in dance. It captures various components such as technical skill, understanding of dance theory, and emotional connection to the art form. By providing a nuanced view of dance-related skills and experiences, this tool allows for the examination of individual differences in dance expertise (Rose et al., 2022).

The Interpersonal Reactivity Index (IRI) is a psychometric questionnaire that measures empathy across multiple dimensions, including cognitive empathy, emotional empathy, and tendencies toward empathic concern or personal distress in social situations. IRI is widely used to assess individual variations in empathic skills and tendencies. These individual differences in empathy have implications for a range of psychological and social outcomes, including interpersonal relationships, prosocial behaviour, and even certain psychiatric conditions (Davis, 1983).

The Toronto Alexithymia Scale-20 (TAS-20) is a widely used psychometric measure designed to assess alexithymia, which is a personality construct characterized by difficulties in identifying and describing one's own emotions. The TAS-20 is a 20-item self-report questionnaire that captures three major facets of alexithymia: difficulty identifying feelings, difficulty describing feelings, and externally oriented thinking. It has been extensively validated and is applied in both clinical and research settings to understand individual differences in emotional awareness and expression (Bagby et al., 1994).

1.4.3. Neural methods

Electroencephalography (EEG) is a non-invasive neuroimaging technique used to record electrical activity generated by neurons in the brain (Niedermeyer & da Silva, 2004). The electric currents generated by active neurons create voltage differences both intra- and extracellularly, which can be recorded as electric fields (Murakami et al., 2002). These fields are influenced by factors like the alignment and properties of the neurons (cytoarchitecture) (Buzsáki et al., 2012). In EEG research, neural responses to specific events are averaged to produce ERPs. Averaging trials helps to minimize the influence of non-specific brain activity while amplifying the signal related to the task (Glaser and Ruchkin, 1976; Luck, 2014). However, there are limitations to this approach. Extracellular field recordings, like ERPs, are spatiotemporally smoothed, meaning that they don't offer precise spatial resolution (Niedermayer and Lopes da Silva, 2005; Buzsáki et al., 2012). As a result, determining the specific origin of observed brain activity remains a challenge.

Event-Related Potentials (ERPs) focus on identifying and studying the time-locked responses of the brain to particular events or stimuli (Luck, 2014). ERPs are especially useful as they allow for the exploration of temporal dynamics of cognitive processes with high temporal resolution of millisecond-level accuracy (Hillyard & Anllo-Vento, 1998). ERPs measure voltage fluctuations in the EEG signal, time-locked to the onset of the event of interest, usually a sensory stimulus or a motor response (Rugg & Coles, 1995). By averaging the EEG signal across many trials, it is possible to isolate neural components from the ongoing, stochastic background activity (Woody, 1967), thus providing a cleaner measure of brain activity associated with the cognitive or perceptual process under investigation (Dien, 2010). Different ERP components are sensitive to various cognitive processes and can be identified based on their polarity (positive or negative deflection), latency (timing), and topography (spatial distribution across the scalp) (Picton et al., 2000).

ERP components are generated in specific brain regions during mental tasks and recorded at the scalp. They are a composite of multiple underlying neural signals, making it challenging to pinpoint their exact origins (Cohen, 2014; Luck, 2014). This issue, known as the 'hard problem, affects the spatial resolution and estimations of the neural generators. The superimposed neural signals can cancel each other out, reducing the overall observed signal, especially when averaged against noise. Furthermore, the voltage propagation from different brain regions varies due to underlying brain tissues and electrode positions (Ahlfors et al., 2010; Irimia et al., 2012; Tenke and Kayser, 2012).

However, ERPs offer excellent temporal resolution, capturing components peaking as early as 20-40ms post-stimulus (Urbano et al., 1997; Giard and Peronnet, 1999). This is valuable for studying rapid information processing and making comparisons across different populations. Additionally, methods exist to mitigate some of the issues with ERP, such as examining waveforms from multiple electrodes to differentiate component contributions (Kappenman and Luck, 2012).

Transcranial Magnetic Stimulation (TMS) is a non-invasive neurostimulation technique used for both research and clinical applications. In TMS, a magnetic coil is placed over the scalp, and rapidly changing magnetic fields are used to induce electrical currents in specific regions of the brain. This induction can transiently modulate neural activity to investigate the causal relationships between neural circuits and cognitive or motor functions. The two primary forms of TMS are single-pulse and repetitive TMS (rTMS). Single-pulse TMS is often used to probe the functionality and connectivity of neural circuits in real-time. In contrast, rTMS involves the application of repetitive pulses to modulate neural activity over a more extended period. While TMS is generally considered safe, it can have side effects like transient headaches or scalp discomfort. Safety guidelines, including limits on stimulation intensity and frequency, have been established to minimize risks (Rossi, Hallett, Rossini, & Pascual-Leone, 2009).

TMS is particularly valuable for establishing causality. Unlike observational methods such as fMRI or EEG, TMS allows for the manipulation of neural activity, making it possible to move beyond correlation to establish causal links between brain and behaviour (Sack, 2006). The technique is often combined with neuroimaging methods to enable simultaneous investigation of neural and behavioural responses. These combined approaches provide more comprehensive insights into the temporospatial dynamics of brain activity (Siebner et al., 2009).

SECTION A: Working Memory (WM) and Top-Down modulations.

Working memory was first coined as a term in the 1960s by Miller, Galanter, and Pribram, in the context of information processing from a Computational Theory of Mind perspective and the well-known computer metaphor to study the brain. A few years later, Atkinson & Shiffrin (1968) were already using this term to describe the “short-term store” within their model of memory, where they claimed hippocampal lesion studies on patients with amnesia like H.M. (Milner, 1966) as compelling evidence for a separation of the short-term and long-term memories. The classical Atkinson–Shiffrin model proposed that human memory has three separate components: a sensory register, a short-term store, and a long-term store. From this early model, the idea of short-term memory evolved into working memory (WM), involving temporary maintenance of information and its manipulation (Baddeley, 2012). WM is defined as the system or mechanism underlying the maintenance of relevant information during the performance of a cognitive task (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980), and directly linked to attention. The initial multi-component model of WM (Baddeley and Hitch, 1974) included three systems: a central executive and two dependent systems storing semantic and visuo-spatial information. Overall, since its original proposal, the multicomponent model of WM has been revisited, challenged, and expanded. Newer proposals suggest additional systems to deal with a wider range of stimuli that seem to differ in the nature of the information to be stored.

Throughout this section of the thesis, and building on the general framework above, the concept of WM will be revised, challenged, updated, and forwarded to include a wider perspective integrating attentional modulations and somatosensory processing.

Chapter 2: Working memory for bodies & attentional mechanisms.

Working memory is a short-term storage system that allows for the manipulation of information needed for complex cognitive tasks such as reasoning, planning, and understanding. While a significant body of research focuses on working memory for objects, numbers, or verbal information, there is growing interest in how our brains handle more complex stimuli like human bodies. For instance, Oberauer and colleagues had worked on differentiating structures within working memory, especially focusing on the distinction between declarative and procedural memory components (Oberauer, 2009). They have posited that working memory consists of multiple components or 'slots,' each holding information either in a declarative or procedural form, highlighting the role of attentional control in determining what information gains access to the working memory system. This control is exercised in a domain-specific manner, meaning that the mechanisms for declarative information may be different from those for procedural information (Oberauer & Hein, 2012). This line of research emphasized that procedural memory in working memory is primarily concerned with the 'how-to' aspects of information, including skills and habits. They showed how procedural memories are encoded and retrieved differently compared to declarative memories (Oberauer, 2019).

Models of working memory (WM) focused on sensory recruitment propose a degree of commonality between the neural systems engaged in perception and memory retention. These models have gained support from research typically using arbitrary visuals like shapes or colours as the items to be remembered (Harris et al., 2002; Serences et al., 2009). While the perception of these simple stimuli doesn't necessitate complex, multi-regional processing, the act of encoding body related information and actions is more complex, involving fronto-parietal networks and even the individual's own body representation within the sensorimotor and somatosensory cortex (Caspers et al., 2010; Molenberghs et al., 2012; Galvez-Pol et al., 2018a; 2018b). This intricate neural process aligns with the theory of embodiment, which emphasizes the role of one's own bodily and motor representations in encoding body-related information (Calvo-Merino et al., 2005; Sel et al., 2014; Urgesi et al., 2007). Moreover, prior behavioural research focusing on working memory mechanisms for visually presented information related to the body indicates that internal sensorimotor representations play a role in both perceiving and retaining such information in memory (Shen et al., 2014; Smyth et al., 1988; Smyth and Pendleton, 1990, 1989; Wood, 2007).

In this context, working memory for bodies examines how we encode, maintain, and retrieve information related to human bodies and their movements. Empirical studies often investigate which brain areas are activated during these tasks, revealing the complex networks involved in body perception and action representation. One could even argue that working memory for bodies has evolutionary roots, as being able to quickly understand and remember bodily actions could be advantageous in both cooperative and competitive social situations.

2.1. Sensory recruitment models

Neuronal activity associated with the retention of information in working memory (WM) has been extensively characterized within the prefrontal cortex (PFC) (Fuster and Alexander, 1971; Romo and Salinas, 2003; Kostopoulos et al., 2007). Specifically, the dorsolateral region of the PFC contributes to the sustenance of information by focusing attention on internal sensory representations, which are generally localized in more posterior brain regions

(Curtis and D’Esposito, 2003). These posterior regions encompass various sensory cortices, including those responsible for visual, somatosensory, and auditory perception. Notably, these areas are implicated not only in perception but also in the encoding phase of memory. For example, the visual cortex and adjacent regions appear to be instrumental in the recognition, perception, and retention of visual information within WM (Becke et al., 2015; Tood and Marois, 2004). A similar modality-specific neural substrate for the upkeep of information in WM is observed in other sensory domains, including tactile (Zhou and Fuster, 1996; Kaas et al., 2013; Katus, Grubert, and Eimer, 2015) and auditory modalities (Arnott et al., 2005).

2.1.1. Evidence from visual working memory (VWM)

While early behavioural studies lent credence to the idea of modality-specific short-term memory stores, the role of the visual cortex in working memory was substantiated much later. This delay could be attributed to the limitations of initial neuroimaging methods, which focused solely on "univariate" analyses, which identify clusters of voxels that show similar responses (Adam et al., 2021). Early work using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) unveiled univariate markers of working memory in regions such as the parietal, frontal, and prefrontal cortex, reinforcing prior evidence of the crucial role of the prefrontal cortex in working memory, particularly its maintenance phase (Curtis & D’Esposito, 2004; Funahashi et al., 1989). Conversely, the visual cortex displayed little or no persistent univariate activity during the working memory delay, aligning with the belief that it was important mainly for perception but not for the maintenance of working memory (Adam et al., 2021).

However, advances in neuroimaging technology, including improved spatial resolution and signal-to-noise ratios, allowed researchers to adopt multivariate approaches. This led to the discovery that the activity in the early visual cortex could be used to decode specific features of visual stimuli, a feat unattainable by standard univariate methods (Kamitani & Tong, 2005). Research by Harrison and Tong (2009) and Serences et al. (2009) revealed that early visual cortex activity could indeed be employed to decipher an orientation held in working memory. These findings occurred even though there was little or no sustained univariate activity, potentially owing to the inherent dynamics of the visual cortex, where enhancing neuronal responses to attended items is offset by the suppression of other neuronal activities (Martinez-Trujillo & Treue, 2004; Serences et al., 2009). Therefore, these multivariate methods opened new avenues for understanding the role of the visual cortex in working memory, a role that might have been masked by the limitations of earlier univariate analyses.

From the EEG literature, there has been parallel evidence in support of sensory recruitment models, supporting the hypothesis that perceptual cortices with a role in perception are also involved in encoding and maintaining that information in working memory. Using change detection paradigms and looking at persistent activity over sensory cortices, increasing activity in brain areas with a role in perceptual processing has been observed between the presentation of the stimuli to-be-remembered and their recall. For instance, Vogel and Luck (2004) showed how mean amplitude of EEG waveforms are modulated according to the number of coloured lines and squares to be-remembered in a visual WM task. The authors showed how after a short presentation phase, when the stimuli disappeared, brain activity over posterior sites started to increase around 300ms after the onset of the stimuli, steadily increasing from memory load 1 to 4; load referring to the number of items to be remembered. The increase in mean amplitudes seemed to also correlate with the memory capacity of the participants and with the capability to handle the information. A follow-up paper showed that compared to participants with low memory capacity, higher performers exhibited smaller increases in brain amplitudes, presumably reflecting less depletion of computational resources under similar tasks (Vogel et al., 2005).

2.1.2. Working Memory for Actions

While traditional research has primarily focused on working memory for static objects or simple visual stimuli, understanding how actions are encoded and maintained in working memory is vital for our interactions with a dynamic world. The prefrontal cortex is often implicated in working memory tasks, including those involving action representation (Curtis & D'Esposito, 2003; D'Esposito & Postle, 2015). However, more specialized regions such as the premotor cortex and the parietal lobe have also been shown to be involved in maintaining action-relevant information (Cisek & Kalaska, 2004; Vesia & Crawford, 2012). Studies using functional magnetic resonance imaging (fMRI) and other neuroimaging techniques have examined how action-related visual stimuli are encoded and maintained. One interesting finding is that the same neural circuits involved in planning and executing actions may also be implicated in maintaining these action representations in working memory (Wadsworth & Kana, 2011; Zacks, 2008).

The complexity of the action also appears to influence how it is stored in working memory. Longer or more complicated action sequences may require different cognitive and neural resources compared to simple, singular actions (Schubotz, 2007; Zacks et al., 2007). The literature also points out that the representation in working memory may not just be a static image but could be more dynamic, evolving with the sequence of the action (Schubotz, 2007). Studies have also investigated differences between working memory for actions versus objects. Some findings suggest that actions are encoded in a more abstract, goal-directed manner, whereas objects are encoded based on their visual features (Rizzolatti et al., 2001; Umiltà et al., 2001). A study by Fiehler et al. (2008) investigated WM maintenance of kinesthetic information with blindfolded participants encoding hand movements. They found that maintenance elicited load-dependent activity, encoding hand movements activated somatosensory areas and engaged networks within the ventrodorsal stream.

A variety of body-related elements, including actions' kinematics, are typically processed in a sequential fashion. This suggests that there is a single representation and corresponding action executed at any given moment, shifting between different movements in an organized, hierarchical, and goal-oriented sequence (Endress and Wood, 2011). The cognitive processing pathways for body shapes (structured postures) and body actions (continuous movements) appear to be somewhat divergent. Vicary et al. (2014) demonstrated that the recognition of structured and continuous body movements is specifically influenced by interference that is pattern-based or motion-based, respectively.

2.1.3. Visual working memory (VWM) for bodies

Visual working memory (VWM) of body stimuli distinguishes itself from the more general VWM research, which has commonly used geometric shapes, colours, or faces as the objects of study. Neural substrates specific to body parts, such as hands or limbs, have been identified, with particular emphasis on the extrastriate body area (EBA) (Downing et al., 2001). Research also points to the involvement of the superior parietal lobule and the dorsal premotor cortex, areas often implicated in action observation and planning, when working memory tasks involve body parts (Astafiev et al., 2004; Ferri et al., 2015).

In a recent EEG study by Galvez-Pol et al (2018a), using new methodologies to isolate somatosensory components from visual carry over effects, they showed how visual activity was modulated by load when maintaining shapes, the same way somatosensory activity was modulated by load when maintaining hand images. Unlike faces, which are often processed holistically, other body parts may be encoded into VWM with different strategies. While it has been argued that it follows feature-based encoding (Reed et al., 2003), there is also evidence that body postures could be encoded holistically, depending on the task requirements (Reed et al., 2006). Clinical populations with motor disorders or body image issues may also demonstrate atypical patterns in VWM for body parts (Case et al., 2012; Urgesi et al., 2014).

Working memory for body stimuli appears to be influenced significantly by the action implied by the body parts or by the context in which they are seen. For instance, the perception of a hand holding a tool elicits different cognitive processes compared to the perception of a relaxed hand. This difference in processing is supported by research studies, such as those conducted by Urgesi et al. (2006) and Zimmermann et al. (2017), which suggest that action-related postures and the context in which body parts are observed significantly modulate working memory representations. When it comes to action-related context, a hand holding a tool might not just activate neural pathways associated with hand recognition, but also those linked to tool use and even the anticipated actions using that tool. The action implied by the posture may tap into motor-related brain regions, potentially facilitating the encoding and storage of the specific body posture in working memory. This could be mediated through mirror neuron systems, which have been implicated in understanding actions and are thought to bridge perception and action (Rizzolatti & Craighero, 2004). Regarding the broader context in which the body stimulus is placed, factors like spatial location, social interactions, or environmental cues may further modulate memory performance. For example, a hand holding a tool might be remembered differently depending on whether it is seen in a workshop or a kitchen, as the associated cognitive schemas for these environments could influence the encoding and retrieval processes in working memory (Bar, 2004).

Additionally, the role of semantic networks cannot be understated. The meaning derived from the action, or the context could interact with existing semantic knowledge, thereby affecting how the body-related stimulus is stored in working memory (Binder & Desai, 2011). In other words, the hand holding a tool may be deeply integrated into a web of meanings and associations that extend beyond the visual characteristics of the hand itself, thereby influencing its representation in working memory. Emotional or social relevance of body stimuli can also impact their processing in VWM. For instance, a study found that participants are better at remembering body postures that imply social interactions or emotional states (Aviezer et al., 2012). Attentional mechanisms also seem to play a role in working memory for actions, perhaps even more than for static objects. Top-down control signals from the prefrontal cortex could help focus attention on relevant aspects of the action, affecting both encoding and maintenance (Awh & Jonides, 2001; Nobre et al., 2004).

Finally, experience with body movements (such as dance or sports), can affect the efficiency of VWM for body stimuli (Stevens et al., 2012). Furthermore, experts in dance or sports often engage in "motor simulation," mentally rehearsing movements as a form of practice (Jeannerod, 2001). This cognitive practice could strengthen the neural circuits involved in the visual and motoric processing of body movements, thereby enhancing VWM capabilities for these stimuli. Experience might modulate attentional processes, allowing for quicker and more accurate allocation of attentional resources to body-related stimuli. This is particularly relevant given that attention plays a significant role in the functioning of working memory (Awh & Jonides, 2001). In experts, this attentional focus is even more refined, allowing for the filtering of irrelevant information and the prioritization of essential features, thus contributing to enhanced VWM performance for body stimuli (Cowan, 2001).

In summary, the Sensory Recruitment Hypothesis posits that specialized memory stores in the sensory cortex are engaged by control signals originating from the prefrontal cortex (Awh & Jonides, 2001; D'Esposito, 2007; D'Esposito & Postle, 2015; Jonides et al., 2005; Pasternak & Greenlee, 2005; Postle, 2006). This theoretical framework is substantiated by two main bodies of evidence: firstly, neuroimaging studies have demonstrated that the content of working memory can be decoded from activity in the early visual cortex (Christophel et al., 2012; Emrich et al., 2013; Ester et al., 2009; Harrison & Tong, 2009; LaRocque et al., 2016; Riggall & Postle, 2012; Serences et al., 2009). Secondly, there exists a well-documented body of research underscoring the critical role of sustained activity in the prefrontal cortex for working memory functions. The integration of control signals from the prefrontal cortex with sensory representations could offer several advantages. For instance, because the size of receptive fields in later visual regions tends to be more expansive and less precise, top-down signals may be required to recruit early visual areas for tasks that necessitate precise visual representations or for comparing perceptual input with stored memories (Merrikhi et al., 2017). Specific studies on VWM for body stimuli had already disentangled the role of somatosensory areas when it comes to maintain body related information beyond visual activity. Future research could also explore how VWM for body stimuli is modulated by attention, or how it is influenced by expertise.

2.2. Attentional modulation in Working Memory

Several leading theories of attention have traditionally concentrated on the selection and modulation of external cues. However, more recent research interest has begun to broaden this scope to include attention directed toward internal cognitive representations (Chun et al., 2011). This updated classification of attention delineates its function based on the kinds of data it can modulate—referred to as the "targets of attention." These targets may be either external, such as sensory input from the environment, or internal, like stored memories or mental schemas. In this context, "attention" represents an integrated network of neural systems and cognitive functions that reconcile these diverse factors, giving precedence to the most heavily weighted information at any given moment for specialized processing (Narhi-Martinez et al., 2022).

2.2.1. External attention: Balancing top-down vs. bottom-up influences

Information in our external environments constantly competes for our conscious attention. Previous research has distinguished two primary modes of attention that arbitrate this contest: top-down, goal-directed attention and bottom-up, stimulus-driven attention (Beck & Kastner, 2009; Carrasco, 2011). When looking for a green ball, for example, top-down attention allows us to focus on circular shapes and green coloured figures, thereby enhancing search efficiency. This aligns with the biased competition model of attention, proposing that neurons specialized in processing goal-relevant features are activated to prioritize these features over others (Desimone & Duncan, 1995; Reynolds et al., 1999). Human behavioural experiments have investigated differences between these attention systems in temporal processing (Carrasco, 2011; Ling & Carrasco, 2006) and in the extent of automatic versus voluntary control (Folk et al., 1992; Pashler, 1988). Furthermore, distinct neural circuits underlie these systems: the dorsal attention network (DAN) supports goal-oriented attention, while the ventral attention network (VAN) interrupts the DAN to quickly redirect attention toward novel or salient events (Corbetta et al., 2000; Downar et al., 2001). Neurophysiological investigations in primates have revealed neural priority maps in several brain regions (Gottlieb et al., 2009; Li, 2002; Mazer & Gallant, 2003). The activation in these maps is influenced by both goal-relevance and stimulus salience, guiding attention toward the most crucial information.

In the current conceptualization as a multi-level system of weights and balances by Narhi-Martinez, Dube & Golomb (2022), external inputs are assigned weights based on both their goal relevance and their salience. Goal-consistent information is weighted more heavily, favouring it in the attentional competition (Narhi-Martinez et al., 2022). Nevertheless, particularly salient or novel stimuli can disrupt this balance, overriding goal-directed focus (Pashler, 1988; Yantis & Jonides, 1984).

2.2.2. Internal attention and the role of working memory

Our interactions with the external world necessitate focused attention on perceptual data. Yet, attention is also directed internally, operating on representations in the absence of external stimuli (Chun et al., 2011). The efficacy of our actions often hinges on internal perceptual and memory-based representations (Carlisle et al., 2011). Importantly, the scope of internal attention is broader than previously conceived, extending to individual objects and features in memory (Cowan, 2010; Luck & Vogel, 1997; Serences et al., 2009), as well as to non-sensory representations like visuospatial and action plans (van Ede, 2020; van Ede et al., 2019).

Working memory is closely tied to attention (Awh et al., 1998; Baddeley, 1993; Chun, 2011), although the nature of this relationship is a subject of ongoing debate (Oberauer, 2019). Some researchers argue that working memory essentially involves prolonged internal attention to a set of elements (Chun, 2011), while others contend that it is a separate mechanism for active storage (Oberauer, 2019). Regardless of the perspective, working memory acts as an intermediary between perception and long-term memory, relying heavily on internal attention for storing, evaluating, and recalling information.

Internal attention is not static; it dynamically prioritizes information in working memory based on the immediate objectives (Garavan, 1998; Gehring et al., 2003). It operates flexibly over its targets, enabling rapid shifts in focus (van Moorselaar et al., 2015). Conceptually, this can be likened to a system of weights and balances, wherein competing internal signals are assessed and prioritized (Narhi-Martinez et al., 2022). It is worth noting that less is known about the mechanisms of internal attention, partially due to the lack of a spatial or map-like structure for most internal attention targets. This presents an interesting challenge to existing models of attention like salience maps or "pop-out" metaphors.

Attention is not solely determined by immediate circumstances; it is also influenced by historical relevance. Current theoretical frameworks widely acknowledge a third factor, driven by experience, which elucidates how learned patterns in the environment guide attention (Anderson et al., 2021; Awh et al., 2012; Hutchinson & Turk-Browne, 2012; Theeuwes, 2019). Specifically, stimuli can gain significance either by being immediately relevant to ongoing goals (top-down) or by their inherent salience (bottom-up). Additionally, past experiences can also guide the allocation of importance to stimuli, such as objects, locations, or sounds, that have proven consistently beneficial. For example, the weight allocated to a specific location could escalate over time if a desirable item frequently appears there (Geng & Behrmann, 2005), or a particular colour could gain greater prominence if consistently associated with higher rewards (Anderson et al., 2013). Consequently, experiential learning can both amplify and attenuate the weights assigned to stimuli based on their relevance or irrelevance, respectively (Leber et al., 2016; Wang & Theeuwes, 2018). Notably, empirical studies suggest that learned experience can be even more potent than explicit directives in de-prioritizing certain stimuli (Beck et al., 2018; Moher & Egeth, 2012).

Furthermore, attention's framework of weights and balances not only integrates past experiences but also accommodates prospective needs (Narhi-Martinez et al., 2022). For example, internal attention can sustain information that will be pertinent in the future while keeping it in a dormant state that does not interfere with present behaviour (Olivers et al., 2006). When encountering a visual cue that will be crucial for an impending task, this information can be preserved in working memory in an inactive but readily accessible state. The weight assigned to this internal information adjusts dynamically, being stronger when immediately relevant and weaker when pertinent to future tasks, thereby adapting to changing task demands and objectives.

2.2.3. Top-down modulation in visual WM paradigms

Top-down attentional modulations in visual working memory (VWM) have garnered substantial interest over the years, as they provide crucial insights into the selective processing and retention of relevant information. Overall, top-down modulatory processes in VWM facilitate a selective focus on pertinent information and have significant implications for understanding human cognition.

The prefrontal cortex plays a central role in these modulations, assisting in both the selective encoding of new information and the suppression of distractors. Studies employing functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) techniques have repeatedly shown that the prefrontal cortex is involved in the top-down modulation of VWM (Awh & Jonides, 2001; Nobre et al., 2004). Specifically, this region has been implicated in controlling what information gains access to working memory and what gets filtered out. Using visual working memory tasks, it has been observed that attention can be directed toward specific features of stimuli, such as colour or orientation, to enhance the fidelity with which these features are represented (Bays & Husain, 2008; Serences et al., 2009). Other studies using memory cues indicated that participants were able to allocate attention to particular items within working memory to prioritize their maintenance or facilitate their retrieval (Griffin & Nobre, 2003).

Advances in neural decoding techniques have also illuminated the fine-grained feature-specific representations that can be maintained through top-down control. fMRI studies have shown the specific features of a stored memory item based on activity patterns in the visual cortex, highlighting the role of top-down modulation in maintaining feature-specific representations (Harrison & Tong, 2009; Serences et al., 2009). These processes are sensitive to a variety of factors, including cognitive load and temporal dynamics, indicating that attention can be dynamically allocated among items within working memory. Some studies have demonstrated that the efficacy of attentional control diminishes as the number of items to be maintained increases, possibly due to limitations in cognitive resources (Todd & Marois, 2004; Lepsien & Nobre, 2007). Other studies looking at temporal dynamics had shown that attention can be dynamically shifted among items in working memory, suggesting a temporal dimension to top-down control (Kiyonaga & Egner, 2013).

A critical feature of top-down modulation in VWM is the ability to suppress potentially distracting information. It has been shown that this suppressive mechanism is particularly important for maintaining high fidelity of stored information (Gazzaley & Nobre, 2012). Attentional mechanisms are considered to play a role in reducing both proactive and retroactive interference in VWM, thus assisting in the maintenance of relevant information (Oberauer et al., 2015). Finally, attentional modulation is known to vary across individuals and to decline with age. There are documented age-related declines and individual differences in the ability to exert top-down control in VWM (Gazzaley et al., 2005; McNab & Klingberg, 2008).

In summary, top-down attentional modulations in visual working memory involve neural networks with the prefrontal cortex as a major hub, and their efficacy is influenced by a variety of factors such as cognitive load, time, and individual differences. These modulatory processes facilitate both the selective encoding and robust maintenance of information in VWM. Furthermore, research has pointed to the significance of top-down attentional control in mitigating interference and enhancing the fidelity of stored information.

Chapter 3: Working memory for bodies and top-down modulation: an EEG study.

3.1. Introduction

Working memory serves as a temporary storage system for manipulating information essential for complex cognitive tasks. Oberauer and colleagues have contributed to the field by distinguishing between declarative and procedural components within working memory (Oberauer, 2009). They suggested that working memory comprises multiple 'slots' for holding information, which can be either declarative or procedural. Attentional control plays a key role in determining what information is stored, and it operates in a domain-specific fashion (Oberauer & Hein, 2012). This line of research underscores that procedural memory, concerned with 'how-to' knowledge like skills and habits, is encoded and retrieved differently than declarative memory (Oberauer, 2019).

Research in Visual Working Memory (VWM) has primarily focused on geometric shapes, colours, and faces as subjects of investigation. However, VWM for body stimuli offers a distinct avenue for exploration, with specialized neural substrates such as the extrastriate body area (EBA) becoming areas of interest (Downing et al., 2001). Additionally, the superior parietal lobule and the dorsal premotor cortex, regions usually associated with action planning and observation, have been implicated in VWM tasks centred on body parts (Astafiev et al., 2004; Ferri et al., 2015). Recent EEG methodologies have been developed to dissociate somatosensory components from visual effects, revealing how both are modulated by cognitive load in VWM tasks when comparing body parts and matched non-body-related stimuli (Galvez-Pol et al., 2018a).

Top-down attentional modulations in visual working memory (VWM) have emerged as a critical research area, providing vital insights into the selective retention and processing of pertinent information. The prefrontal cortex is central to these modulations, aiding in the selective encoding of new information and the suppression of irrelevant distractors (Awh & Jonides, 2001; Nobre et al., 2004). Techniques such as functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) have confirmed the role of this brain region in determining what information is admitted into working memory and what is excluded. Furthermore, these modulatory processes allow for attention to be focused on specific features of stimuli, like colour or orientation, thus enhancing the quality of their mental representation (Bays & Husain, 2008; Serences et al., 2009).

In addition to the role of the prefrontal cortex, advances in neural decoding have elucidated feature-specific representations maintained through top-down control (Harrison & Tong, 2009; Serences et al., 2009). This attentional focus is not static; it can be dynamically allocated among multiple items in VWM, influenced by factors like cognitive load and temporal dynamics (Todd & Marois, 2004; Lepsien & Nobre, 2007; Kiyonaga & Egner, 2013). A pivotal aspect of this top-down modulation is its ability to suppress distracting information, thereby enhancing the fidelity of stored data (Gazzaley & Nobre, 2012). Such suppressive mechanisms also help in mitigating both proactive and retroactive interference, thereby aiding in the maintenance of relevant information in VWM (Oberauer et al., 2015). Based on the previous literature and following up recent work by Galvez-Pol and colleagues (2018a; b) showing neural recruitment of body-related cortices during active maintenance of body-related images in WM, this study investigated whether the active maintenance of body-related images can be modulated by top-down attentional mechanisms. To this aim, persistent activity was measured (a neural marker of attention-based rehearsal) during a WM task for body-related images. Persistent activity, also known as sustained or delayed activity, is usually observed between the memory and test phases (i.e., during the retention interval) in sensory areas relevant for the task. A key feature of this activity is that it persists in the absence of continued sensory input. Moreover, its amplitude is modulated by memory load (Luria et al., 2016; Vogel and Machizawa, 2004, Galvez-Pol et al., 2018a).

Using a change detection paradigm, a series of experimental manipulations were designed. During the memory array (100ms), attention to either left or right visual hemi-field (VHF) was cued by an arrow (fixating vision in the centre of the screen), allowing to look at later contralateral or ipsilateral activity. Stimuli displayed were coloured hands portraying different postures. Task conditions were blocked and instructed to attend and maintain information regarding to either the posture or the colour, allowing later comparison based on the attended feature. The number of items displayed at both sides of the screen could be 1 or 2, allowing to look at modulation by memory load. Half of the trials would contain a bilateral tactile stimulation in the index fingers simultaneously to the visual onset, unrelated to the task and used for a later methodological subtraction allowing to isolate somatosensory activity (SEPs VEP-free, see Galvez-Pol et al., 2020). The retention interval lasted for 900ms, followed by the test array until verbal response. The test array could be same or different from the memory array regarding the attended feature and participants responses were recorded accordingly. The non-attended feature could follow the same trend as the attended feature, resulting in congruent trials where both features change together, or the opposite trend leading to incongruent trials where only one feature changed and the other remained constant. This last manipulation is key to allow us to explore how bottom-up or top-down attentional modulations might be interfering along the main task.

In line with previous research, we predicted somatosensory activity to show a modulation by memory load in the contralateral hemisphere during the retention interval *(hypothesis 1)*. Further, we wanted to explore whether the attention could modulate cortical engagement, hypothesising that colour features would rely more on visual cortices and posture on somatosensory areas for encoding and maintenance of the information to-be-remembered *(hypothesis 2)*. Lastly, we could explore saliency effects (bottom-up attention) interfering with the task instructions (top-down attention), by looking at differences between congruent and incongruent trials *(hypothesis 3)*.

3.2. Methods

3.2.1. Participants

Thirty paid adult participants took part in the experiment. Two participants were excluded from analysis due to technical problems during EEG recording, another five participants were excluded due to poor behavioural performance (overall accuracy lower than 60%). The remaining twenty-three participants (mean age 32.26, SD 8.75, 7 females) had normal or corrected to normal vision, and reported no cognitive, attentional, or neurological impairments. Written and informed consent was obtained from all participants and the study was approved by the Psychology Research Ethics Committee at City, University of London [PSYETH (UPTD) 13/14 18], following guidelines and procedures established in the Declaration of Helsinki. The sample size of the current experiment was based on previous studies using comparable WM paradigms and techniques (Vogel and Machizawa, 2004; Galvez-Pol et al., 2018a, b; Galvez-Pol et al., 2020).

3.2.2. Stimuli

A set of 36 images depicting 6 different meaningless postures of right hands in 6 different colours were used (see Fig. 1a). Mirrored images were used to create the equivalent set of 36 left hands, resulting in a total of 36 pairs of right and left coloured hands. The hand images were adapted from Galvez-Pol et al. (2018a), applying 25% and 50% shades of red, blue or green tones respectively to the original grey-scaled stimuli (using standard RGB filter from Adobe Photoshop).

3.2.3. Experimental design and procedure

The experiment consisted in a repeated measures design with two factors: task and memory load, each with two categorical levels. The first factor was defined by the attended stimuli feature, stating colour or posture tasks as levels, manipulated as separate experimental blocks, and counterbalanced between participants controlling for order effects. The second factor was memory load, operationalized as one or two items to be remembered, and manipulated in a randomised order within task blocks. A change detection paradigm was used as the visual memory task. Participants were shown a first array of items (memory array), followed by a blank retention interval, and a second array (test array) which could display the same or different items compared to the first array. Participants had to maintain in working memory the attended features, either posture or colour, from one or two items, appearing on the left or right side of the screen while fixating in the centre. The stimuli presented in the test array could differ from the memory array in both colour and/or posture, but Participants' responses were measured using a yes/no forced choice to the experimental question on whether the items in the test array portrayed the same attended features (colour or posture) as the memory array.

Participants performed a visual working memory task adapted from Galvez-Pol et al. (2018a) in which items to-be-remembered were coloured hand images (depicting 6 different hand postures and in 6 different colours, see Fig.1a). Participants were cued at the beginning of each trial by a central arrow to attend to items displayed in their left or right VHF. This was followed by a bilateral memory array depicting 1 or 2 items in each hemifield and a blank retention interval lasting 900 milliseconds. Right hand images were shown on the right hemifield while left hand images were displayed on the left. In 50% of the trials, a single tactile tap was applied bilaterally to the index fingers simultaneously with the memory array. A final test array that differed in 50% of the cases from the memory array by one item was displayed until participants verbally responded whether (yes/no) the memory and test arrays were identical regarding the task-relevant feature (colour or posture). Due to the same stimuli been used for both tasks and the only difference being on the attended feature (by instructions, per block and counterbalanced order across participants), half of the trials were congruent meaning both stimuli features (colour and posture) changed or not on the same trend, and the other half were incongruent trials as the feature-to-be-attended and the unattended followed different trends (see Fig.1b). Participants' forearms rested on the top of a table with their hands separated in palm up position while covered by a black surface (see Fig. 2a). Visual stimuli were displayed using E-Prime Software (Psychology Software Tools).

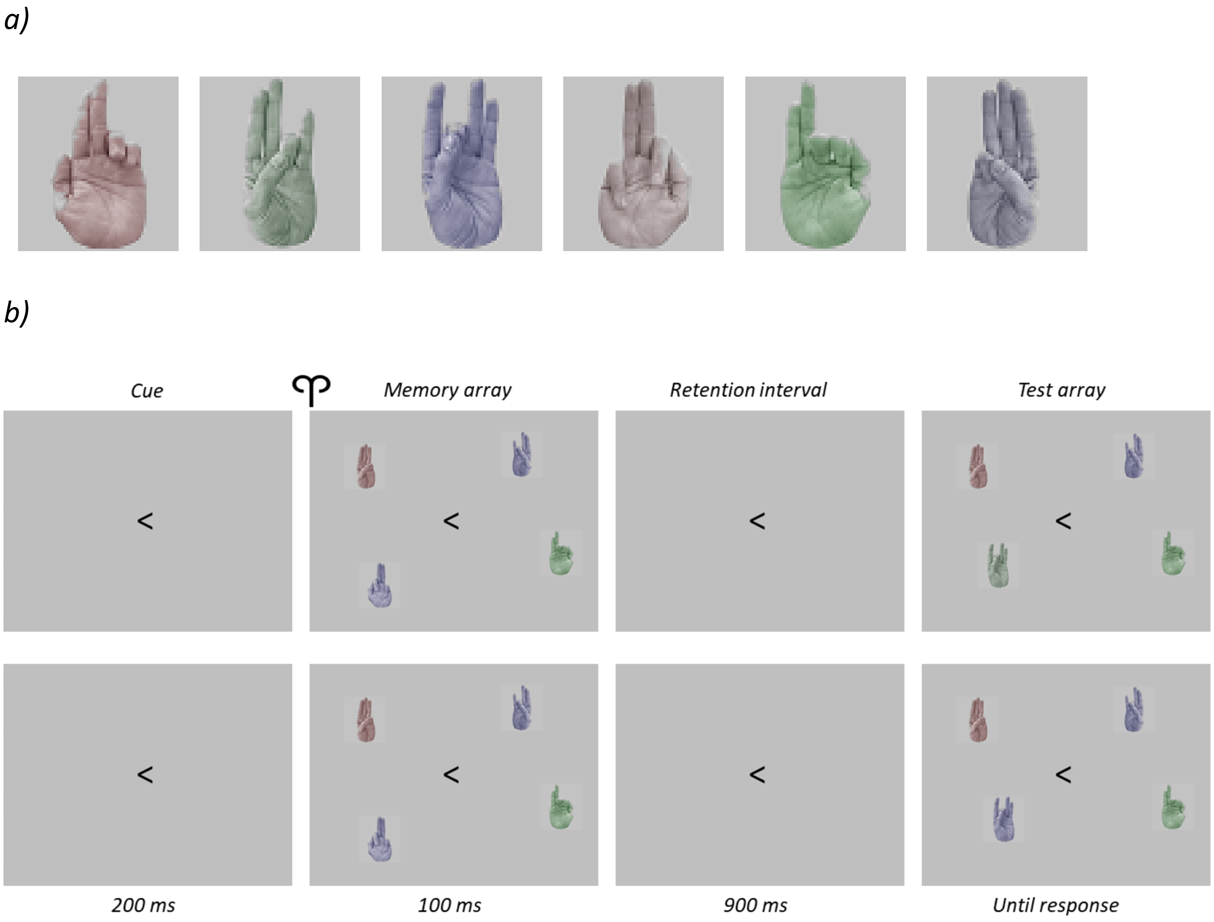


Figure 1. Stimuli and trial design. (a) Left and right hands portraying 6 different postures on 6 different colours (red, green, and blue at 25 or 50% RGB filter). Adapted from greyscale images (Galvez-Pol et al., 2018), using Adobe Photoshop. (b) Change detection paradigm: trial sequence. Participants were instructed to fixate on the centre of the screen, attend to the side pointed by the arrow cue and respond whether the second array was the same as the first regarding the instructed feature: either colour or posture of the hands. A memory array was visible for only 100ms, and in half of the trials a bilateral tactile stimulation in the index fingers would happen simultaneously to the visual onset. The retention interval lasted for 900ms, followed by the test array until verbal response. The arrays could display one or two items on each side of the central arrow, and these could change between the memory and test arrays. Depending on the block, participants were asked to attend to the posture or colour of the hands. Due to using the same visual stimuli but focusing on 2 different features, we decided to control for congruent (top sequence) and incongruent (bottom sequence) trials. The changing item (left-bottom hand on the screen) would be congruent when both features (colour and posture) followed the same trend changing (or not) together, facilitating the perception of same/different in the array. Incongruent trials would only change one of the features of the stimuli (colour or posture) while the other remain constant.

All stimulus arrays were presented within two 4.9° x 9.2° rectangular regions that were centred 6.4° to the left and right of a central fixation cross on a light grey background. The positions of all stimuli were randomized on each trial with the constraint that the distance between stimuli within a VHF was at least 2.4° (centre to centre). Each memory array consisted of 1 or 2 hands (1.9° x 1.7°) in each VHF. Screen resolution was 1280 x 1024 pixels, screen size was 41 x 30 cm, and distance screen to eye was 85cm.

Visual-only trials. In 50% of the trials only VEPs were elicited. These were recorded from the onset of the visual memory array and while participants maintained in working memory the stimuli until the test array appeared. Visual-tactile trials. In the other 50% of the trials, we elicited simultaneously VEPs and SEPs by applying task irrelevant single tactile taps simultaneously delivered to both hands on the tip of the participants' index fingers at the onset of the visual memory array. Tactile stimulation was applied using two 12 V solenoids driving a metal rod with a blunt conical tip that contacted with participants' skin when a current passed through the solenoids. Both solenoids were placed on the tip of the index fingers, one for each hand. To mask sounds made by the tactile stimulators, white noise (65 dB, measured from participants' head) was presented through a loudspeaker centrally positioned 90cm in front of the participants. Participants were instructed to ignore these tactile stimulations.

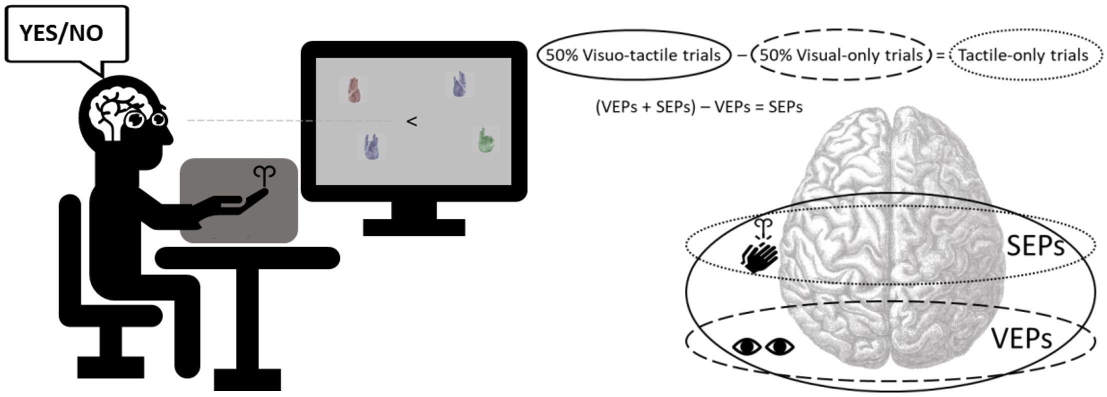


Figure 2 (a) Experimental procedure. Participants had to fixate at the centre of the screen, while attending to the cued side. They were asked to attend the cued side of a visual array displaying 1 or 2 items per side and respond verbally yes or no to whether the test array was the same as the memory array regarding the attended feature (colour or posture). Participants' electrophysiological data was recorded during the experiment, in an electromagnetic-shielded room, where a black screen was covering their hands from sight. In 50% of the trials, a bilateral mechanical tap was applied to the index fingers at the onset of the visual memory array. (b) ERPs subtraction methodology. Tactile stimulation enhancing somatosensory activity was applied in half of the trials, creating visuo-tactile conditions (VEP+SEP). To isolate somatosensory activity, a later subtraction was computed where visual-only trials were subtracted from visuo-tactile trials, leaving tactile-only trials. In these tactile-only trials we can investigate SEPs beyond visual carry-over effects.

Overall, participants performed a total of 1344 trials, 672 per task condition (colour or posture). This is equal to 336 trials for each memory load condition (one or two items). Half of the trials randomly presented task-irrelevant tactile taps (visual-tactile trials) while the other half involved visual stimulation only (visual-only trials). The task condition (i.e., attending to colour or posture) was blocked, instructed by experimenter, given practice trials, and the order was counterbalanced between participants. A change detection paradigm was used, where the memory array and the test array differed on the attended feature 50% of the trials. The unattended feature could change together with the attended one (congruent trials), or it could not follow the change of the feature to-be-attended (incongruent trials).

3.2.4. EEG recording and data analysis.

EEG recording took place in an electromagnetic-shielded room (Faraday cage), using a 75 Hz LCD monitor. Continuous EEG data was recorded using a 64-electrodes equidistant M10 montage (EasyCap) and a BrainAmp amplifier at 500 Hz sampling rate. Electrodes were on-line referenced to the right earlobe and off-line re-referenced to the average reference (excluding ocular electrodes: HEOGL, HEOGR, VEOG). Vertical and bilateral horizontal electrooculogram was recorded to track eye movements and further artifact rejection. Continuous EEG data was submitted to a 30 Hz low-pass filter, and the Ocular Correction transformation algorithm by Gratton & Coles (Gratton, Coles & Donchin, 1983) was implemented in Brain Vision Analyzer 2 to the raw continuous data. This method used EOG channels to estimate the influence of blinks and eye movements on the EEG channels via regression. Ocular artifacts are then subtracted based on a correction factor. Epochs were extracted as 1300 ms intervals and corrected relative to a 200 ms pre-stimulus baseline. Artifact rejection was computed eliminating epochs with amplitudes exceeding +/- 100 µV, and visual inspection to reject remaining blinks and overt lateral eye movements. BrainVision Analyzer 2.2 software was used to analyse EEG data and export values for later statistical analysis.

Contralateral waveforms were calculated as the average from electrode sites opposite from the attended visual hemifield (VHF), namely right hemisphere electrodes when attending left side of the screen and left hemisphere electrodes when attending right side of the screen. Ipsilateral waveforms were computed as the average from electrode sites in the same side as the attended VHF (left hemisphere when attending left VHF, right hemisphere when attending right VHF).

The mechanical stimulation of the tactile taps allowed us to examine the state of the SCx, exposing its underlying processing during memory encoding and maintenance of the visual stimuli by measuring the electrocortical activity (SEPs) elicited by task irrelevant tactile stimulation. These tactile taps probe SCx responses during encoding and maintenance of the visual stimuli. Moreover, to be able to isolate somatosensory processing over corresponding parietal electrode sites from possible carry over visual effects elicited by the visual onset of the stimuli, we subtracted brain activity from those trials containing activity only due to VEPs (visual-only trials) to those trials containing a combination of visual and somatosensory activity due to the combined VEPs-SEPs (visual-tactile trials). This allowed isolating and observing somatosensory processing free of visual evoked activity (SEPs, VEPs-free) (see Fig. 2b, also: Sel et al. 2014, Galvez-Pol et al., 2018a, b; Galvez-Pol et al., 2020).

The choice of electrodes was based in previous work using change detection paradigm looking at visual and tactile WM (Vogel & Machizawa, 2004; Katus & Eimer, 2016; Galvez-Pol et al., 2018). Visual activity was measured at occipital and posterior parietal electrodes (O1/2 and PO7/8 of the 10-20 system) and somatosensory activity was measured at parietal electrodes (CP3/4 and CP5/6 of the 10-20 system). The suitability of this regions was investigated by visual inspection of group average waveforms and corresponding neural maps. Grand averages for visual-only and tactile-only trials were computed separately and visually inspected to confirm the appropriate choice of electrode sites. P2 component was used as a marker for visual spatial attention for visual-only trials, and P50 component was used as a short latency cortical response to somatosensory stimulation for tactile-only trials (Fig. 3). Both presented the expected components with a higher contralateral engagement, as expected from a lateralized attention task. This visualization allowed us to confirm the choice of electrodes from previous studies using similar paradigm and SEP subtractions (Galvez-Pol et al., 2018, 2020).

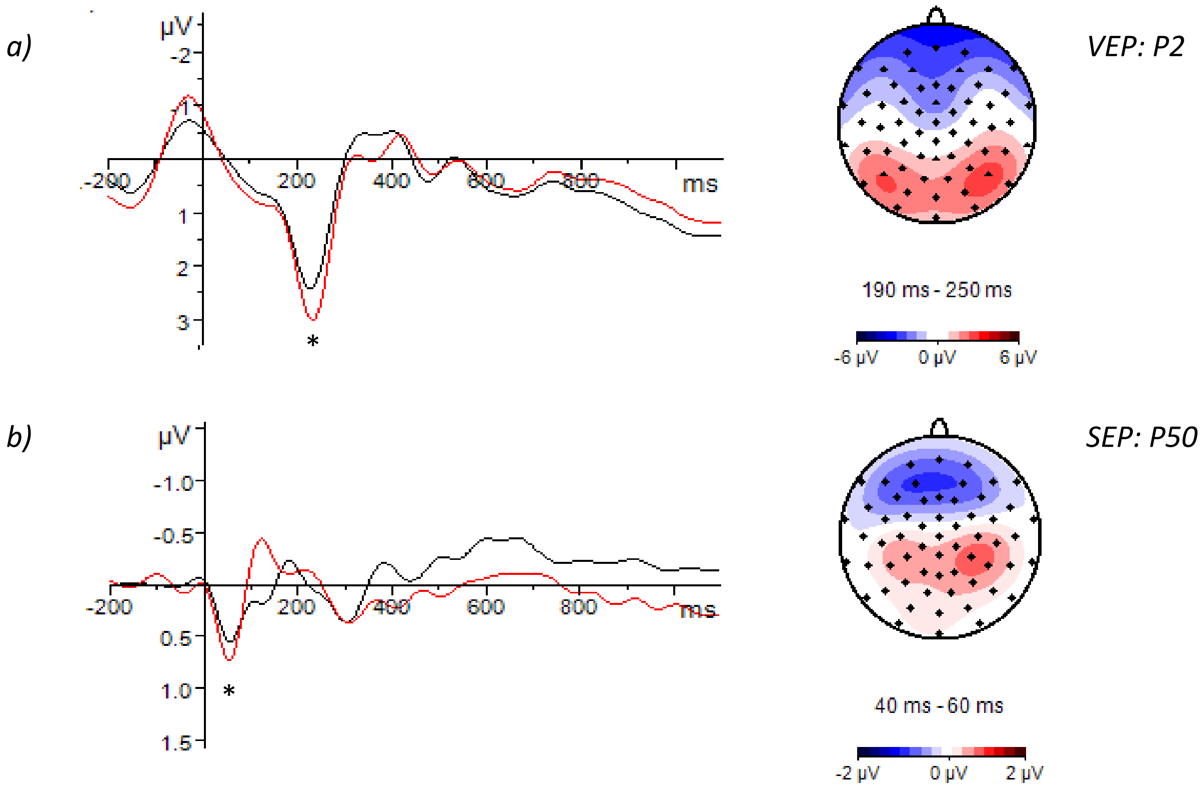


Figure 3. Average ERP waveforms across all experimental conditions. Contralateral waves plotted in red and ipsilateral in black. Negative amplitudes are plotted upright (μV) during the 1300ms time window (including 200ms baseline before onset). a) Visualization of visual-only trials (VEPs) showing P2 component for visual spatial attention, together with the map of neural activity at the peak of the component (190-250 ms), showing higher contralateral activity on visual electrodes. b) Visualization of tactile-only trials (SEPs) showing P50 component for body-related attention over somatosensory electrodes with higher contralateral engagement, together with neural activity map at the peak of the component (40-60 ms). Note that in this SEP average across both tasks, a lateralized difference can already be observed in early components and as a persistent slow wave from 350ms until the end of the trial, suggesting somatosensory engagement in maintenance of body-related stimuli.

Grand averages were computed separately for each experimental condition (colour or posture task, 1 or 2 items to be remembered, congruent or incongruent feature change, visual-only or visuo-tactile trials), and independently for left and right visual hemifields (participants were cued to attend the left or right side of the memory array while fixating in the centre). Two different time windows were exported for statistical analysis. The first one, to look at early attentional processes and potential top-down modulations encoding the feature-to-be-attended, was exported as the mean amplitude from 200 to 300 ms after stimuli onset. The second time interval, to further explore maintenance in working memory as a slow wave modulated by memory load, was exported as the mean amplitude from 300 to 900 milliseconds. Mean amplitudes for both time windows were extracted from visual (in visual-only trials) and somatosensory areas (in tactile-only trials), separately for each experimental condition.

Contralateral Delayed Activity (CDA), measured as the difference waveform subtracting ipsilateral from contralateral activity was used to look at lateralized effects. CDA waveforms had been used as a measure of visual working memory maintenance using the same change detection paradigm in previous experimental work (Vogel & Machizawa, Katus & Eimer, Galvez-Pol et al., 2018). As the memory array displays stimuli on both sides of the screen, while the cue is directing attention only to one side of the screen, both hemispheres are exposed to the visual stimuli but only the contralateral hemisphere to the attended side will be carrying most of the encoding and maintenance of that information in working memory.

Current Source Density Analysis (CSD)

Current Source Density is an independent method to estimate the neuronal generator patterns underlying EEG signal, allowing further exploration of topography from ERP data (Tenke & Kayser, 2013; Katus et al., 2015; Galvez-Pol et al., 2018). Using Brain Vision Analyzer 2, CSD analysis was performed on the EEG data under standard parameters (Lambda: 10^{-5}, order of splines m: 4, Legendre polynomial: 10) to estimate the current source generating the measured potentials. CSD maps were also performed on the early components used for confirmatory visual inspection of electrodes (*see figure 4*). Later, a secondary statistical analysis was ran using CDS data on VEPs and SEPs looking for differences between the different conditions.

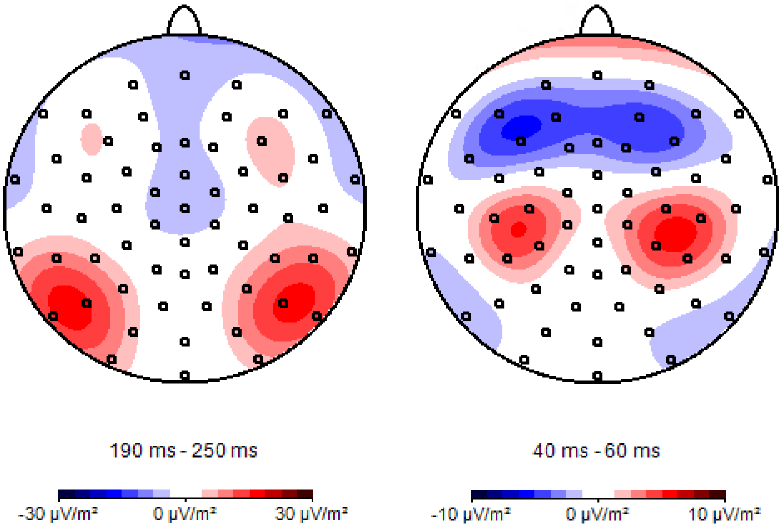


Figure 4. CSD ($\mu V/m^2$) average maps for visual trials (left) showing VEP P2 visual spatial attention component, and tactile trials (right), showing SEP P50 tactile stimulation component. These maps were used to confirm the topography and choice of visual & somatosensory electrodes for the experimental analysis.

To examine the signal to noise ratio, statistical analysis was performed in the number of correct trials used for ERP averages. More specifically, to ensure the suitability of the methodological subtraction of visuo-tactile trials minus visual-only trials, it was key that the number of segments averaged per condition were not significantly different.

All statistical analysis were performed with IBM SPSS Statistics software using Repeated Measures (RM) Analysis of Variance (ANOVA) and paired-samples t-test. When appropriate, degrees of freedom and p values were adjusted using Greenhouse-Geisser and Bonferroni corrections respectively.

3.3. Results

3.3.1. Behavioural performance

Accuracy (ACC) and sensitivity (d') were used for statistical analysis of behavioural data. A 2x2x2x2 Repeated Measures Analysis of Variance (RM ANOVA) was performed for each measure, with factors being task (colour or posture), load (1 or 2 items to be remembered), congruency (congruent or incongruent change of features) and type of trial (visual-only or visuo-tactile trial). See Table 1 for descriptive statistics and Figure 5 for graphical representations per experimental condition for both behavioural measures.

Overall accuracy levels detecting a change in the array were 74.6%. Results showed main effects of congruency F(1,22)=34.096, p<.001, PES=.608, load F(1,22)=216.157, p<.001, PES=.908, and a significant interaction between task and load F(1,22)=52.061, p<.001, PES=.703. Participants were significantly more accurate in congruent trials (M=.762, SE=.013) compared to incongruent trials (M=.730, SE=.014), meaning that when both features of the stimuli changed together this facilitated the detection. Participants were also better at detecting change when there was only one item to remember (M=.785, SE=.014) compared to 2 items (M=.707, SE=.013). Bonferroni corrected pairwise comparisons showed the interaction between task and load was explained by a significant difference in performance between colour and posture tasks when there were two items to remember in the array (p<.001) but no task differences when there was only one item (p=.575).

Table 1: Descriptive statistics representing mean (M) and standard deviation (SD) per experimental condition for accuracy (ACC) and sensitivity (d') measures. Note that accuracy measures are given as a proportion (0-1).

				ACC		**d'**	
Task	**Trial modality**	**Feature-change**	**Load**	M	SD	M	SD
Colour	Visuo-tactile	Congruent	1	0.80	0.08	1.53	0.54
			2	0.75	0.09	1.21	0.56
		Incongruent	1	0.76	0.08	1.42	0.52
			2	0.74	0.09	1.20	0.53
	Visual	Congruent	1	0.80	0.08	1.58	0.61
			2	0.76	0.07	1.28	0.40
		Incongruent	1	0.76	0.08	1.34	0.48
			2	0.74	0.08	1.24	0.56
Posture	Visuo-tactile	Congruent	1	0.81	0.09	1.68	0.70
			2	0.70	0.10	0.92	0.62
		Incongruent	1	0.78	0.11	1.47	0.58
			2	0.65	0.09	0.72	0.44
	Visual	Congruent	1	0.80	0.10	1.59	0.67
			2	0.69	0.08	0.85	0.43
		Incongruent	1	0.77	0.11	1.38	0.67
			2	0.63	0.08	0.62	0.45

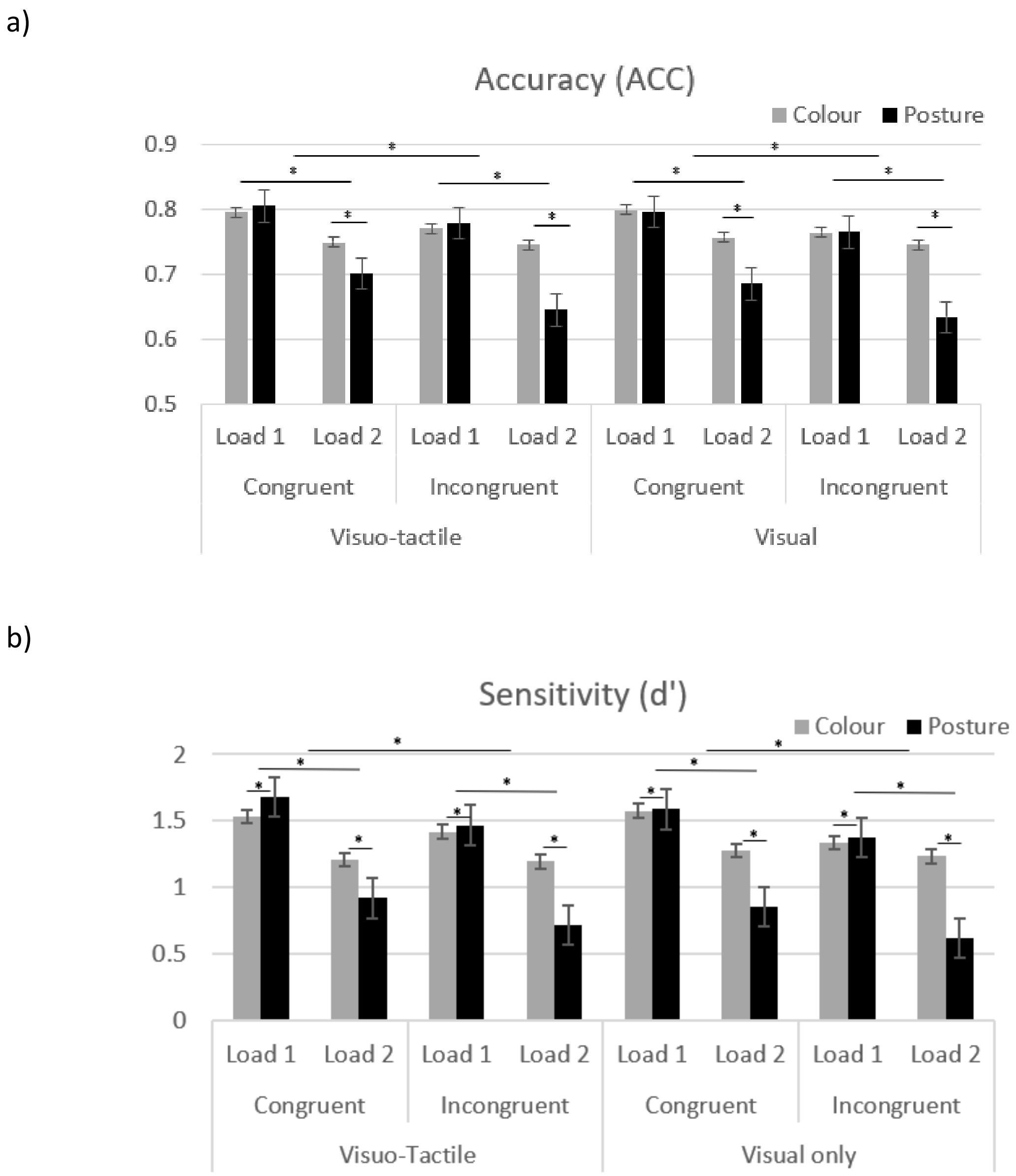


Figure 5. Behavioural plots representing mean values per experimental condition with error bars indicating Standard Errors (SE). a) Accuracy levels as group means for the proportion of correct answers per condition. b) Sensitivity index (d´) displaying group means representing the average discriminability per condition.

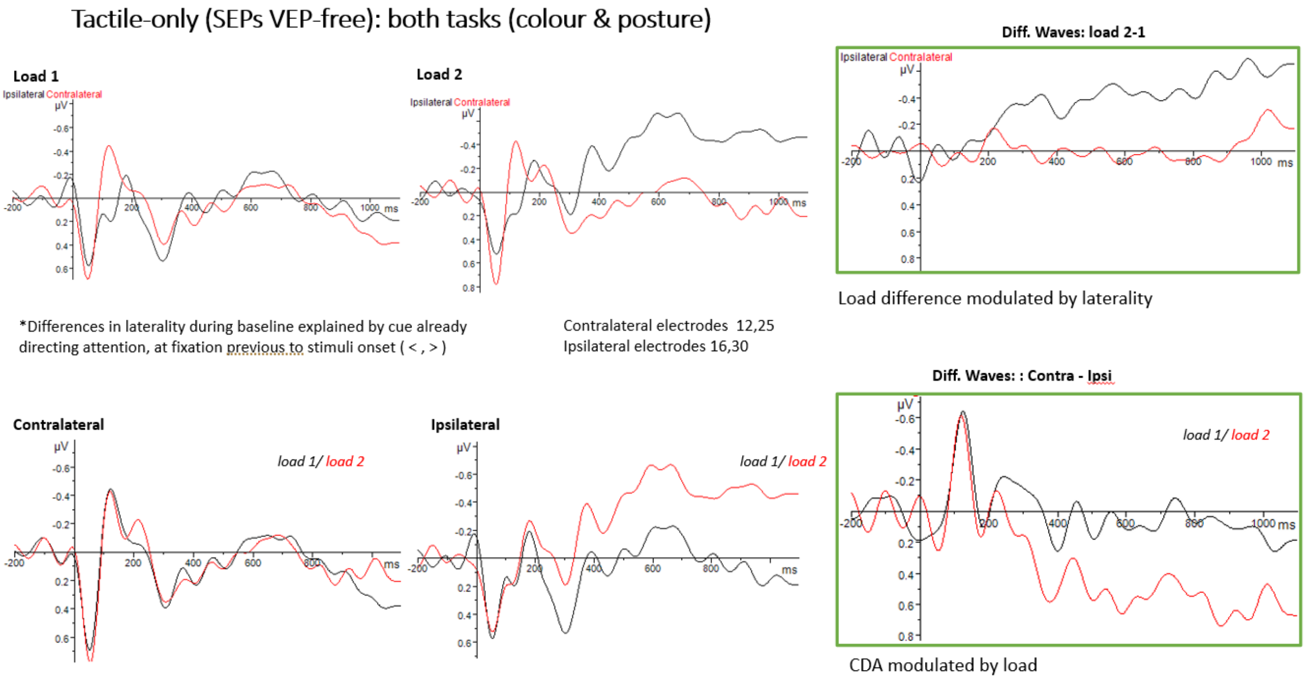
D prime (d’) was calculated as a sensitivity index per participant for each condition. It was computed as the difference between the standard scores for the false-alarm rate and the hit rate and used as a measure discriminability *(see Section 1.4 Measurements & Methods)*.

Results showed main effects of task F(1,22)=4.85, p=.038, PES=.181, load F(1,22)=153.11, p<.001, PES=.874, congruency F(1,22)=21.102, p<.001, PES=.490, and a significant interaction between task and load F(1,22)=23.806, p<.001, PES=.520. Participants were significantly more sensitive to change when attending to the colour feature (M=1.35, SE=.082) compared to attending the posture feature (M=1.15, SE=.101). Change detection was significantly better when attending one item (M=1.50, SE=.089) compared to attending 2 items (M=1.01, SE=.076). Congruency of change between attended features significantly increased discrimination of change (M=1.33, SE=.086) compared to incongruent changes where only one feature changed in the array (M=1.17, SE=.079). Bonferroni corrected pairwise comparisons showed the interaction between task and load was explained by a significant difference in performance between colour and posture tasks when there were two items to remember in the array (p<.001) but no task differences when there was only one item (p=.600).

3.3.2. Somatosensory activity (SEP, VEP-free)

To ensure the suitability of the ERP subtraction method, the number of segments used for ERP averages were analysed comparing visual-only and visuo-tactile trials per condition. Paired samples t-tests were performed for each condition, revealing no significant differences (all p>.05) between the number of visual trials and visuo-tactile trials. These results validated the use of the ERP subtraction (visuo-tactile trials – visual-only trials = tactile-only) demonstrating a good fit on the signal to noise ratio for the subsequent SEP analysis. In this section SEPs were analysed as somatosensory activity free from any visual carry over effect. Average waveforms were computed separately for each experimental condition, looking at posture or colour task, memory load 1 or 2 items, contralateral or ipsilateral activity, and congruent or incongruent change between stimuli features. Values from Mean amplitudes (µM) and Current Source Density (µV/m2) were used for parallel analysis at each time interval.

Visually inspecting the SEP average waves collapsing both tasks, a glimpse of the expected memory load and laterality modulations can be quickly identified. This confirms WM for body related information is engaging somatosensory areas, therefore both paradigm and stimuli seemed to work as expected *(see figure 6).*


Tactile-only (SEPs VEP-free): both tasks (colour & posture)

*Differences in laterality during baseline explained by cue already directing attention, at fixation previous to stimuli onset (< , >)

Contralateral electrodes 12,25
Ipsilateral electrodes 16,30

Load difference modulated by laterality

CDA modulated by load

Figure 6. Overview of SEPs (VEP-free, for both tasks averaged) showed expected modulations by load and laterality, confirming sound methodology when using body-related stimuli.

3.3.2.1. Attentional component (200-300ms)

Somatosensory activity (SEPs, VEP-free) was extracted after preprocessing, as mean amplitudes from the time window corresponding to 200 to 300 ms. RM 2x2x2x2 ANOVA was used for analysis, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended VHF), Load (1 or 2 items) and Congruency (congruent or incongruent change between features). A significant 3-way interaction was found between task, laterality and load F(1,22)=5.303, p=.031, PES=.194 *(see Figure 7)*. To further explore this interaction, two RM 2x2x2 ANOVA were computed separately for colour and posture tasks, with remaining factors being laterality, load and congruency. Results on the colour task revealed a main effect of congruency F(1,22)=5.013, p=.036, PES=.186, meaning there was a significant difference between levels of somatosensory engagement depending on whether the posture feature was changing congruently with the colour or not. A significant interaction was also found between laterality and load F(1,22)=8.249, p=.009, PES=.273. Follow up analysis looked into congruent and incongruent conditions separately using RM 2x2 ANOVA with remaining factors load and laterality in the colour task, revealing a load effect for congruent trials F(1,22)=4.814, p=.039, PES=.311, and a significant interaction between laterality and load for incongruent trials F(1,22)=9.939, p=.005, PES=.311 explained by load differences in contralateral sites (p=.018) but not ipsilateral (p>.05). Results on the posture task showed no significant effects.

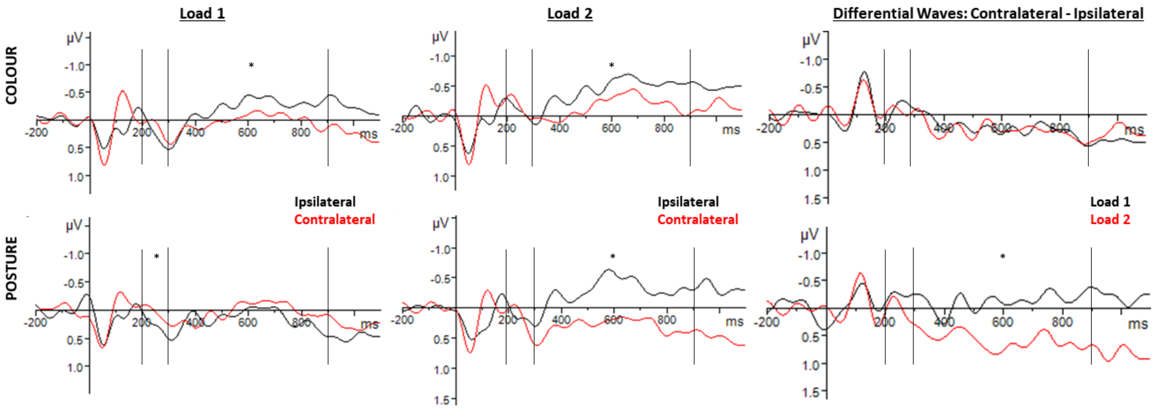


Figure 7. Somatosensory average waveforms representing the 3-way interaction between task, laterality, and load. SEP plots represent separately colour (top) and posture (bottom) tasks, when remembering 1 item (left) or 2 items (centre), showing ipsilateral (black) and contralateral (red) activity. On the right side of the panel, differential waves subtracting ipsilateral from contralateral (CDA) activity were plotted for load 1 (black) and load 2 (red) separately. This hemispherical difference represents the CDA waves, and increases only in the posture task, modulated by memory load. Significant differences are represented (*) for the attentional window (200-300ms) and WM window (300-900ms).

Secondary analysis of SEPs over attentional window (200-300ms) was done using CSD transformations (μV/m2). As before, RM 2x2x2x2 ANOVA was used for analysis, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended VHF), Load (1 or 2 items) and Congruency (congruent or incongruent change between features). Results showed a main effect of congruency F(1,22)=5.628, p=.027, PES=.204, meaning there were significantly different levels of somatosensory activity depending on whether both stimuli features were changing congruently or incongruently within them, and a significant 3-way interaction between task, laterality and congruency F(1,22)=6.39, p=.019, PES=.225 was also found (*see figure 8*). Follow up analysis investigated congruent and incongruent conditions separately using RM 2x2x2 ANOVA with remaining factors task, load and laterality *(see also figure 9).*

Congruent trials showed a significant interaction between task and laterality F(1,22)=4.385, p=.048, PES=.166, where there was a bigger difference between the tasks in the contralateral hemisphere compared to the ipsilateral, although none of the differences were significant after Bonferroni correction for multiple comparisons (all p>.05). Analysis of incongruent trials revealed a main effect of task F(1,22)=7.494, p=.012, PES=.254, showing significantly different levels of somatosensory activity depending on the task, and a significant interaction between task and load F(1,22)=9.669, p=.005, PES=.305. This interaction was explained by a significant difference between colour and posture tasks when there were 2 items to remember (p<.001) but no difference between the tasks when only one item appeared in the memory array. After finding a main effect of task for incongruent trials, a follow up analysis was run with factors load and laterality, performing a RM 2x2 ANOVA for each task separately. A main effect of laterality F(1,22)=5.326, p=.031, PES=.195 and load F(1,22)=5.243, p=.032, PES=.192 were found for the colour task, but no significant results were seen for the posture task.

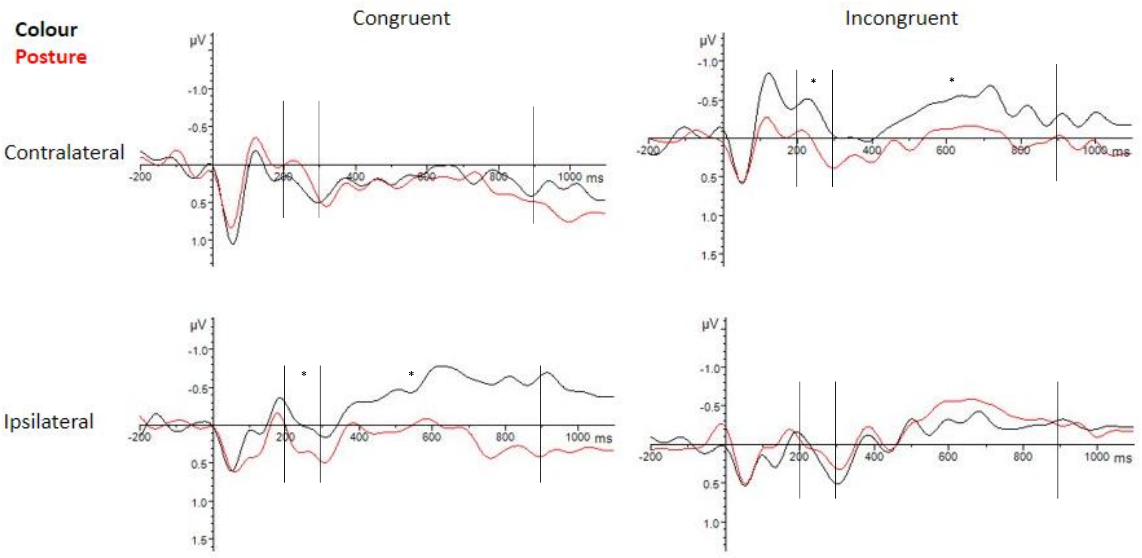


Figure 8. Somatosensory average waveforms representing the 3-way interaction between task, laterality, and congruency. Somatosensory Evoked Potentials (SEPs) comparing tasks as Colour (black) and Posture (red), representing Congruent (left side), and Incongruent (right) conditions. Top: Contralateral activity, showing task differences for the incongruent trials only, but not for congruent trials. Bottom: Ipsilateral activity, showing the reversed pattern with sustained differences between tasks on the congruent trials only. Significant differences are represented (*) for the attentional window (200-300ms) and WM window (300-900ms).

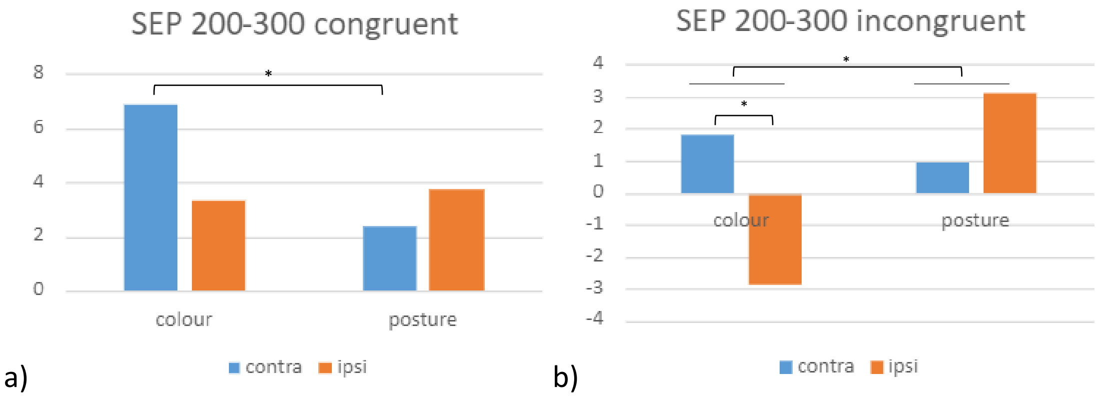


Figure 9. SEPs CSD mean values (µV/m2) over attentional window (200-300), showing main effect of congruency and 3-way interaction between congruency, laterality, and task. Contralateral activity

plotted in blue, Ipsilateral in orange, both graphs comparing colour task on the left side and posture task over the right side. a) Congruent trials showing a contralateral difference between tasks. b) Incongruent trials showing a main effect of task and stronger laterality differences on the colour task. Significant differences are represented (*).

3.3.2.2. Contralateral Delayed Activity (CDA) (300-900ms)

Somatosensory activity (SEPs, VEP-free) was extracted as mean amplitudes from the time window corresponding to 300 to 900 ms. RM 2x2x2x2 ANOVA was used for analysis, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended side of the screen), Load (1 or 2 items) and Congruency (congruent or incongruent change between stimuli features). Results showed a main effect of load F(1,22)=7.243, p=.013, PES=.248, where mean amplitudes were more negative when remembering 2 items (M= -.188, SE=.092) compared to 1 item (M=.092, SE=.112). When posture and colour tasks were analysed separately (following a priori planned analysis based on the hypothesis) using RM 2x2x2 ANOVA with remaining factors load, laterality and congruency, results showed memory load was only a significant factor in the posture task F(1,22)=11.395, p=.003, PES=.341, but it did not remain a significant effect in the colour task (*see Fig. 6).* Secondary analysis of SEPs over retention interval (300-900ms) using CSD values (µV/m2) was performed. As before, RM 2x2x2x2 ANOVA, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended side of the screen), Load (1 or 2 items) and Congruency (congruent or incongruent change between features) were entered for analysis. Results showed no significant effects.

3.3.3. Visual activity (VEPs)

Average waveforms were computed separately for each experimental condition, looking at posture or colour task, memory load 1 or 2 items, contralateral or ipsilateral activity, and congruent or incongruent change between stimuli features. Values from Mean amplitudes (µM) and Current Source Density (µV/m2) were used for parallel analysis at each time interval.

3.3.3.1. Attentional component (200-300ms).

Visual activity (VEP) was extracted as mean amplitudes from the time window corresponding to 200 to 300 ms. Repeated Measures ANOVA was used for analysis, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended side of the screen), Load (1 or 2 items) and Congruency (congruent or incongruent change between features). Results showed an overall effect of load F(1,22)=15.55, p<.001, PES=.414, meaning there was a sustained difference in visual activity modulated by memory load *(see figure 10).*

Same RM ANOVA analysis was performed on CSD values (µV/m2), finding no significant differences on visual activity for the 200 to 300ms time interval.

3.3.3.2. CDA (300-900ms)

Visual activity (VEP) was extracted as mean amplitudes from the time window corresponding to 300 to 900 ms. Repeated Measures ANOVA was used for analysis, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended side of the screen), Load (1 or 2 items) and Congruency (congruent or incongruent change between features). Results showed an overall effect of load F(1,22)=9.728, p=.005, PES=.307, meaning there was a sustained difference in visual activity modulated by memory load *(see figure 10).*

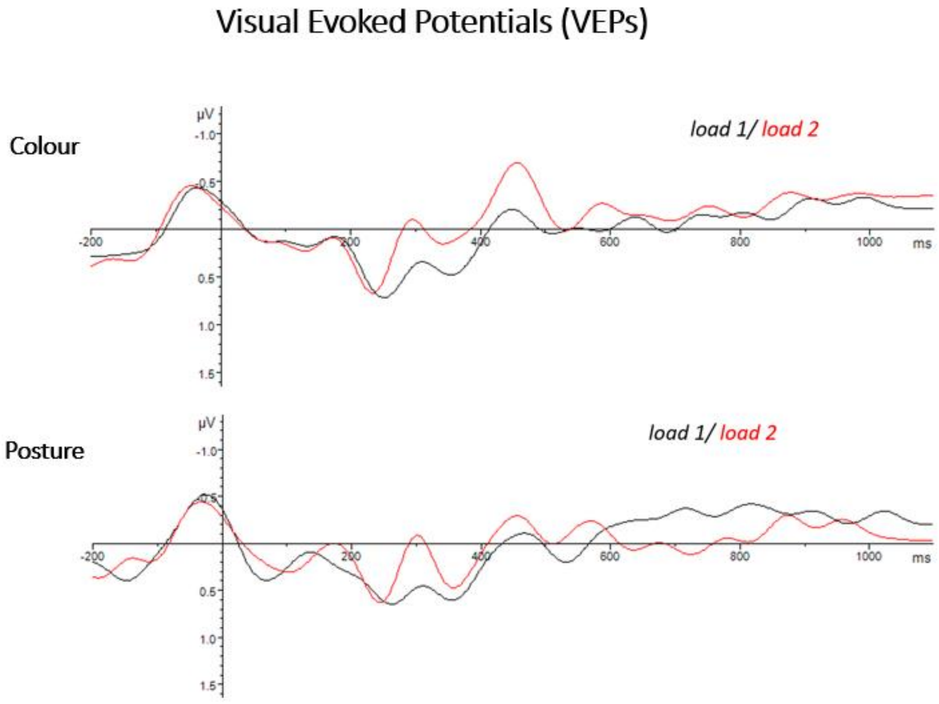


Figure 10. VEPs showing modulation by memory load for each tasks on visual-only trials.

Secondary analysis of VEPs over retention interval (300-900ms) using Current Source Density values (µV/m2) was performed. As before, RM 2x2x2x2 ANOVA, with factors Task (colour or posture), Laterality (contralateral or ipsilateral to the attended side of the screen), Load (1 or 2 items) and Congruency (congruent or incongruent change between features) were entered for analysis. Results showed a significant interaction between laterality and congruency F(1,22)= 4.512, p = .045, PES = .170. There was a laterality difference in congruent trials but not incongruent, although follow-up analysis on this interaction showed no significant simple main effects using Bonferroni corrections (all p.>05). VEPs were visually inspected and mean CSD values were plotted to further understand the interaction *(see figure 11).*

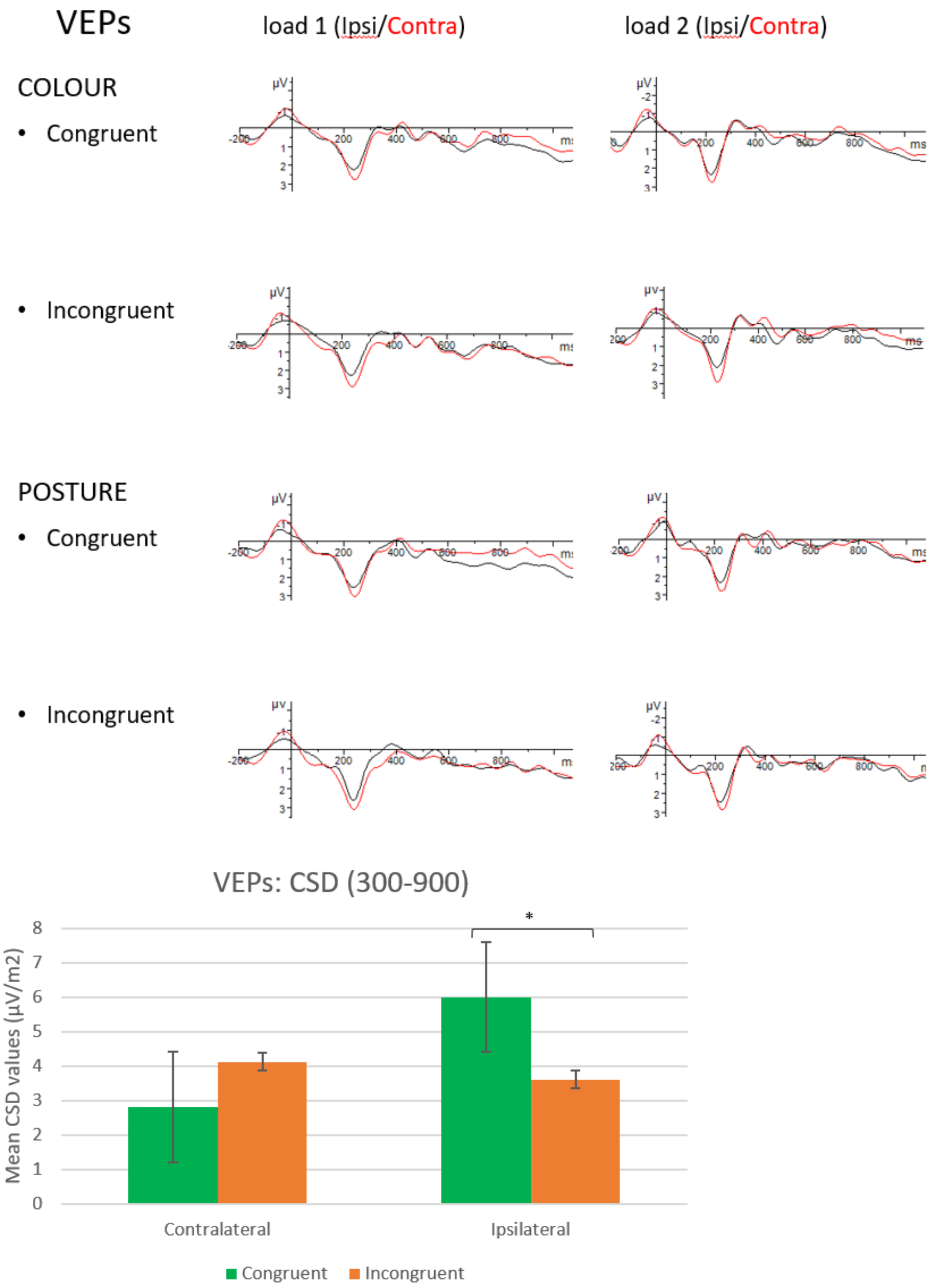


Figure 11. Top: VEPs average waveforms visualised for each experimental condition. Bottom: Bar graph shows mean CSD values, explaining interaction between congruency and laterality.

3.4. Discussion

The behavioural results from this study already showed a variety of effects. Both analysis on accuracy and sensitivity revealed better performance with congruent trials and low memory load, meaning the least parameters changing at a time, the easier it was for participants to maintain the information in WM. It was also evident that participants' performance when holding 2 items in WM showed a task difference, where the colour feature was remembered better than the posture. Interestingly, sensitivity analysis also showed there was a higher discriminability for colour features, which explains not only why colour was less affected by increasing memory load, but also how, as the colour feature might be more salient. Interpreting behavioural results, and according to our third hypothesis, it could already be argued that there was a saliency effect from the colour feature (bottom-up attention), either facilitating performance during colour task or interfering with posture task (top-down attention).

Neurophysiological data from VEPs showed that visual activity was modulated by memory load, both at the attentional window and as persistent activity during maintenance. Overall, this could be interpreted as visual areas engaging in maintenance of visually presented stimuli. Results from visual activity also showed a persistent laterality difference in congruent trials but not for incongruent ones, meaning visual maintenance was more prominent during congruent trials. This demonstrates different engagement of visual areas modulated by attention. In line with our second hypothesis and based on behavioural results showing a saliency effect for colour, this could be interpreted as stronger visual recruitment when the task could be resolved by attending to the colour feature (both colour task, and congruent trials in posture task).

SEPs showed overall somatosensory engagement modulated by memory load, stronger in the contralateral hemisphere, during the whole the retention interval. These results could be interpreted in line with our first hypothesis, where it was predicted an overall somatosensory engagement on WM for body stimuli. During the attentional window, somatosensory cortex showed different patterns of engagement during congruent and incongruent trials. This demonstrates different engagement of somatosensory areas modulated by attention. Congruent trials showed a modulation over contralateral somatosensory activity depending on the attended feature. Incongruent trials revealed a main difference on somatosensory engagement modulated by the attended feature. This result could be interpreted, in combination with the analogous result observed for visual engagement, as strong support for the second hypothesis, where it was predicted that attention could modulate the engagement between visual and somatosensory areas. Incongruent trials also showed stronger laterality differences and a modulation by memory load when attending to colour. When looking at the persistent activity over the retention interval on somatosensory areas, SEPs revealed a strong modulation by memory load. Overall, this could be interpreted as somatosensory areas engaging in maintenance of body-related stimuli. Planned follow-up analyses by task showed that memory load played a bigger role during the maintenance of posture features compared to colour. This result could be interpreted as confirmatory evidence for the second hypothesis, showing body-related areas of the brain were critically involved in maintaining body-related features, while not so critically non-body-related features.

The behavioural data generated in this study adds nuance to our understanding of the interplay between attention and memory load in Visual Working Memory (VWM).

Congruent with Oberauer's delineation of working memory into 'slots' (Oberauer, 2009), the performance metrics indicated that congruent trials and low memory load optimized the retention of information. This may be understood as a specific manifestation of Oberauer's principle, where attentional control dictates the nature and amount of information retained in these 'slots' (Oberauer & Hein, 2012).

Interestingly, the study's neurophysiological data buttresses and extends the existing literature on the recruitment of specific brain areas for VWM tasks. Our study aligns with previous research by Galvez-Pol et al. (2018a; b) indicating neural engagement in body-related cortices during VWM tasks involving body-related images. Furthermore, the stronger recruitment of visual areas when attending to colour features corroborates Downing et al.'s (2001) work on specialized neural substrates like the extrastriate body area (EBA).

The modulation of somatosensory and visual areas by attention, particularly in congruent and incongruent trials, lends credence to the study's second hypothesis. This is consistent with seminal work on top-down modulation in VWM and its influence on feature-specific attention (Serences et al., 2009; Harrison & Tong, 2009). Our findings also echo Todd & Marois (2004) and Lepsien & Nobre (2007), in that they reflect the complex interplay between cognitive load and attentional focus, reaffirming the dynamic nature of attentional allocation.

The study's third hypothesis concerning the impact of bottom-up saliency effects was substantiated by the behavioural data. Specifically, colour features emerged as more salient, corroborating Gazzaley & Nobre's (2012) notion of the suppressive role of attention in enhancing the fidelity of stored data in VWM. This suggests a nuanced interaction between bottom-up and top-down attentional mechanisms, enriching our understanding of the competing influences that modulate visual attention and, consequently, memory performance.

In conclusion, this study provides an integrative understanding of how top-down and bottom-up attentional mechanisms jointly modulate encoding and maintenance in WM, contributing to the growing body of research on body representation and attentional control. Future work may extend these findings by incorporating more diverse stimuli and examining additional neural markers of attention-based rehearsal.

Chapter 4: Investigating motor encoding of body-related information using TMS.

NOTE: In this Thesis we prepare the initial description of the study and our predicted results. While we were waiting for ethics approval to initiate the experiment in early 2020, the TMS laboratories closed due to the pandemic covid-19 and the data for this study has not been collected. The TMS lab did not re-open on time to collect data for this thesis as the TMS head lab (Dr Kielan Yarrow) was reallocated overseas during this period and his following sabbatical.

Abstract

Research in visual memory has highlighted the significance of perceptual regions in short-term memory encoding and maintenance. Recent studies employing electroencephalography (EEG) have indicated that brain areas beyond visual cortices, such as somatosensory and motor cortices, participate in encoding and retaining bodily information. This study aims to investigate the role of motor regions in encoding and maintaining visual memory of bodily information through transcranial magnetic stimulation (TMS). The following study proposes a TMS experiment with multiple sessions, where magnetic pulses are applied during the retention interval of a visual memory task involving body (hand images) and non-body images (geometrical shapes). Motor evoked potentials (MEPs), reaction time, accuracy, and D prime will be recorded. TMS will be applied over the motor cortex at different intervals, following a within-subjects design with factors including load, stimuli type, interval, visual cue, and TMS stimulation site. The study predicts that MEP

amplitude and task performance will be affected by the type of stimuli being processed, with lower performance when dealing with hand images compared to geometrical shapes. An interaction between stimuli type and load is expected, where TMS over the motor cortex will disrupt performance more in higher load conditions for hand images, highlighting the role of sensorimotor processing in encoding bodily information. Overall, this study aims to demonstrate the causal role of the motor cortex in maintaining bodily information in working memory, challenging existing cognitive models that traditionally overlook the involvement of primary sensory areas, like the motor cortex, in cognitive processes. The findings will contribute to a more comprehensive understanding of the role of motor cortices in cognition beyond their established functions in body perception and action execution.

4.1. Introduction

Working memory (WM) is a fundamental aspect of cognition that allows relevant information to be manipulated and actively maintained in memory (Baddeley, 2003). Recent accounts postulate that temporary maintenance of stimuli in WM occurs in brain areas such as sensory cortices, which also process these same stimuli in the absence of working memory demands. This provides sensory cortices with a strong role in WM and characterizes it as a reestablishment of the perceptual experience (Tsubomi et al., 2014; D'Esposito & Postle, 2015). Specifically, sensory cortices have shown a modulation by memory capacity performance locked to the modality in which information to be remembered has been perceived (Vogel & Machizawa, 2004; Vogel et al., 2005; Katus et al., 2015) (e.g. increasing visual cortex activity when maintaining a higher number of visual stimuli in visual memory -such as colours and shapes).

Importantly, the nature of the perceived stimuli strongly influences where and how relevant information is processed (Kanwisher 2010; Pitcher et al., 2009; Urgesi et al., 2007a), providing additional cortical regions tightly linked to an early sensory response with functional properties related to human behaviour (Sel et al., 2014; Meyer et al., 2011; Postle 2006). For example, it has been shown that beyond those sensory areas originally used to perceive visual information (i.e., visual cortex), when the visual stimuli relate to body images and actions, its perception elicits activity in regions representing our own body and actions such as sensorimotor areas (motor and somatosensory cortices, beyond visual regions) (Calvo-Merino et al., 2005; 2006; Urgesi et al., 2007a). Research has shown that not only body-related brain areas play a fundamental role in action perception and action execution (i.e., sensorimotor cortices), but also they have been linked to crucial processes in human cognition such as theorizing about others' mental states (Rizzolatti et al. 2001; Rizzolatti and Sinigaglia 2008), understanding kinematics implied in perceived movements (Calvo-Merino et al. 2005; Cross et al. 2009) or participating in human judgements such as aesthetic decisions (Calvo-Merino et al., 2010) or action predictions (Aglioti et al., 2008).

Interestingly, all the above-mentioned processes involve not only perceiving visual stimuli but maintaining this information in working memory while further decisions are taken (i.e. judgements). However, until recently, no studies have studied the direct role of sensorimotor cortices in visual working memory of bodies or actions. During the last years, Galvez-Pol and colleagues have developed a new methodology using electroencephalography (EEG) that combines visual evoked potentials (VEP) recorded during a visual working memory task, with neural signatures classically associated with somatosensory and motor cortices, namely somatosensory evoked potentials (SEP) and lateral readiness potential (LRP) (Galvez-Pol et al., 2018a; Galvez-Pol et al., 2018b). This methodology has been extensively described in our previous chapter and has been replicated in the earlier studies presented in the thesis (Abad-Hernando et al., *in preparation*). These studies have allowed us to find initial evidence for the role of sensorimotor cortices in holding visual body-related information in memory. Specifically, using EEG, Galvez-Pol et al (2018a) recorded brain activity while participants maintained in memory visually presented body postures and control stimuli (matched geometrical shapes) in a similar paradigm to those employed by visual working memory studies (Vogel & Machizawa, 2004; D'Esposito & Postle, 2015). These results show that compared to non-body-related stimuli (shapes), maintaining bodies in memory involves initial visual processing but importantly this process is fast recoding in our own neural body representation, as shown by a modulation of the somatosensory evoked potential recorded simultaneously (by applying irrelevant tactile stimulation to the fingertips) during the visual task (Galvez-Pol et al., 2018a). Similarly, a second study showed a similar pattern of results (stronger modulation during the maintenance of body information than for control shapes) when recording lateralized readiness potential (a classic neural signature of motor cortex activity) during the same visual working memory task (Galvez-Pol et al., 2018b). Crucially, these studies employed a newly developed methodology that allows isolating the specific involvement in somatosensory and motor cortices from the visual carry over effects during the visual task, by using a subtraction methodology to cancel out all potential activity related to exclusively visual processing carry-over effects (for a methodological description see Galvez-Pol et al. 2020).

These studies provided the first evidence for a new role of sensorimotor regions not only in perception by also in encoding and maintaining body-related information in a visual working memory task. However, what remained unclear is whether this activity is indeed necessary (i.e. causally involved) for memory encoding or rather just a correlated signal.

The current study aims.

This project aims to investigate using transcranial magnetic stimulation (TMS) if motor regions involved in body perception and action observation play a crucial role in encoding and successfully maintaining memory bodily information. Or if this sensorimotor response just reflects an ongoing carry-over activity that correlates with the visual memory mechanism but does not play a functional and causal role in visual memory for body information. The project proposes a main TMS experiment divided into several sessions. The study will employ TMS to measure ongoing activity in the motor cortex by applying magnetic pulses during the retaining interval during a visual memory task when the stimuli to be remembered are body (hand images) and non-body images (geometrical shapes) *(Figure 12).* We will use a change detection task with memory arrays presenting 1 or 2 items (low load, high load; following Galvez-Pol et al., 2018a; 2018b, Fig S2). Motor evoked potentials (MEPs), reaction time and accuracy will be recorded as dependent variables. D prime will also be calculated for the analysis. We will apply single-pulse TMS at a rate of 1 every 5 seconds, at different intervals over the 600 ms retention interval, over the motor cortex and record motor evoked potentials (MEPs). We plan a within-subjects design 2x2x2x2x2 that will include the following factors: load (1/2 items), stimuli (hands/control shapes used in Galvez-Pol et al., 2018a), interval (we will apply the magnetic stimulation at

500 and 700 ms SOA -based on Galvez-Pol et al., 2018b) visual cue (presented in the left or right visual hemisphere) and TMS stimulation site (right/left motor cortex).

Similar predictions are made for modulation of the MEPs amplitude collected over left/right-hand muscle during the visual memory task, and performance in the visual memory task. Specifically, we predict a main effect of stimuli type, suggesting that MEP amplitude will be modulated by the type of stimuli being processed, and performance will be lower when information to be maintained are hands than geometrical shapes. This will imply that interfering with motor processing via TMS affects memory when information to be remembered is bodily related (hands) but not for control geometrical shapes. Following this main effect, we expect an interaction of stimuli type and load. TMS over the motor cortex will modulate MEPs and disrupt performance encoding and memory of hand images, but not geometrical shapes, and this disruption will be stronger in the higher load than in the lower load. These predictions follow nicely on previous EEG work results and will suggest that sensorimotor processing participates in encoding bodily information. These results will thus identify a crucial role for the motor cortex in working memory, over and above those processes provided for visual regions.

Overall, the present study will show that motor cortices are relevant in cognition (i.e., memory) beyond their well-established role in body perception and action execution. Demonstrating the causal role of the motor cortex in maintaining bodily information, will be a stepping stone towards revising existing cognitive models that generally do not consider a role for primary sensory areas (i.e., motor cortex) beyond the basic processing of sensory information.

4.2. Methods

4.2.1. Design

The experimental design will be a within-subjects repeated measured (2x2x2x2x2) and will include the following conditions:

A) Stimulus type (2): Visual stimuli to be remembered will be body (hand images) and non-body images (visually controlled stimuli: geometrical shapes). We have previously validated and visually matched the stimuli for physical visual properties and task difficulty (Galvez-Pol et al., 2018a; 2018b; see Fig 1).

(B) Memory load (2): number of stimuli to be remembered in the target array (1 or 2). This represents low load vs high load. Prior studies have shown that remembering two items may well lead to limits in WM capacity (Alvarez and Cavanagh, 2004; Luria et al., 2010; Olsson and Poom, 2005); therefore, memory load 1 and 2 (low/high) would allow observing increasing activity related to memory encoding and maintenance.

(C) Lateralized cues (Left, Right): As previous work on visual short-term memory (Vogel & Machizawa, 2004) or our previous work (Galvez et al., 2018a; 2018b), it is necessary to cue one visual hemifield to show differences between ipsilateral and contralateral neural processing of the stimuli to be remembered in the contralateral hemisphere. At trial onset, visual cues ((left/right arrow) will indicate which stimuli should be remembered (left/right visual hemifield).

(D) TMS stimulation site: single-pulse TMS will be applied over the left and right motor cortex in different condition blocks. The rationale behind this choice is based on the clear contralateral brain representation of the hands in the motor cortex. This allows the possibility of measuring persistent contralateral activity over motor cortices when seeing and remembering the stimuli (in a similar fashion to the EEG paradigms employed by Vogel et al., (2005).

(E) TMS timing (3): TMS will be applied over motor cortex at two different intervals (we will apply the magnetic stimulation at 500 and 700 ms SOA). This timing information is based on the timing effects observed in our previous work with EEG (Galvez-Pol et al., 2008a).

Dependent Variables: We are going to handle two types of data that will be analysed separately. Physiological data will be obtained from the TMS over left and right motor cortex (in the shape of motor evoked potentials). Behavioural data we will obtain from the visual working memory task. Here we will explore reaction times, and accuracy and we will calculate d prime to assess perceptual sensitivity.

4.2.2. Materials

Participants will perform a visual memory task similar to that of Vogel and Machizawa (2004) and Galvez-Pol et al., (2018a). Items to be remembered were hand images (depicting different finger/hand positions) and analogous geometrical shapes (Fig. 1B). Participants were cued on each trial by a central arrow to attend to items displayed in their left or right hemifield. This was followed by a bilateral memory array depicting 1 or 2 items (low and high memory load conditions) in each hemifield and a blank retention interval lasting 900 ms. A final test array that differed in 50% of the cases from the memory array by one item was displayed until participants verbally responded whether or not the memory and test arrays were identical (Fig. 1A). The participants' verbal responses were reported through a microphone to the experimenter, who entered the responses manually from outside of the electromagnetically shielded room. Participants' forearms rested on the top of a table with their hands separated in palm up position while covered by a black surface. Visual stimuli were displayed using E-Prime2 Software (Psychology Software Tools, Pittsburgh, PA).

All stimulus arrays were presented within two 5 x 8.5 degrees rectangular regions that were centred 5.3degrees to the left and right of a central fixation cross on a grey background. The positions of all stimuli were randomized on each trial with the constraint that the distance between stimuli within a hemifield was at least 2.4 degrees (centre to centre). Each memory array consisted of 1 or 2 hands (1.3 x 0.8 degrees) in each hemifield. These were randomly selected from a set of twelve hands. Right hand images were shown on the right hemifield while left hand images were displayed on the left. The rationale behind this latter choice is based on the clear contralateral brain representation of the hands in the somatosensory cortex. This allows the possibility of measuring persistent contralateral activity over SCx, which can be isolated from concomitant visually evoked activity when seeing and remembering the stimuli. In the control condition, 1 or 2 polygonal shapes (1.3 x0.8) were selected and shown in a similar fashion. Since prior studies have shown that remembering two items may well lead to limits in WM capacity (Alvarez and Cavanagh, 2004; Luria et al., 2010; Olsson and Poom, 2005); memory load 1 and 2 (low/high) would allow observing increasing activity related to memory encoding and maintenance.

During the trials, a task irrelevant TMS pulse will be applied over left or right motor cortex at a given interval (500 or 700 ms) and MEPs over the contralateral hand to the TMS stimulation site will be recorded. On a basis of 100 trials per condition and assuming a design 2x2x2x2, there will be 1600 trials (1 pulse/5sec x 1600 = 2.2 hours approx.) with 4 blocks and probably 2 sessions of 1.5 hours (number of sessions will be pilot to ensure we respect health and safety regulations). Based on previous EEG studies finding effects using same paradigm and design (Galvez-Pol et al., 2018a,b) we estimated a sample size of 40 participants.

We have collapsed two factors over one, to create a congruency factor between the visual stimulation site (left, right hemifield) and TMS stimulation site (right, left motor cortex).

TMS stimulation and EMG recording: We will use a Magstim Super Rapid Stimulator and a 70 mm figure-of-eight coil and apply single-pulse TMS over the primary motor cortex during the visual working memory task. We will employ a classical measurement of motor activity: Motor Evoked Potentials (MEPs). When single-pulse TMS is applied over the primary motor cortex, this will activate descending motor pathways and result in MEPs (i.e. muscular twitches). The intensity of pulses will be set around 110–120% of resting motor threshold (RMT) to elicit MEPs of around 1 mV amplitude in the FDI. Individual RMTs will be determined prior to the experiment as the minimal intensity required to elicit an MEP 50 mV in amplitude in around 3 out of 6 single pulses when the hand is fully relaxed. The MEPs will be recorded using electromyography (EMG, via electrodes placed on the skin over the relevant hand muscles) to measure covert motor processing. Specifically, surface Ag/AgCl EMG electrodes will be placed approximately 2–3 cm apart, over the first dorsal interosseous FDI muscle of the left and right hand (to record MEP from TMS over left and right hemisphere) and a nearby reference site. EMG will be collected and stored on a second dedicated PC. Participants will be instructed to perform the visual working memory task and ignore the TMS stimulation. Digital data will be exported and analysed offline.

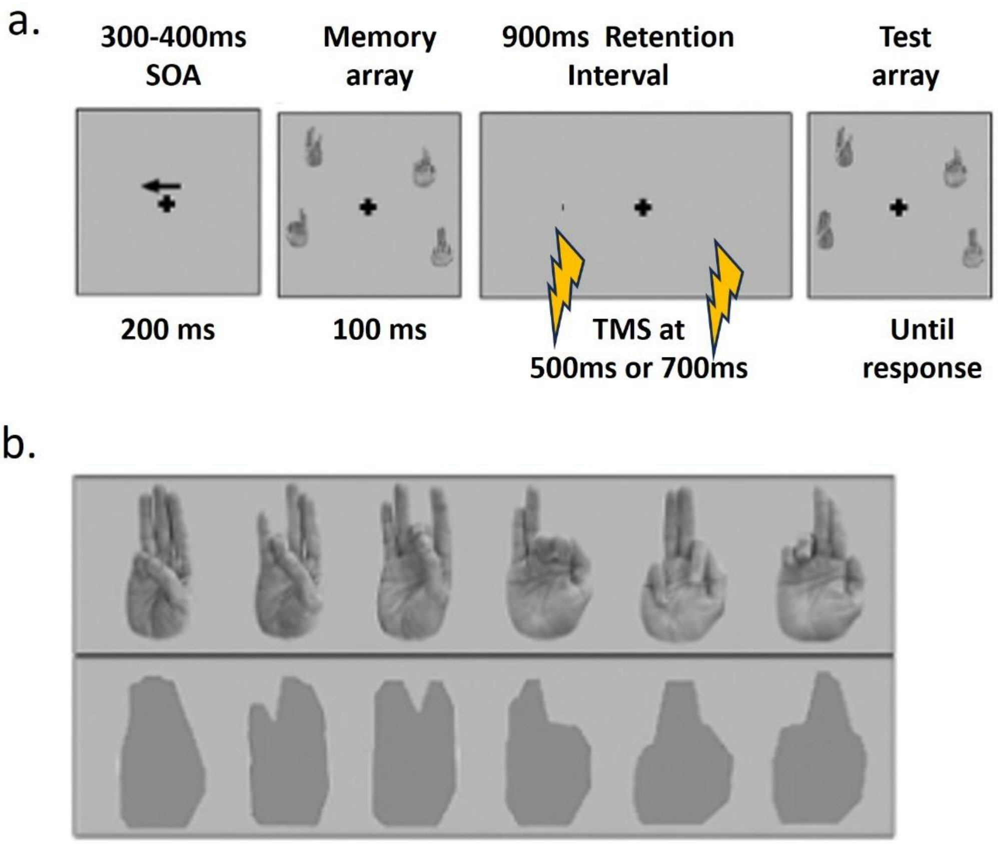


Figure 12. (a) Trial structure and stimuli example. (a) example of a trial indicating the TMS timings (500 or 700 ms) during the retention interval. (b) example of hands and shapes used as stimuli. (Adapted from Galvez-Pol 2018a).

4.3. Predicted results & Discussion.

Here we present a description of the main planned analysis and prediction. For the analysis, Repeated Measures ANOVA will be used for behavioural (accuracy), sensitivity (d-prime) and physiologic data (MEPs) separately. We will calculate performance (accuracy, RT a d-prime -visual sensitivity index) for each type of stimulus, load, lateralized cue, stimulation time and stimulation site). We will combine the factors lateralized cue and stimulations site, to create a new variable that represents the congruency between the visual hemifield that hold the visual information (left, right) and the TMS stimulation site (right, left respectively - as visual information is cross over) (similar to Galvez-Pol et al., 2018a, 2018b). We predict a significant modulation of the MEP amplitude affected by the interaction between the type of stimuli to be remembered (hands, shapes) and the memory load (high, low load). Importantly, this interaction will be accompanied by a main effect type of stimuli, suggesting that the activity evoked in the motor cortex via MEPs is affected by the type of stimuli that is been maintained in memory (hands or shapes). A main effect or interaction with the factor time of stimulation will provide us with information about the timing of this effect. Similarly, we predict a significant decrease in performance (increased RT, reduced d prime) during the hand condition as compared to the shape condition. This stimuli effect can also interact with load, as we have described above, suggesting the motor processing is not only sensitive the type of stimuli being encoded and maintained but also is sensitive to the amount of information being stored when this is bodily related (but not when the stimuli are shapes). A main effect or interaction with the factor stimulation time will allow us to see at what time of the retention interval is more critical the involvement of motor cortex in this kind of processing. Overall, this will establish how sensorimotor processing plays a crucial and necessary role in visual working memory for bodily information. This experiment will provide direct causal evidence for the role and necessary involvement of motor areas during a visual working memory task. This project will show that sensorimotor cortices are relevant in cognition (i.e. memory) beyond their well-established role in body perception. This will be a stepping-stone towards revising existing cognitive models that generally do not consider a role for the primary sensory area beyond the basic processing of sensory information. A potential pattern of results for accuracy is presented in *figure 13.*

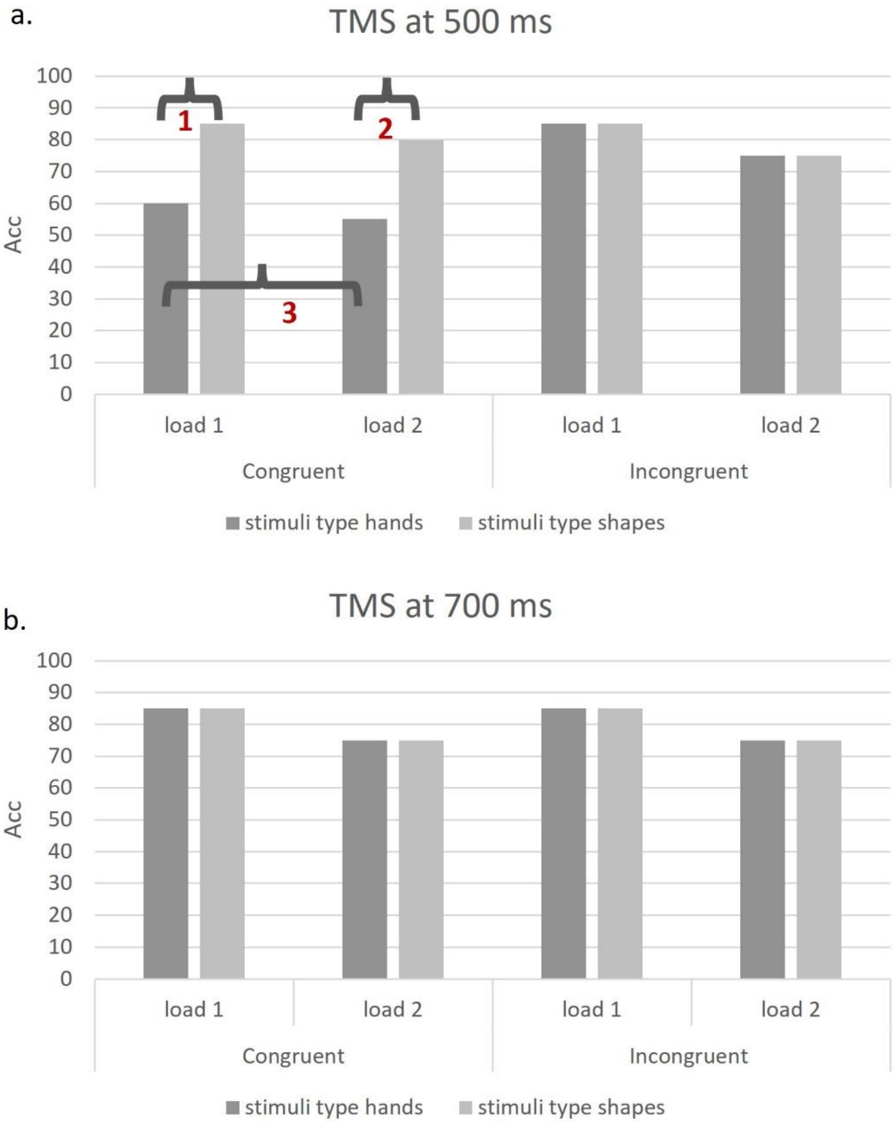


Figure 13. Figure illustrating a potential performance pattern for accuracy during the two TMS conditions. (a) condition TMs stimulation at 500 ms. The figure shows an interaction of load and stimuli type, showing lower performance for hands than for shapes (value 1 and 2). There is also a significant difference in the load only for hand conditions (value 3), suggesting that the TMS at 500 ms affect hands in general, but it is also modulated by load. No significant effects were observed when the TMS stimulation was not congruent with the visual field to be stimulated. This scenario shows effects at the early part of the encoding process (500 ms). (b) This scenario shows similar effects to the body and shapes during the TMS stimulation at 700 ms. Differences between TMS timing will allows to understand the temporal pattern of the sensorimotor encoding in visual WM. The figures proposed this may be relevant to the early involvement.

SECTION B: Action Prediction & Movement Fluency

In cognitive psychology, action prediction and movement fluency serve as interconnected pillars that help illuminate how perception and action are seamlessly integrated in human behaviour. Action prediction refers to our ability to foresee future events by interpreting sensory data, a skill that equips us to navigate a continually changing environment effectively. Motor fluency, conversely, pertains to the smooth and efficient execution of physical tasks, which often improves through practice and learning. These two components are interdependent: accurate action prediction can lead to smoother motor responses, and a high degree of motor fluency can offer more reliable data for effective action prediction. As Ericsson (2014) and Ericsson and Pool (2016) have argued, the mastery of a skill relies on identifying key performance elements, receiving immediate feedback, and continual refinement through repetition and problem-solving.

Biological motion as a sensory-motor construct. The concept of biological motion (BM), initially defined by Johansson in 1973 as the unique movement pattern of living organisms, plays a pivotal role in linking perceptual and motor fluency. Recognizing and understanding this form of motion is essential for survival and social interaction. One of the most common ways to study biological motion in action prediction is using point-light displays, as a way to disentangle biological kinematics from biological appearance. Originally developed also by Johansson in 1973, point-light displays are visual representations where key joints of a human or animal figure are marked by small, luminous dots against a dark background. The movement of these dots simulates the natural movement of the body, stripping away all additional visual cues like facial expression, attire, or body shape. Point-light displays are particularly useful for studying action prediction because they isolate the raw kinematic data, enabling researchers to discern which specific aspects of movement are crucial for accurate prediction. For example, in the field of sports psychology, researchers have utilized point-light displays to evaluate how skilled athletes are better at predicting the outcome of an opponent's actions merely by observing these simple visual cues (Aglioti et al., 2008; Urgesi et al., 2012). Similarly, studies involving dance have used point-light displays to investigate how professional dancers are more adept at predicting complex choreographed sequences compared to novices (Bläsing et al., 2012; Orgs et al., 2013). Importantly, point-light displays have also been instrumental in neural studies exploring the mechanisms of action prediction. Researchers use techniques like functional magnetic resonance imaging (fMRI) or electroencephalography (EEG) to monitor brain activity while participants watch point-light displays (Balser et al, 2014; Zhao et al., 2021). This approach has highlighted the roles of the Action Observation Network (AON) and mirror neuron systems in facilitating predictive processes. Overall, point-light displays offer a streamlined yet effective methodology for understanding the nuances of biological motion and action prediction, significantly contributing to our understanding of how humans interpret and interact with the world around them.

On the perceptual front, humans excel at rapidly and accurately recognizing biological motion, whether it's discerning the gait of a person from minimal cues (Blake & Shiffrar, 2007) or decoding subtle facial expressions (Ekman & Friesen, 1971). This aptitude for quick recognition, an element of perceptual fluency, also enables us to predict future actions like the path of a ball based on a player's kicking motion. On the flip side, executing intricate movements, such as dance sequences or athletic feats, requires a high degree of motor fluency (Sevdalis & Keller, 2011). Motor fluency can be refined by engaging in a feedback loop that involves observing biological motion, either of oneself or others, and subsequently practicing to improve. This observation process is thought to activate neural pathways that are also engaged during the execution of the observed movements, a phenomenon known as motor resonance (Rizzolatti & Craighero, 2004; Calvo-Merino et al., 2005; 2006). This dual activation may encourage the mutual enhancement of both perceptual and motor fluency. Clinical conditions like autism spectrum disorder and Parkinson's disease, which affect both the perception and execution of biological motion, underscore the essential interconnectedness of these two forms of fluency (Cook et al., 2009; Casjens et al., 2013). They serve as a reminder of how deeply entwined they are in our understanding and interaction with the surrounding world, revealing a critical connection between the two, potentially mediated by shared neural mechanisms or networks.

This upcoming section will focus on the dynamic nature of bodies as a key characteristic of how we understand and predict actions. Chapter 5 was written as a review paper on action prediction, based on latest state of the art, proposing an integrative and holistic view on the field from a combination of cognitive, behavioural, and neuroanatomic and computational perspectives. Action prediction refers to the cognitive processes by which individuals anticipate and forecast the actions of others based on various cues and prior knowledge, and it involves making inferences about the likely next actions in each situation. Action Prediction will be explored based on 3 cognitive stages with different underlying mechanisms: sensory cues and detection of animacy, sensorimotor representation of actions as an internal experience-based model, and top-down modulations from intentions and expectations based on the context. Chapters 6 & 7 consist of a series of experimental studies where action prediction and fluency will be empirically explored using dance movements as complex stimuli, in an attempt to improve ecological validity and explore this multidimensional relationship in a more naturalistic way. The first study will stablish the behavioural and psychophysical baseline measurements and experimental conditions, building a robust paradigm to explore fluency from an action prediction framework. The second experiment will explore further the relationship between action prediction and fluency when comparing with experts, looking at aesthetic appreciation and individual differences as potential modulators adding to the richness and complexity of human movement.

Chapter 5 will elaborate on how action prediction stands at a complex intersection between attention, perception, sensorimotor representations, implicit memory, and contextual information such as socio-affective factors or cultural background. At this intersection, prediction means not just survival, but better adjustment in everyday life. Research on action prediction contributes to our understanding of social cognition, empathy, and theory of mind—the ability to attribute mental states to oneself and others. We are constantly anticipating other’s actions, making predictions about the most likely goal or intention, based on our personal experience and available information. Action prediction (AP) is based on sensorimotor representations of actions, including procedural memory for motor skills. But it also improves with contextual information, which aids in constructing top-down models that help us interpret everyday situations and interactions. First, we perceive sensory information in the environment and our attention might be driven towards certain characteristics implying animacy. Those are generally key features like faces, biological motion (BM) or interaction contingency that would indicate agency behind movements. Agency and animacy detection via those attentional and perceptual cues would be the first processing stage guiding our understanding towards goal-directed actions. Secondly, we use our internal sensorimotor representations of those actions, which are constructed based on our own experiences of perceiving and performing them. In this sense, action prediction is built upon both action perception and action execution. These models have been developed through sensorimotor experience, adjusted using self-motion interoceptive information and consolidated in memory. Although, predicting actions goes beyond simply forecasting the physical movements, it also involves inferring the intentions and goals underlying those actions. Individuals use their context-based knowledge of the situation, cultural norms, and the socio-affective background information, to make more accurate predictions about others' intentions. Prediction Error (PE), a key concept from the predictive coding account, is proposed here from different angles depending on the type of information, top-down modulations would feed into the sensorimotor loop or broader social cognition frameworks.

Chapter 6 will explore fluency, a multifaceted concept, which has been shown to significantly impact various aspects of human cognition and behaviour, from perception and motor execution to memory and decision-making. Perceptual fluency relates to the ease with which sensory information is processed, influencing judgments about aesthetics, truthfulness, and memorability. Motor fluency, on the other hand, deals with the execution of physical tasks and is indicative of skill level in diverse domains like sports and music. These two types of fluency are not mutually exclusive; rather, they often interact in complex ways. For example, the study of biological motion serves as a unique crossroads where perceptual and motor fluency meet. It's not just about how effortlessly one can interpret or predict biological movements (perceptual), but also how seamlessly one can execute them (motor). Observing such motion can even activate neural pathways used in executing the same movements, highlighting a bidirectional relationship between perception and action. This first behavioural study aims to delve into this intricate relationship by focusing on the intersection between perceptual and motor fluency through an action prediction task. The focus will be on human dance movements, providing ecological validity while also challenging participants with complex whole-body movements they might not have motor experience with. The study proposes that participants will more accurately and quickly identify fluent movements as the "correct" continuation in a temporal occlusion paradigm, suggesting that fluency has both perceptual and motor implications. This work is part of a broader research trajectory that aims to explore the nuances of fluency and its impact on our cognitive and behavioural functions, offering insights that could extend to various domains including social neuroscience, psychology, and even clinical applications.

Chapter 7 elaborates on fluency and outlines an experiment designed to explore the relationship between aesthetic appreciation and action prediction in the context of dance. Providing not only a review of existing literature that highlights the role of perceptual and motor experience in predicting others' actions, particularly in skilled performances. There's also evidence that experts and non-experts engage different neural regions when watching dance, depending on their aesthetic preferences and other individual differences. Movement fluency, or the smoothness and precision of dance moves, is emphasized as a key factor in how audiences perceive and appreciate dance. Fluent movements are linked to a dancer's emotional expression, skill level, and the aesthetic quality of the performance. This applies universally across different styles and cultural forms of dance. The field of neuroaesthetics provides a backdrop, studying how the brain processes aesthetic experiences. This work builds on previous research showing that body-related brain areas play a role in appreciating the aesthetic aspects of bodily movement in space. The experiment aims to measure whether aesthetic appreciation of dance enhances our ability to predict fluency in dance movements. It also seeks to investigate whether individual differences, such as dance experience or alexithymia, affect this predictive ability. Two hypotheses are proposed: one suggesting that performance will be better on fluent trials and the other exploring if aesthetic appreciation or other individual differences could relate to, or even predict, performance.

Chapter 5: The Advantage of Prediction: a review

5.1. Introduction

Action, defined as a goal-directed movement, ranging from walking to communication, can be better understood through the lens of prediction. To navigate the world around us, efficient visuomotor control of our actions requires the estimation (on-line prediction) of one's own body state prior to movement execution, which is based on internal sensorimotor representations. These internal representations or models, allow for anticipating the sensory consequences of one's own movements in real time based on motor commands (i.e., efference copies) (Wolpert & Flanagan, 2001). In the context of social interaction, it is imperative to anticipate the forthcoming actions of other individuals or agents. Those same internal models may also be applied to predict the actions observed in others (Blakemore & Frith, 2005; Kilner, Friston, & Frith, 2007; Prinz, 2006; Wolpert & Flanagan, 2001).

Perceiving, understanding, and contextualizing actions occur at different levels, which can be utilized to predict events and determine the need for action. These processing levels appear to follow a temporal order, beginning with the perception of low-level sensory features, progressing to the understanding of action kinematics and goals, and culminating with broader conceptualizing via contextual information such as social-affective frameworks (Tarhan et al., 2021; Dima et al., 2022).

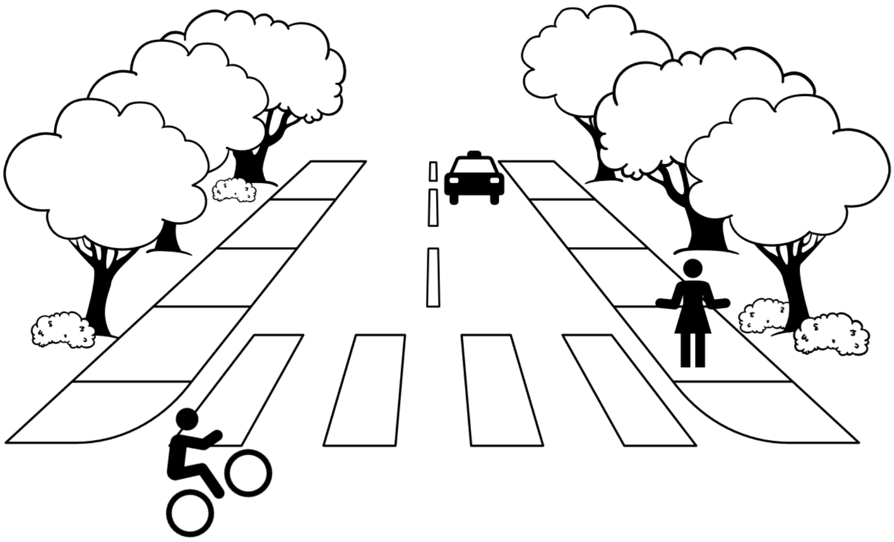


Figure 14. Action prediction scenario. Cyclist will first detect there is an agent on the right side of the road (bottom-up sensory cues). Then, they will use their own sensorimotor experience (internal representation / model of action) to read body posture and kinematics, trying to predict whether the person is more likely to move and cross the road or stand and stop the cab. At last, knowing that the person is standing by a pedestrian crossing (top-down modulation by context) will increase the probability of the cyclist predicting the action of crossing.

Action prediction is essential. Action stands at a complex intersection between attention, perception, sensorimotor representations, implicit memory and contextual information such as socio-affective factors or cultural background. At this intersection, prediction means not just survival, but better adjustment in everyday life. Research on action prediction contributes to our understanding of social cognition, empathy, and theory of mind—the ability to attribute mental states to oneself and others. We are constantly anticipating other's actions, making predictions about the most likely goal or intention, based on our personal experience and available sensory information. Action prediction (AP) is based on sensorimotor representations of actions, including procedural memory for motor skills. But it also improves with contextual information, which aids in constructing top-down models that help us interpret everyday situations and interactions. First, we perceive sensory information in the environment and our attention might be driven towards certain characteristics implying animacy (Saito et al., 2023). Those are generally key features like faces, biological motion (BM) or interaction contingency that would indicate agency behind movements (for a review, see Schultz & Frith, 2022). Agency and animacy detection via those attentional and perceptual cues would be the first processing stage guiding our understanding towards goal-directed actions. Secondly, we use our internal sensorimotor representations of those actions, which are constructed based on our own experiences of perceiving and performing them. In this sense, action prediction is built upon both action perception and action execution. The accuracy of these internal representations or models is enhanced by both visual and motor familiarity with the specific action (Calvo-Merino et al., 2006). These models have been developed through sensorimotor experience, adjusted using self-motion interoceptive information (e.g., oculovestibular feedback, proprioception, and motor efference copy), and consolidated as procedural memory. Although, predicting actions goes beyond simply forecasting the physical movements, it also involves inferring the intentions and goals underlying those actions. Individuals use their context-based knowledge of the situation, cultural norms, and the socio-affective framework, to make more accurate predictions about others' intentions. For instance, when a pedestrian stands by the roadside, a cyclist must predict whether that person intends to cross the road or maybe hail a taxi *(see Fig 1 for minimalist example scenario).*

The first section of this review will explore the concept of prediction error, brought from the intersection of action prediction and predictive coding theories, to understand how we update both sensorimotor and context models. The second section will look into perception of action-related sensory information, focusing on the detection of agency, and exploring the role of body-related features and biological kinematics as saliency cues. This is a first and necessary step where attentional and perceptual mechanisms will play a decisive role in recognizing “natural agents”. Throughout this section, moving “agents” will be referred as opposed to moving objects, like a bouncing ball or a falling leaf. Agency cues will determine perceptual pathways, with great impact on action prediction performance. A review on action prediction studies focusing on animacy, with an integrative view on biologically salient information processing will be discussed, including neuroanatomical basis and pathways key on agency detection. The third section will elaborate on the main body of research on action prediction, which is centred in action features and mechanisms underlying sensorimotor representations and memory for actions. It will focus on how we acquire our sensorimotor representations through experience, and why this experience-based sensorimotor knowledge is the most accurate model to make better predictions. A variety of approaches to study Action Prediction (AP) will be covered, from developmental studies to visual vs motor familiarity and expertise. We will discuss the motor hypothesis with the core AP experimental studies pointing towards Mirror Neurons (MN) and Action Observation Network (AON) as neural substrates. The last section will give an overview on the latest advances on Action Prediction (AP) research, moving towards understanding the role that contextual information can play modulating how we perceive and interpret actions. Framing action prediction within social and emotional contexts might have been the missing piece for a more holistic understanding.

5.2. A more integrative view using predictive frameworks.

5.2.1. Predictive Coding (PC), an integrative framework

Considering action from the perspective of prediction can provide a multi-faceted, coherent, and integrative framework that offers explanatory power across different levels of understanding, from neural mechanisms to behaviour and social interaction. Prediction allows for a more comprehensive understanding of how sensory information (e.g., visual or auditory cues) can inform and guide motor actions. This is a cornerstone of adaptive behaviour, also covered as the proactive brain hypothesis (Bar 2007; 2009). Predictive models suggest that the brain constantly makes forecasts about future states. If the brain can predict the outcome of an action successfully, it can more efficiently allocate cognitive and motor resources. A predictive framework is inherently adaptive, meaning it can adjust to new information or changes in context, which is essential for survival in a complex and dynamic world. There is growing evidence supporting the role of predictive coding in brain function (Parr & Friston, 2018; Friston & Frith, 2015; Aitchison & Lengyel, 2017; Palmer et al., 2020). Brain regions responsible for actions, like motor cortex or basal ganglia, are closely connected with areas responsible for prediction and expectation, such as prefrontal cortex (Fuster, 2001; Passingham et al., 2002).

The concept of action has been compellingly reframed through the predictive coding account, as articulated in the work of Karl Friston and colleagues. Under the predictive coding framework, actions are not mere responses to stimuli but are guided by the brain's ceaseless effort to minimize *prediction errors*—the discrepancies between predicted and actual sensory input. According to Friston's free-energy formulation, actions serve to adjust the environment or our perception of it, in such a way that future sensory input aligns with the brain's predictions (Friston et al., 2010). This perspective integrates seamlessly with other cognitive processes; for example, the same predictive models that guide our actions can also be used for interpreting the actions of others, a concept that gains traction in Clark's behavioural and brain sciences paper (Clark, 2013). In this view, our motor acts, as well as our perception of actions, emerge from a unified predictive framework, significantly enriching our understanding of agency, intentionality, and social interaction. This account of action informed by predictive coding offers a more nuanced understanding that encapsulates not just motor output but also intention, planning, and social understanding, supported by a brain network that is inherently geared for prediction.

5.2.2. Prediction Error (PE): the real deal for AP

Action prediction refers to the cognitive process through which individuals anticipate and forecast the actions of others based on various cues and prior knowledge. It involves making inferences about the ongoing action. Predictive error refers to the discrepancy or mismatch between the predicted action and the actual observed action or outcome and can be used as a learning signal to update the predicted action to match observed action more closely. People make anticipatory predictions about the future actions or events based on available information, contextual cues, and prior knowledge. Predictive error occurs when the predicted action or outcome does not align with what happens, and it represents the error or inaccuracy in the prediction made.

Prediction error involves anticipating the unfolding of observed actions (Avenanti et al., 2018; Stadler et al., 2012; Urgesi et al., 2010; Kilner, 2011). Neurophysiological evidence, such as the study by Whitford et al. (2017) on efference copies in inner speech, supports the concept of sensory attenuation, where self-generated sensations elicit a smaller neurophysiological response compared to externally generated sensations. Sensory attenuation is believed to occur due to predictions generated by an internal forward model (IFM) based on an efference copy of motor commands. The predicted sensations are then compared to actual sensations, and any discrepancy is processed at higher levels in the neural hierarchy. Consequently, self-generated sensations feel less salient than externally generated sensations. This phenomenon is thought to be mediated by the cerebellum and the IFM (Blakemore et al., 2000a; Wolpert and Miall, 1996).

An exceedingly encompassing and impactful postulate emphasizing the significance of prediction in neural functionality is the predictive coding theory, significantly developed and expanded by Friston and his team over the course of two decades (2002, 2005, 2010). The fundamental tenet of predictive coding posits that neural computation involves the gathering of substantiating data for competing anticipations about the causes of sensory experiences in a hierarchically ascending, or bottom-up, fashion. Concurrently, it tests these anticipations by generating predictions in a hierarchically descending, or top-down, manner. A core part of this process, belief updating, hinges on the upward transmission of prediction errors (defined as the discrepancy between rising sensory information and descending predictions) through the neural hierarchy. This results in a revision of expectations, facilitating the generation of improved predictions that are then propagated down the hierarchy (Siman-Tov et al., 2019). Friston's (2009) theory of active inference introduces the concept that perceptions and actions are inextricably intertwined for the express purpose of minimizing prediction error. An additional facet of active inference is the forecasting of the precision (the reciprocal of variance) of prediction errors, which are often conceptualized as computational equivalents of attention (Feldman and Friston, 2010; Brown et al., 2011).

In conclusion, the framework of Predictive Coding (PC) offers a profound and integrative perspective on action, encompassing various levels of understanding from neural mechanisms to behaviour and social interaction. PC highlights the brain's inherent ability to predict and minimize prediction errors, thereby efficiently allocating cognitive and motor resources. This framework has found support in extensive research, emphasizing the close connections between brain regions responsible for actions and those involved in prediction and expectation. Karl Friston's work, in particular, has reframed our understanding of action by depicting it as an ongoing process aimed at minimizing prediction errors, fundamentally altering the traditional stimulus-response paradigm.

Furthermore, this predictive coding account extends seamlessly to our understanding of the actions of others, enriching our comprehension of agency, intentionality, and social interaction. It underscores the unity of predictive models that guide both our own actions and our interpretations of the actions of others. Moreover, within this framework, action prediction and prediction error assume vital roles. Action prediction involves anticipating others' actions based on contextual cues and prior knowledge, while prediction error represents the discrepancy between predicted and observed actions, serving as a critical learning signal. Neurophysiological evidence supports the concept of sensory attenuation, wherein self-generated sensations elicit reduced neurophysiological responses compared to externally generated ones. Predictive coding, with its emphasis on prediction and error minimization, has emerged as a pivotal concept in understanding neural functionality and the interplay between perception and action. It provides a unified and comprehensive perspective that encompasses not only motor output but also intention, planning, and social understanding, highlighting the brain's innate propensity for prediction and adaptation in a dynamic world.

Prediction Error (PE) will be elaborated in later sections from different angles depending on the type of information. It will help explain how top-down modulations feed into the sensorimotor loop or broader social cognition frameworks.

5.3. Agent or object? Biological cues and detection of animacy

In order to predict others' actions, the first step would be to identify agency behind the observed movements, also called “animacy” or “life motion” (for review, see Schultz & Frith, 2022; Troje & Chan, 2023). Animacy perception is a necessary component of social interaction (Rutherford and Kuhlmeier, 2013), and evidence shows that such perceptions emerge even in infancy (Adam et al., 2016; Daum et al., 2016; Elsner & Adam, 2021). Notably, animate objects capture attention (Schultz & Frith, 2022; Saito et al., 2023), and thus reflect that detecting animacy is vital in ancestral hunter-gatherer environments and is consistent with the animate-monitoring hypothesis (New et al., 2007). Previous research on factors driving the perception of animacy mainly focused on the visual properties of target stimuli (Scholl and Tremoulet, 2000, Scholl and Gao, 2013, Gao et al., 2019), such as human-like appearances (e.g., faces or bodies) and motion (e.g., body kinematics or interactive motion between geometrical shapes) (Blake and Shiffrar, 2007). Depending on our familiarity with the scenarios, we may employ different strategies to recognize "natural agents", ranging from body-related features to agency cues (Adam et al., 2016; Daum et al., 2016; Elsner & Adam, 2021), up to more complex biological motion cues extracted from body kinematics (Saygin & Stadler, 2012; Schultz and Bülthoff, 2013; Schultz & Frith, 2022). Traditionally, this has been studied using vision as the dominant sensory modality (see Schultz & Frith, 2022), although we can also recognize agency through other sensory modalities such as auditory (Bidet-Caulet et al., 2005) and tactile cues (Tame & Longo, 2023).

Yiltiz and Chen (2015) showed that task-irrelevant tactile input can resolve ambiguous perception of biological motion, and this cross-modal bias was mediated by social-cognitive factors. Bidet-Caulet and colleagues (2005) showed evidence for an auditory attentional network specialised in biological motion related to low-level sound processing. All these studies reporting agency recognition through different sensory modalities, including cross-modal learning, converge in the direction towards a multisensory processing of agency/animacy from early on.

Neural correlates from action research highlight two key areas implicated in the extraction of visual body features: STS (superior temporal sulcus) as sensitive to whole-body kinematics independently from motion perception (Puce and Perrett 2003; Blake and Shiffrar 2007), and extra-striate visual areas like FBA (fusiform body area) specifically tuned to human bodies, or EBA (extra-striate body area) involved in the representation of one's own body, contributing to the 'body schema' even when the bodies are depicted schematically in the form of Point-Light Displays (PLD) and without heads (Cross et al., 2010; Orgs et al., 2016). Downing et al. (2006) showed that incoherently ordered sequences of body postures were associated with greater activity in EBA than coherently ordered sequences, whereas the reverse effect was observed in pSTS. Yet, EBA does not appear to be specifically involved in representing or distinguishing body actions (Peelen & Downing, 2007).

STS has been reliably identified using functional magnetic resonance imaging (fMRI) to measure neural activation upon presentation of biological motion displays (e.g., Grossman et al., 2000). Moreover, brain damage in the STS region has been shown to impair the ability to recognize biological motion animations but spare other aspects of motion perception (Schenk & Zihl, 1997a, 1997b). A recent study by Karakose-Akbiyik and colleagues (2023) found distinct neural representations of actions and object events in the pSTS/TPJ (posterior superior temporal sulcus/temporoparietal junction) and superior parietal lobes. The higher sensitivity to action information in the right pSTS/TPJ aligned with previous studies linking these regions to human-specific event information such as animacy, intentionality, social interactions, and biological motion (Han et al., 2013; Gao et al., 2012; Isik et al., 2017; Grossman et al., 2005; Tarhan & Konkle, 2020). Sours and colleagues (2018) using resting state functional MRI (fMRI) in neonate, showed STS significant functional connectivity with the visual association areas, primary auditory cortex, and somatosensory association areas. This demonstrates the presence of cortical areas with converging sensory inputs, representing that the functional architecture needed for multisensory processing is already present within the first weeks of life (Sours et al., 2018). All these neuroscientific studies point towards STS as the best candidate for detection of agency, not just visually but from any sensory modality.

Box 1. STS and a third visual pathway specialized in Social Perception

Recently, Pitcher and Ungerleider (2021) proposed a third visual pathway that projects from early visual cortex, via motion-selective areas, into the superior temporal sulcus (STS). This pathway would be responsible for processing dynamic social cues and encompasses various higher socio-cognitive functions, involving tasks such as recognizing facial expressions, discriminating eye gaze, integrating audiovisual speech, and interpreting the actions and behaviours of other biological organisms (Pitcher & Ungerleider, 2021).

Dynamic social stimuli such as bodies and actions are initially processed in MT/V5 and the extrastriate body area (EBA), before information segregates between the STS and the parietofrontal regions (Urgesi et al., 2007; Kilner et al, 2007). While EBA plays a pivotal role between dorsal and ventral streams (Zimmermann et al., 2018), parietofrontal regions are part of the human action observation network (AON), which has a strong degree of overlap with the dorsal pathway. Within the AON, the STS, which is proposed to be at the heart of the new third proposed pathway, provides a centre for visual representations within a visuo-motor system engaged in processing kinematics and intentions of actions to understand other's movements (Kilner et al., 2007). This anatomical-functional overlap between the proposed third-visual pathway and multiple regions within the AON, which are dedicated to the processing of social dynamic stimuli, invites the revision and integration of action and *body processing theories. For example, a detailed reappraisal of the functional role of EBA and STS and whether their contribution changes dynamically with the demands of the social perception context, analogous to proposals for the dorsal stream made by Galletti & Fattori (2018).*

In addition to the STS' fundamental role in the third-visual stream pathway, and beyond processing faces, bodies and actions, this area also contributes to audio-visual integration, speech and language processing, touch perception and theory of mind (ToM). Yang and colleagues (2015) proposed that the pSTS lives at the intersection of social perception, action understanding, and ToM systems, providing the necessary multilevel spatiotemporal integration of social information and making it an integrative centre for social cognition. Other theories propose that social perception and cognition can be better understood by considering integrative large-scale brain networks approaches (Barret & Satpute, 2013), suggesting that the pSTS orchestrates at the junction of face perception, biological motion, and social cognition (Dasgupta et al., 2017). In addition, large-scale brain networks approaches are becoming useful to explain the consequences of dysfunction to the STS that can be observed in some psychiatric disorders such as autism (Yang et al., 2015). Research has uncovered various social processing functions in the brain (Sours et al., 2017). Studies have pinpointed activation in the STS due to five distinct social stimuli, leading to the presumption that the STS is involved in social perception. This activation is observed in relation to theory of mind (when exposed to false belief stories compared to false physical stories), voices as opposed to environmental sounds, stories in contrast to nonsense speech, moving faces as opposed to moving objects, and biological motion (Grossman & Blake (2001; Beauchamp, 2015). The STS plays a role in discerning the direction of others' gaze (known as joint attention) and is crucial for identifying the target of others' emotions

(Campbell et al., 1990). Studies looking at impaired social cognition in neurological conditions, such as high-functioning autism, the superior temporal sulcus has been identified as the key mechanism underlying these deficits in social interpretation (Pelphrey & Carter, 2008).

5.4. Sensorimotor representation: experience-based model for actions

Our internal models for sensorimotor representation of actions play a crucial role not only while learning how to navigate the world around us, but also understanding others' goal-directed actions and allowing interactions and coordination in joint-actions. This sensorimotor representations or memory of actions, traditionally called procedural memory (see Box 2), embodies the "know how" to perform specific actions, which helps to identify the underlying goals. Throughout this section, different theoretical and experimental perspectives focusing on different levels of experience with actions will be presented, with the aim to provide an integrative view on the role of internal action models for action prediction *(see Table 2 for an overview of the main body of empirical studies on sensorimotor representation).*

5.4.1. Neural basis

Action execution, action understanding and action observation. Action prediction is supported by associative learning between observed and executed actions, as indicated by neuroimaging studies showing sensorimotor involvement through the AON during action prediction (Southgate et al., 2009; Grafton, 2009; Kilner, 2010; Abreu et al., 2012; Balser et al., 2014). Engagement of the Action Observation Network (AON) during action observation facilitates effortless and efficient prediction of ongoing movements, with greater recruitment of AON regions supporting the prediction of less familiar action sequences (Cross et al., 2013). The activation of the Mirror Neuron System (MNS) has also been shown to be implicated in predicting subsequent actions or consequences of observed actions (Ramnani & Miall, 2004; Kilner et al., 2004; Iacoboni et al., 2005; Grafton & Hamilton, 2007; Cattaneo & Rizzolatti, 2009). The observation of an action that is incongruent with the one that was previously predicted would lead to higher activity in the MNS network (Kilner et al., 2007), showing higher activation in MNS regions when erroneous, impossible or extraordinary actions are observed (Costantini et al., 2005; Koelewijn et al., 2008; Stapel et al., 2010; Abreu et al., 2012). But other studies had shown higher activity in MNS regions during the observation of correct actions (Van Schie et al., 2004; Shimada & Abe, 2009, 2010; Bello et al., 2015).

Neuroimaging studies have provided insights into the neural correlates of action prediction, in line with previous studies looking at AON and MNS. In an EEG study, de Bruijn et al. (2007) brought evidence for a neural correlate when observing execution errors in everyday actions. They found faster reaction times for the categorization of execution errors, matching with a significant increment on P300 amplitude, which suggest “a more general monitoring process that signals that the occurrence of unexpected events is involved in the detection of execution errors” (de Bruijn et al., 2007). According to this, a recent TMS study from de Beukelaar et al. (2015) supported with evidence the role of the primary motor cortex (M1) and the ventral premotor cortex (vPM) in the anticipatory activity of action observation, showing a facilitatory effect of vPM over M1 in grasp predictions during action observation. Koelewijn et al. (2007) studied via MEG possible contributions of the motor system for evaluating the correctness of others' actions and demonstrate that not only is cortical motor activity modulated by action observation, but the modulation increases when the observed action is erroneous. This suggests that the motor system is engaged in evaluating the correctness of the actions of others.

Jacquet and colleagues (2016) investigated motor resonance during an action prediction task, employing transcranial magnetic stimulation (TMS) to measure corticospinal excitability (CSE). Their study revealed that experimentally induced changes in observers' prior expectations influence CSE, particularly under conditions of perceptual uncertainty. Specifically, they found that these expectations are updated based on both biomechanical and probabilistic prior information. Biomechanical information pertains to the physical effort and efficiency associated with an action, while probabilistic information involves the likelihood of behaviours derived from statistical patterns and past experiences. The degree of CSE modulation across participants corresponded with the extent of change in their prior expectations. These findings provide the first evidence that motor resonance mechanisms adapt to fluctuations in observers' expectations about others' intentions. The researchers suggest that this adaptive adjustment may function as a regulatory control mechanism, similar to those involved in action selection. This mechanism potentially helps balance and arbitrate between biomechanical and probabilistic information, optimizing the prediction of others' intentions. This balancing act is crucial when the two types of information conflict, ensuring more accurate and contextually appropriate predictions.

Table 2. Table of empirical studies used as main body of research reviewed on Action Prediction. The table classifies studies by task or paradigm, providing reference to paper, details on task/paradigm and main results.

Task/Paradigm	Paper	Details on Task/Paradigm	Results
Behavioural Developmental			
gaze	Adam et al., 2016	goal-directed gaze shifts were recorded as they observed 14 trials in which either a human hand or a mechanical claw reached for a small goal area (low-saliency goal) or a large goal area (high-saliency goal).	Only infants who had observed the human hand reaching for a high-saliency goal fixated the goal object ahead of time, and they rapidly learned to predict the action goal across trials.
	Daum et al., 2016	Twelve-month-old infants were presented with reach-and-transport actions performed by a human agent. Movement distance, duration, and – resulting from the two – movement velocity was systematically varied. Action prediction was measured via the latency of gaze arrival at target in relation to agent's hand	showing that infants are not only sensitive to differences in distances, durations, and velocities at early age but that these factors have a strong impact on the prediction of the goal of observed actions
	Kanakogi & Itakura, 2011	Time of the gaze arrival at the goal relative to the arrival of the observed agents' actions as an index of action prediction ability. Comparing gazing and grasping responses to interesting objects in 4- to 10-month-old infants and adults.	The onset of infants' ability to predict the goal of others' action was found to be synchronized with the onset of their own ability to perform that action.
gaze/spatiotemporal judgement task	Stapel et al., 2016	Spatial occlusion paradigm and eye-tracker. Short movies displaying either an object or an infant actor moving from left to right or right to left in the scene, Part of the stimulus was occluded by a black rectangle so that the actor's or object's movements were hidden from view for 280–720 ms.	Results show that infants who were proficient crawlers, but inexperienced walkers predicted crawling more accurately than walking, whereas age groups mastering both skills (i.e., toddlers and adults) were equally accurate in predicting walking and crawling. Regardless of experience, human movements were predicted more accurately by all age groups than non-human movement control stimuli.
Behavioural Expertise			
spatiotemporal judgement task	Aglioti et al., 2008	(Psychophysical) 3 Alternatives Forced Choice: in/out/uncertain at 10 different times during the video (from 400-1600ms, being 718ms the time when the ball leaves the player's hands)	Motor predicted the success of free shots at a basket earlier and more accurately than visual experts and novices, even before the ball was seen to leave the model's hands, suggesting they predicted by reading the body kinematics.
	Sebanz & Shiffrar, 2009	(Psychophysical) 2AFC: fake or pass, using experts and novices with videos, static images and PLD (kinematics vs posture) of basketball players	Significant interaction between expertise and condition. Experts made more accurate predictions than did novices only when dynamic movement information was available indicates that experts may have tapped into their own action repertoire

temporal occlusion paradigm	Cañal-Bruland & Schmidt, 2009	fake or shot, temporal occlusion paradigm, one frame before the ball left the hand. Handball players: field players vs. goalkeepers	only significant differences compare to novices, but none between goalkeepers and field players
	Cañal-Bruland et al., 2010	fake or shot, temporal occlusion paradigm, one frame before the ball left the hand. Handball players: field players vs. goalkeepers AND VIEW side vs. front (goalkeeper perspective)	significant differences for perspective (better front), but none between groups of experts
	Cañal-Bruland et al., 2012	strike or ball, temporal occlusion paradigm, response verbally or motorically. Baseball players, umpires, and novices	independent of response, players were significantly more accurate than novices, and tended to outperform umpires
	Cañal-Bruland et al., 2015	online occlusion with goggles at the time the ball left the hand, pairs of basketball players (performer, observer), from 5 locations (foul line and others) and 4 responses (2 in 2 out, touching or not the ring), Signal Detection Theory Paradigm	performers were significantly biased to overestimate balls in from the foul line compared with other locations
	Chen et al *	Temporal occlusion paradigm (early, mid and late ball flights), baseball batters and pitchers, advance and intermediate levels, motoric response (swinging or not), video from a batter point of view	advanced batters where significantly better (more accurate and less uncertain) than intermediate players, but no difference for the pitchers (motor experts). Greater perceptual sensitivity for advanced batters compared to the other three groups for early occlusions
	Jackson et al., 2006	temporal occlusion paradigm, 5 conditions from 600ms before to 120 ms after the foot that initiates direction change touches the ground, 28 skilled and novices' rugby players	Novices were found to be susceptible to deceptive movement whereas skilled participants were not; however, both skilled and novice participants were more confident on trials containing deceptive movement. The data suggest that the skill-level difference in sensitivity to advance visual information extends to deceptive information
	Mori & Shimada., 2013	RT's and Temporal-occlusion paradigm with eye movement recordings, using video clips and PLD of rugby players. Task: determine the direction of the movement.	Expert players anticipated better deceptive actions but worse non-deceptive (expectations in rugby players), compared to novices. Experts responded faster in all conditions.
Behavioural motion/agent familiarity/intentions			
temporal occlusion paradigm	Springer et al., 2011	Occlusion paradigm and PLD with partial, full or no overlap between observed and executed actions. VVII: side of motor execution, movement pattern, occluder, time, the pose time. Exp 1 front view and Exp 2 back view instructions for the same PLD stimuli.	RM ANOVA from the error rates reveal significant interaction occluder time x pose time, main effect of pose time (low error rates for longer pose times). 4way ANOVA: significant interaction between 4VV, but no occluder x pose x side, neither occluder x pose x pattern. 3way RM ANOVA significant interaction occluder time x pose time x representational overlap. No effect of front or back view. Main effect of representational overlap, showing less errors when executed and observed movement fully corresponded, compared with partial overlap and no overlap.

spatiotemporal judgement task	Brattan et al., 2015	spatiotemporal judgement task, 1000ms of an action sequence video. The action video was then replaced by the grey-scale occlusion image, presented for a fixed duration of 500ms. The occluder was then followed by the continuation of the action video until the end of the sequence. The point from which the action sequence continued post-occlusion was manipulated in 13 time-steps	confirm that we draw on our motor experience for the accurate simulation and prediction of action, results suggest that such experience facilitates more accurate state estimation for actions perceived in the 1PP which map more closely onto visual input of self-generated action. More forward prediction error is retained for 3PP viewed actions, which may however have the benefit of compensating for the uncertainty involved in interacting with others.
spatiotemporal judgement task	Hudson et al., 2016	(Spatial recognition paradigm) Action depicted hands reach for or withdraw from objects. After the sequence, participants have to judge whether a probe stimulus corresponded with the hands' final position. Participants first see static frame of the object and the arm in a neutral position, and they decide the direction of the movement depending on the object.	Main effect of probe, i.e., more "different" responses for unpredicted displacements than predicted displacements. Participants were more accurate judging the misallocation of the unpredicted trajectory, compared with the misallocation of the predicted trajectory.
spatiotemporal judgement task	Saygin & Stadler, 2012	Temporal occlusion paradigm (early, late or in-time), using videos of familiar actions by human, android and robot (appearance vs. kinematics). Task: determine time coherence.	Task performance interact with observed agent. For early continuations: accuracy was highest for human and lowest for robot. For later ones the pattern was reversed.
Spatiotemporal judgement task/Temporal alignment paradigm	Colling et al., 2016	Temporal alignment paradigm (press the button in-time with a change in the movement). Two groups, the "experienced" one was trained in the motor dimension of the experimental actions. They tested action and non-action movements.	Greater accuracy for predictions about action stimuli (mannequin with human kinematics) relative to non-action (point in motion) condition in the experienced group.
temporal occlusion paradigm	Mulligan et al., 2016	A motor-visual group physically practiced throwing darts and a perceptual training group learned to associate dart throw actions (occluded video clips) with landing outcomes. A final control group did not practice. Accuracy was assessed on related prediction tests before and after practice (involving temporally occluded video clips). These tests were performed while additionally performing simple, action-incongruent secondary motor tasks	Both trained groups improved their prediction accuracy after training, only the motor-visual group showed interference associated with the right-arm secondary motor task after practice, effects were evidenced regardless of whether predictions were made in response to video stimuli or static clips

EEG expertise/motion/agent familiarity/observation/intentions

observation/ MEP	Panasiti et al., 2016	Recording electro-encephalographic (EEG) activity in expert pianists, non-pianist musicians and musically naïve individuals while they observed correct or incorrect mute piano sequences (2sec movies of different scales). Incorrect sequences were made with the model pressing a given key with a finger different from the one expected in a standard execution of that scale	Superior error detection in pianists was paralleled by a larger Pe, hinting at the selective activation of the parietal error-monitoring system in visuo-motor experts. Moreover, only in pianists did action observation induce left lateralized mu suppression in the 10–12 Hz band, reflecting somatotopic sensorimotor simulation. A mediation analysis showed that mu suppression and performance (indexed by d') were mediated by Pe

			amplitude, indicating that the higher the simulation, the higher the sensitivity to errors for large Pe amplitude. This study shows that specific electrocortical indices link motor simulation and detection of errors in the actions of others.
spatiotemporal judgement task	Amoruso et al 2014	2AFC: correct/incorrect, ERP's time-locked to the error onset at 200ms before the end of the video. VVII. Within-factor: Condition 5 (lower body ending): Congruent, Incog: Male/Fem: Gross/Subtle. Between-factor Group 3: experts, beginners, naives. VVDD. Behaviour (accuracy), EEG: anticipatory, P300, N400, SW	behavioural: effects group, condition, and interaction (expertise modulate error detection). EEG- Anticipatory: group, condition, and interaction within groups. P300: group, among each group there are effects of condition. N400: condition, interaction within group. SW: condition, interaction within group
spatiotemporal judgement task	De Bruijn et al., 2007	Sequence of 6 pictures, the last one is correct or incorrect execution, 2s delayed response. ERP time-onset on the last picture.	RT faster for categorization of execution errors, sig bigger parietal P300 for observing execution errors, ERN no difference in amplitude (suggest the observed actions where no relevant for their own performance)

fMRI / MEG aging, motion/agent familiarity/observation/intentions

temporal occlusion paradigm	Diersch et al., 2016	using fMRI, brain activity was measured while participants observed partly occluded actions and judged the temporal coherence of the action continuation that was manipulated	Lower sensitivity in action prediction resulted in activity increases in the caudate. These results imply that the neural implementation of predicting actions undergoes similar changes as the neural process of executing actions in older adults. The comparison between internal predictions and sensory input seems to become less precise with age leading to difficulties in anticipating observed actions accurately
temporal occlusion paradigm	Balser et al., 2014	Occlusion paradigm: tennis videos occluded at the ball/racket contact. 2x2: within condition action anticipation task spatial (expected flight direction of the ball) or motor (decide their own motor response), & level of expertise. 2 control conditions: observation-only (player bouncing the ball with the racket) and ball-only (ball on a sinusoidal trajectory)	Main effect of expertise in accuracy, and condition in RT. Contrast anticipation>observation tasks: higher activation in dPMC, SPL, IFG, S1, insula, thalamus, SMG, MCC, hippocampus, caudate nucleus, V1, V2, V5, and cerebellum (Crus I, II, Lobule VI, VIIIb, IX, X). Between-subjects, anticipation>observation show higher activation in experts for IFG, IPS, SPL and cerebellum, but no differences in novices. No differences between groups in the ball-only condition. Within-subject contrast motor>spatial anticipation: IFG, S1, IPL, mSPL; the opposite comparison: lateral and anterior SPL. In experts, motor>spatial activates IPL, the opposite SPL. In novices, motor>spatial activates vPMC, S1, pre-SMA and SPL. Between-subject, novices showed stronger activation in vPMC, SMA and SPL for motor>spatial compared to experts.

spatiotemporal judgement task	Koelewijn et al., 2007	Action observation	Demonstrate that not only is cortical motor activity modulated by action observation, but the modulation increases when the observed action is erroneous. This suggests that the motor system is engaged in evaluating the correctness of the actions of others.
TMS expertise motion/agent familiarity/observation/intentions			
spatiotemporal judgement task	Jacquet et al., 2016	watching movies in which the agent grasped an unfamiliar tool with two different grips (power or precision) and then manipulated it to achieve two different goals (open the box or turn on the light). Thus, among the four possible action combinations, two were biomechanically optimal and two were suboptimal. We also varied the amount of visual information provided by the movies in such a way that in one condition only the type of grip was visible (goal-hidden movies) while in another condition both the grip and the final goal were visible (goal-visible movies)	Show that prior expectations are updated on the basis of both biomechanical and probabilistic prior information and that the magnitude of the CSE modulation observed across participants is explained by the magnitude of change in their prior expectations.
spatiotemporal judgement task	Makris & Urgesi, 2015	outfield players, goalkeepers, and novices to predict the direction of the ball after perceiving the initial phases of penalty kicks that contained or not incongruent body kinematics. During the task, we applied repetitive transcranial magnetic stimulation (rTMS) over the superior temporal sulcus (STS) and the dorsal premotor cortex (PMd)	STS-rTMS disrupted performance in both experts and novices, especially in those with greater visual expertise (i.e., goalkeepers). Conversely, PMd-rTMS impaired performance only in expert players
observation/ MEP	De Beukelaar et al., 2015	Three experiments showing two types of actions (whole hand grip, WHG, or precision grip, PG) of right-hand grasping actions (reach, grasp, lift), exp 2&3 include a pre-cue for the action. 1&2exps single-pulse stim 130% over M1, 3exp M1 120% and half of the trials paired-pulse with PMv 80% 7 sec earlier. MEP were recorded for two hand muscles (FDI-hotspot, ADM)	Exp1: muscle by phase interaction, during grasp and lift ADM was facilitated more by WHG, during lift more facilitation of the FDI for PG. Exp2: the pre-cue evoked anticipatory muscle-specific facilitation, muscle by cue significant interaction, particularly for ADM. Exp3: only during grasp phase a net facilitatory effect of PMv over M1 was measured in ADM muscle.

Box 2. Procedural memory: how memory for actions can inform AP

Procedural memory refers to cognitive and sensorimotor habits and skills learned through repetition. The first convincing experimental evidence for a dissociation between declarative memory ("knowing what") and non-declarative or procedural ("knowing how") memory was from Brenda Milner (1962), by demonstrating that a severely amnesic patient (H.M.), could learn a hand–eye coordination skill (mirror drawing) in the absence of any memory of having practiced the task before. The repetition of an activity leads to the progressive consolidation of the memory trace of the involved skills, leading to their automatization. The neural correlates of procedural memory have typically focused on the basal ganglia (BG), in particular the dorsal striatum (comprising the caudate and the putamen) and the cerebellum. Patients with degeneration of the BG due to Parkinson's or Huntington's disease, can experience marked abnormalities of force/load coupling, high variability in static force and higher than needed grip forces (Prodoehl et al. 2009). The basal ganglia (BG) are involved in learning sensory-motor contingencies, organizing sequential movements, reinforcement-based learning (including emotionally competent reward-based learning), motor planning (especially precise timing), and multiple motor programs. This is key for action prediction because BG is responsible for implicit knowledge and automatic motor patterns we take for granted, but it is engrained as the basis for prediction. Recent research used fMRI activation patterns to illuminate the representation and organization of procedural knowledge, and such procedural signatures were found in frontal, parietal, motor, and cerebellar regions (Mason et al., 2020). An alternative hypothesis is that the basal ganglia are needed for the formation of experience-dependent short-term sensorimotor memories (Grafton, 2010). Object knowledge acquired in one trial could then be used in the planning of the next trial, as has been shown (Weiss et al. 2009). This short-

range adaptation might require a reward-based learning mechanism tied to the context or motor goal (for review, see Shadmehr and Krakauer 2008; Shadmehr et al. 2010).

5.4.2. Onset of AP: Developmental approach & familiarity with actions.

The onset and chronological development of action prediction skills can be examined through developmental approach, by studying the spatial and temporal dynamics of infants' gaze during the observation of goal-directed actions. By comparing predictive gaze and grasping responses from 4 months-old to adults, critical stages of development can be identified, particularly in relation to the onset of motor abilities and whether the individual's familiarity with both agent and action. Previous research has shown that from an early age, factors such as the type of agent (human or non-human), saliency of action-relevant features, distances, durations, and velocities are rapidly learned from visual exposure and strongly impact the ability to predict the goals of observed actions (Adam et al., 2016; Daum et al., 2016; Adam & Elsner, 2018). Studies have also indicated the role of sensorimotor experience in action prediction, revealing that infants' ability to predict the goals of others' actions emerges simultaneously with their own ability to perform those actions (Kanakogi & Itakura, 2011; Stapel et al., 2016).

Using time of gaze arrival at the goal relative to the arrival of the observed agents' actions as an index of action prediction ability, Kanakogi & Itakura (2011) showed the onset of infants' ability to predict the goal of others' action was synchronized with the onset of their own ability to perform that action. Supporting these findings, Stapel et al. (2015) studied to what extent motor experience of an observer determined prediction accuracy for others' actions. Results showed that infants who were proficient crawlers but inexperienced walkers, predicted crawling more accurately than walking, whereas age groups mastering both skills (i.e., toddlers and adults) were equally accurate in predicting walking and crawling. Looking at developmental studies, it is understood that in order to predict other's actions there is a critical stage, but more importantly the onset of action prediction is directly related to the infant's idiosyncratic experience not just seeing but doing those actions.

5.4.3. The role of experience in AP

The importance of previous visual and motor knowledge. From action observation studies, it has been shown how action understanding increases in experts by the "motor resonance". This statement enables the use of "high-skilled performers" as a way to dissociate between the visual and motor domains in action prediction studies.

Visuo-motor dissociation in adults: Acquired skills and expertise. As previously shown by developmental studies, action prediction in adults is also modulated by the visual and motor experience of the observer. This line of research in action prediction explores the advantage expertise on acquired skills might have when anticipating those particular actions. Studies have shown that motor experts in specific domains exhibit enhanced neural responses when anticipating highly skilled actions (De Brujin et al., 2007; Aglioti et al., 2008; Amoruso et al., 2014; Balser et al., 2014; Panasiti et al., 2016). Studies comparing perceptual and motor experts have found motor enhancement effects when predicting the movements related to their expertise (Paull & Glencross, 1997; Ranganathan & Carlton, 2007; Aglioti et al., 2008; Rammami & Miall, 2004; Urgesi et al., 2012; Verfaillie & Daems, 2002).

The well-known study from Aglioti et al. (2008) used motor and visual experts on basketball to disentangle how each domain modulated accuracy in judging the fate of free shots. They found that motor experts predicted the success of free shots at a basket earlier and more accurately than visual experts and novices, even before the ball was seen to leave the model’s hands, suggesting they predicted by reading the body kinematics. The results are usually interpreted as indicating faster or more accurate anticipation by experienced players (as motor experts) compared with experienced viewers (visual-only experts). Mori & Shimada (2013) addressed action prediction as the ability to anticipate forthcoming events from the observation of a player's action, and assessed whether it was applicable to deceptive action. Using video clips and PLD of rugby players, they reveal expert players anticipated more accurately and faster deceptive actions in all conditions. In a similar study from Jackson et al. (2006), Novices were found to be susceptible to deceptive movement whereas skilled participants were not; however, both skilled and novice participants were more confident on trials containing deceptive movement. The data suggest that the skill-level difference in sensitivity to advance visual information extends to deceptive information.

A different approach has been recently used involving sports like handball in which there are different action roles, where there are all highly skilled professionals but depending on the action is possible to separate between motor experts and perceptual experts.

Surprisingly, they did not find any significant differences between both groups of experts, but they did for the perspective point of view (Schmidt, 2009; Cañal-Bruland et al., 2010). On the other hand, it has also been reported that when expert players are allowed to respond with their skilled sports actions, they take longer to respond than novices do, to minimize prediction errors (Brault et al., 2012; Dessing & Craig, 2010).

Neural correlates of expertise.

Different studies (Abreu et al., 2012; Balser et al., 2014) have explored with fMRI techniques how expertise modulates the AON activation pattern during anticipation tasks, suggesting that the neural processing of different anticipation tasks depends on the expertise level. Balser et al. (2014) used an occlusion paradigm, in which tennis videos where occluded at the time of ball-racket contact, with two action anticipation tasks and an observation-only condition. They found higher activation in dPMC, SPL, IFG, S1, insula, thalamus, SMG, MCC, hippocampus, caudate nucleus, V1, V2, V5, and cerebellum (Crus I, II, Lobule VI, VIIIb, IX, X) when comparing anticipation>observation tasks. Between-subjects, the same contrast show higher activation in experts for IFG, IPS, SPL and cerebellum, but no differences in novices.

In a recent study from Makris & Urgesi (2015), where they use rTMS to search for the causative role of visual and motor representations of the experts' ability to predict soccer actions, they found that when the rTMS was applied to STS it disrupted performance in both experts and novices, especially in those with greater visual expertise (i.e. goalkeepers). Conversely, when applied to the PMd, it impaired performance only in expert players.

To define how the level of expertise modulates prediction and interpretation during action observation, Amoruso et al. (2014) used electroencephalography (EEG) to compare anticipatory activity, P300, N400 and Slow Wave (SW), from experts, beginners and novices tango dancers. They found that anticipatory activity from the fronto-parieto-occipital network, early discriminated between levels of motor expertise and error detection, and significantly predicted semantic integration (N400 and SW) in temporal and motor regions. In a more recent study, Panasiti et al. (2016) explore the link between sensorimotor expertise, the ability to detect another's erroneous action (indexed by positivity error, Pe) and action simulation (indexed by mu frequency suppression), recording electro-encephalographic (EEG) activity in expert pianists, non-pianist musicians and musically naïve individuals while they observed correct or incorrect mute piano sequences. Superior error detection in pianists was paralleled by a larger Pe, hinting at the selective activation of the parietal error-monitoring system in visuo-motor experts. Moreover, only in pianists, action observation induced left lateralized mu suppression in the 10–12 Hz band, reflecting somatotopic sensorimotor simulation. A mediation analysis showed that mu suppression and performance (indexed by d 0) were mediated by Pe amplitude, indicating that the higher the simulation, the higher the sensitivity to errors for large Pe amplitude. This study shows that specific electrocortical indices link motor simulation and detection of errors in the actions of others.

Motor interference/facilitation: Cognitive approach theories and behavioural studies.

Cognitive approach theories, such as the "perceptual experience hypothesis" and "motor experience hypothesis," provide insights into the influence of perceptual and motor experience on action prediction. The perceptual experience hypothesis suggests that skilled individuals are better anticipating actions because they have greater perceptual experience due to their frequent encounters with the observed actions (Abernethy & Zawi, 2007; Abernethy, Zawi, & Jackson, 2008; Williams et al., 1999). Common coding theory, supporting the motor experience hypothesis, suggests that anticipatory readiness potentials over the motor cortex can predict the onset of others' movements (Kilner et al., 2004). Motor resonance, characterized by the activation of similar motor representations in the observer when perceiving an action, has also been explored in action prediction paradigms (Springer et al., 2011; De Beukelaar et al., 2015; Jacquet et al., 2016). Motor interference, a paradigm that involves simultaneous action execution and therefore an interaction between motor and visual processes, has also been investigated in the context of action prediction.

In order to assess how real time simulation is affected by simultaneous action execution, Springer et al. (2011), used an occlusion paradigm displaying actions performed by an actor or Point Light Display (PLD) with partial, full or no overlap between observed and executed actions. They manipulated side of motor execution, movement pattern, occlude time and the pose time, as well as front (Experiment 1) or back (Experiment 2) view instructions. They manifested a main effect of representational overlap, showing less errors when executed and observed movement fully corresponded, compared with partial overlap and no overlap, suggesting that action prediction and real-time simulation can be affected by motor processes, producing interference or facilitatory effects. Mulligan et al. (2016) examined the effects of motor interference in action prediction, where a motor-visual group physically practiced throwing darts and a perceptual training group learned to associate dart throw actions (occluded video clips) with landing outcomes. Both trained groups improved their prediction accuracy after training, only the motor-visual group showed interference associated with the right-arm secondary motor task after practice (Mulligan et al., 2016). Moreover, there is evidence for a motor facilitation effect of representational overlap, showing less errors when executed and observed movement fully corresponded, compared with partial overlap and no overlap (Springer et al., 2011). Colling et al., (2016) showed how familiarity with movement kinematics is cross modal and motor experience alone can indeed enhance visual prediction.

5.4.4. Updating the sensorimotor model

Proprioception and vestibulo-ocular feedback play a crucial role in updating our sensorimotor model. Proprioception, which involves the awareness of body posture and movements through receptors in skeletal muscles, joint capsules, and the skin, contributes to coordinating and adapting movements. Constant updating of the sensorimotor model through prediction error occurs with every action performed or sport practiced.

Some EEG studies have been proposing different indexes to measure prediction error in the brain. De Bruijn et al. (2007) identified parietal P300 associated with observing execution errors during everyday actions. Amoruso et al (2014) proposed that anticipatory activity, with sources in a fronto-parieto-occipital network, early discriminated between expert and non-expert groups according to their level of expertise. Furthermore, this early activity significantly predicted subsequent semantic integration indexed by semantic responses

(N400 and SW, sourced in temporal and motor regions), which also predicted motor expertise (Amoruso et al., 2014).

Panasiti and colleagues (2016) explored the link between sensorimotor expertise, the ability to detect another’s erroneous action (indexed by positivity error, Pe) and action simulation (indexed by mu frequency suppression). The study recorded EEG activity in expert pianists, non-pianist musicians and musically naïve individuals while they observed correct or incorrect mute piano sequences. Superior error detection in pianists was paralleled by a larger Pe, hinting at the selective activation of the parietal error-monitoring system in visuo-motor experts. Moreover, only in pianists did action observation induce left lateralized mu suppression in the 10–12 Hz band, reflecting somatotopic sensorimotor simulation. A mediation analysis showed that mu suppression and performance were mediated by Pe amplitude, indicating that the higher the simulation, the higher the sensitivity to errors for large Pe amplitude. This study shows that specific electrocortical indices link motor simulation and detection of errors in the actions of others (Panasiti et al., 2016).

Another study by Orlandi and colleagues (2020) found that observation of effortful compared with effortless movements resulted in a larger P300 over frontal sites in dancers only, likely because of their visuomotor expertise with the specific steps. Moreover, an enhanced Late Positivity was also identified over posterior sites in response to effortful stimuli in both groups, possibly reflecting the processing of larger quantities of visual kinematic information. Late Positivity component showed greater engagement of frontoparietal regions in dancers, while task-related frontal and occipitotemporal visual regions were more active in controls (Orlandi et al., 2020).

As we are constantly updating our own sensorimotor model, every time we perform an action, or we practice a particular sport, prediction error is therefore tailored to our own sensorimotor experience. Cerebellum has been shown to be involved in signalling the sensory discrepancy between the predicted and actual sensory consequences of movements (Blakemore et al 2001). These findings show there is cerebellar updating during on-line control. Cerebellar involvement is likely implemented in part via thalamo-cortical pathways to motor cortex. Current hypotheses posit that the cerebellum predicts current and future states of the body using stored models of motion patterns that typically occur in specific contexts—a necessary tool given the delay with which proprioceptive data regarding the body’s state return to the central nervous system.

There is a general consensus that both cerebellum and parietal cortex play critical roles in forming internal models that involve state estimation for tasks such as pointing and grasping (Nowak et al. 2007b; Tunik et al. 2007; Andersen and Cui 2009). Single neuronal recordings of cerebellar Purkinje cells show object-specific modulation of signals appearing within the reach phase or at grasp onset in a precision reach-to-grasp task (Mason et al. 2006). Responses within motor cortex and cerebellum track monotonically with overall grip force (Keisker et al. 2009). Exposure to an object where the dynamics have been learned causes increased activity in cerebellum and motor cortex prior to movement onset, suggesting that the physical priors of the object (aka affordance) are triggering internal models stored in these two areas (Bursztyn et al. 2006). When the dynamics of an object are experimentally manipulated by making subjects balance a flexible ruler compared to a simple grip task, there is also relatively greater activity in cerebellum and motor cortex, suggesting that these areas are either implementing or representing internal dynamics (Milner et al. 2007).

5.5. The importance of context

5.5.1. Intentions & expectations: top-down modulations by context

By predicting others' actions, individuals can anticipate their needs, intentions, and responses, thereby facilitating social interactions and adaptive behaviour. Action observation, like perception in general, is inherently predictive and happens relative to top-down expectations (Bach, Nicholson & Hudson, 2014; Csibra, 2007; Kilner, 2011). Hudson and colleagues (2016) showed how expected goals can bias perceived kinematics of other’s familiar actions (i.e., affordance). Mori & Shimada (2013) found that even though experienced players anticipated better deceptive actions, they were also more susceptible towards context-driven expectations.

5.5.2. AP and Context - emotion (behavioural and neural evidence)

The socio-affective framework, which encompasses conceptual and emotional contexts related to intentions (the "why" aspect), plays a significant role on how we predict and interpret our daily actions and interactions. A recent study by Dima and colleagues (2022) used naturalistic videos of everyday actions and found that social-affective features predict action similarity judgments better than, and independently of, visual and action-related characteristics. Neural patterns indicated that behaviourally relevant features are automatically extracted by the brain in a temporal gradient from visual to action to socio-affective features (Dima et al., 2022).

Neural studies had highlighted the involvement of the posterior cerebellum in reconstructing and predicting social action sequences. (Van Overwalle et al.,2022; Siciliano et al., 2023). The role of the cerebellum as a sequential and prediction processor has been investigated in the social domain, where the construction of internal models reflects the correct implementation of the sequence of social actions (Siciliano et al., 2023), thus allowing people to predict one's own and others' behaviors and reactions, and to adjust unexpected events when violations from predicted scenarios are met to finally adjust social interaction accordingly by modulating cerebral cortex activity (Clausi et al., 2019; Heleven et al., 2019; van Overwalle et al., 2019a,b, 2022). Meanwhile, the proposed third visual pathway computes higher socio-cognitive functions based on dynamic social cues (see Box. 1). These functions include recognizing facial expressions, discerning eye gaze, integrating audiovisual speech, and interpreting the actions and behaviours of other organisms (Pitcher & Ungerleider, 2021). The superior temporal sulcus (STS) is suggested to be central to this third pathway and considered within AON and MN studies as a hub for visual representations within the visuo-motor system involved in processing the kinematics and intentions of actions to understand others' movements (Kilner et al., 2007, Yang et al., 2015).

5.5.3. AP and Context - perspective

Social function: First- vs third-person perspective. The perspective from which actions are perceived, whether first-person or third person, is closely linked to social function. It can be noted that good social function relies on accurate action prediction. We draw on our motor experience for the accurate simulation and prediction of others' actions. A study by Brattan et al. (2015) suggested that such experience facilitates more accurate state estimation for actions perceived in the 1PP which map more closely onto visual input of self-generated action. More forward prediction error was retained for 3PP viewed actions, which may however have the benefit of compensating for the uncertainty involved in interacting with others (Brattan et al., 2015).

Additionally, there may be variations in action prediction abilities among individuals with different disorders, such as autism spectrum disorder (ASD) (see Box 3). Balsters and colleagues (2016) investigated third-person perspective in ASD group, showing reduced awareness of prediction accompanied by reduced prediction-error-related frontal BOLD activity when predicted reward is experienced by others. The examination of such differences could warrant a dedicated section or inclusion within the context of social function. Amoruso et al. (2019) investigated the modulation of action prediction by contextual priors in children with autism and found no significant effect.

Box 3. ASD & AP studies

Bolis and Schilbach (2018) define autism as "a developmental condition, characterized by difficulties of social interaction and communication, as well as restricted interests and *repetitive behaviours". Schuwerk and colleagues (2016) suggested that impaired action prediction is at the core of social interaction deficits in autism spectrum disorder (ASD). However, the social difficulties experienced on ASD might be more related to the complexity of the situation, with social interactions being of highest complexity and related to dynamic cues. It has been proposed that sensory and social processes in autism are more tightly interconnected than traditionally thought, meaning a mismatch in sensory abilities can lead to difficulties on a social level (Bolis and Schilbach, 2018).*

There are several studies pointing towards social sensory cues, which could lead to cascade effects (Fulceri et al., 2018 --- ASD group showed impaired joint action when they had to rely on partner's kinematic cues in the absence of goal cues.; Barzy, Mahsa; Black, Jo; Williams, David; Ferguson, Heather J., 2019) . If there is an attentional deficit for agency, this would impact the saliency of dynamic social cues, potentially disrupting sensorimotor and social learning. This has also been reported as ASD population using a more proprioceptive-driven approach, showing stronger than normal association between self-generated motor commands and proprioceptive feedback, meaning they rely less on visual cues (Haswell et al., 2009). Moreover, the same study found that the greater the proprioceptive-driven generalization in our task, the greater the impairments in general motor function, social interaction, and imitation/praxis. ...within predictive processing, attention is understood as a mechanism for precision modulation, with top-down attention corresponding to prospective precision modulation based on inferred context (Feldman & Friston, 2010). Balsters et al., (2016) ---- ASD group demonstrates reduced awareness of prediction accompanied by reduced prediction-error-related frontal BOLD activity when predicted reward is experienced by others.

Cannon et al (2021), Four studies found predictive differences in ASD that were specific to social stimuli and did not appear for stimuli outside the social domain: reduced neural habituation to repeated faces but not repeated shapes (Ewbank et al., 2017), reduced EEG evidence of prediction for action sounds (Grisoni et al., 2019), reduced accuracy of perceptual inferences when they were based on social priors (von Der Lühe et al., 2016), and reduced fMRI activity related to social (but not non-social) reward prediction error differences (Kinard et al., 2020). Although, Amoruso et al. (2019) investigated the modulation of action prediction by contextual priors in children with autism and found no significant effect. Other studies, when looked directly at attended vs. unattended predictions (with the target of attention determined by task instructions), found neural responses were stronger in the ASD group for attended prediction errors but weaker for unattended prediction errors (Gonzalez-Gadea et al., 2015; Westerfield et al., 2015). This suggests that attention plays a powerful mediating role in predictive processes or prediction-modulated responses in ASD (Cannon et al., 2021).

5.5.4. Updating contextual models

Neurophysiological evidence suggests the presence of prediction error in the superior temporal sulcus (STS), a visual area that is selective for biological motion. Recurrent feedback from fronto-parietal mirror areas to the STS has been observed in functional magnetic resonance imaging (fMRI) studies, indicating top-down input from the mirror system to the STS (Iacoboni et al., 2001, 2005). The cerebellum is proposed to assist in learning and understanding social action sequences, thereby facilitating social cognition, and enabling optimal predictions about imminent or future social interactions and cooperation (Overwalle et al., 2020).

Specific environmental scenarios are often indicative of which actions are likely to occur in them, thus constraining predictive processing. Amoruso et al. (2019) suggest that top-down predictions are not solely based on past visual or motor experiences (Aglioti et al., 2008; Amoruso et al., 2014; Calvo-Merino et al., 2005), but also on prior knowledge about the context in which actions are typically observed (Amoruso et al., 2016; Wurm & Schubotz, 2012). Furthermore, contextual information plays a crucial role in aiding action recognition, particularly when perceptual information is limited (Wurm & Schubotz, 2016).

Sitnikova and colleagues (2008) highlighted ERP findings in a variety of contextual congruency paradigms, suggesting people map visual images on graded semantic representations within approximately 400 ms after stimulus onset. Again, the amplitude of this N400 effect was proportional to the relationship strength between the prime and target pictures (e.g., McPherson & Holcomb, 1999). Both linguistic and picture stimuli evoked the N400 component, the distribution of this waveform across the surface of the scalp was different for pictures than for words, suggesting distinct underlying neuronal sources. Whereas the N400 evoked by verbal stimuli is characterized by a parieto-occipital scalp topography (Friederici, Pfeifer, & Hahne, 1993; Hagoort & Brown, 2000b; Holcomb et al., 1999; Kutas & Van Petten, 1994; van Berkum et al., 1999), the negativities elicited by pictures are typically distributed over more anterior electrode sites (Barrett & Rugg, 1990;

Hamm, Johnson, & Kirk, 2002; Holcomb & McPherson, 1994; McPherson & Holcomb, 1999; West & Holcomb, 2002).

5.6. Conclusion

Action prediction is a complex cognitive process that involves detecting agents, utilizing internal representations, and inferring intentions from contextual cues in a dynamic fashion. It draws upon sensorimotor experiences, social knowledge, and contextual expectations. The ability to predict others' actions contributes to our understanding of social cognition and empathy, facilitating social interactions and adaptive behaviour. Action prediction initiates with the perception of sensory cues that indicate agency or animacy, such as faces and biological motion. These cues guide the subsequent application of internal sensorimotor models, which have been fine-tuned through individual experiences and self-awareness. These models not only help anticipate physical movements but also serve as a basis for inferring the underlying intentions and goals. Contextual information—including socio-affective factors and cultural background—further refines these predictions, adding an additional layer of depth to our understanding of human actions and interactions. Prediction Error serves as an essential mechanism in this process, offering insights into the flexibility and adaptability of both sensorimotor predictions and broader social cognitive frameworks. Through the lens of action prediction, we gain a comprehensive understanding of the intricate tapestry of factors that contribute to human social interaction and cognition.

In summary, the multidimensional nature of action prediction necessitates an integrative perspective that unifies attention, perception, sensorimotor functions, implicit memory, and context such as socio-affective elements and cultural norms. This integrative view is not merely an academic exercise but a critical framework for a more complete and nuanced understanding of complex human behaviours. By compartmentalizing these domains, we risk missing the synergistic effects that these elements have on our ability to predict actions and intentions. Moreover, the concept of Prediction Error underscores the adaptability of this integrated system, serving as a key mechanism that recalibrates and refines our anticipatory models based on the different types of information and top-down influences. This integrative approach provides a comprehensive lens through which we can explore not only the mechanics of action prediction but also its broader implications for social cognition, empathy, and theory of mind. Future research should continue to explore the neural mechanisms underlying action prediction and its implications for various domains, such as social neuroscience, developmental psychology, and clinical populations. Enhanced understanding of action prediction will contribute to a deeper comprehension of human behaviour and social interactions.

Chapter 6: Action prediction and fluency: a behavioural study

6.1. Introduction

Fluency has been established by researchers of different fields as a transversal quality with significant impact on perception, motor execution, language comprehension, memory, aesthetic appreciation and even decision making. In perception, fluency has been long reported to play a role in the ease of processing visual stimuli, such as shapes and patterns (Reber, Winkielman, & Schwarz, 1998). In motor execution, fluency is considered a measure of mastery in tasks, including sports and musical performance (Keller, 2012; Palmer, 1997). In language comprehension, fluency is considered essential for reading and linguistic processing (Frazier & Rayner, 1990; Segalowitz & Hulstijn, 2005). In memory research, fluency impacts both the encoding and retrieval of information (Jacoby & Dallas, 1981; Whittlesea, Jacoby, & Girard, 1990). Aesthetic appreciation is also affected by perceptual fluency, with smoother, more symmetric objects generally preferred (Reber, Schwarz, & Winkielman, 2004). Even in decision making studies, fluency has been shown to influence judgments and choices, both trivial and consequential (Alter & Oppenheimer, 2009; Schwartz, 2004). By incorporating such broad impacts, fluency emerges as a quality with transversal implications, influencing various aspects of human cognition and behaviour.

Our everyday lives consist of a continual stream of sensory information that informs our choices, followed by the planning and carrying out of corresponding actions. It is therefore logical to explore how perception affects action. While it might seem less intuitive to look at this relationship in reverse, there is mounting evidence that action and perception are mutually dependent, and fluency might be an intersectional feature. Some studies had shown that the way we prepare for specific actions can alter our perception of the relevant attributes of objects (Bekkering & Neggers, 2002; Craighero et al., 1999; Gutteling et al., 2013). Other studies had also confirmed that a person's ability to perform particular actions can shape their perception around it (see Witt, 2011 for a review). But more importantly, there is research evidence showing that the biomechanical exertion tied to the outcomes of actions influences perceptual judgments (Orgs et al., 2016). For example, perceptual decisions were less prone to change when altering a pre-planned action that would result in a greater physical effort (Burk et al., 2014; Moher & Song, 2014). Moreover, Hagura et al. (2017) found that the effort associated with an action could not only sway perceptual decision-making thresholds but also the perception itself: responses from participants tended to veer away from choices that required substantial physical effort in a motor task, suggesting that the ease or difficulty of performing an action has repercussions for perceptual processes.

Perceptual fluency and motor fluency are distinct but interconnected aspects of human cognition and behaviour (Reber, Winkielman, & Schwarz, 1998). Both refer to the ease and efficiency with which certain types of tasks are executed—perceptual fluency deals with the processing of sensory information while motor fluency relates to the performance of motor tasks. *Perceptual fluency* refers to the ease with which information is processed. In cognitive psychology, this concept is used to describe how quickly and effortlessly a stimulus can be understood (Reber, Schwarz, & Winkielman, 2004). When something is perceptually fluent, the mind requires less cognitive "effort" to interpret it (Alter & Oppenheimer, 2009). This can apply to a wide range of stimuli, including but not limited to visual images, sounds, or even conceptual ideas. Perceptual fluency is often associated with a range of psychological effects. For instance, studies have shown that perceptually fluent stimuli are more likely to be judged as more beautiful, more truthful (Winkielman & Cacioppo, 2001), or more easily remembered compared to less fluent stimuli (Rhodes & Castel, 2008). This phenomenon is closely related to the idea of cognitive ease, which suggests that the mind prefers situations or stimuli that are easier to process.

Motor fluency refers to the ease and efficiency with which motor tasks are executed (Wulf, Shea, & Lewthwaite, 2010). In essence, it describes how "smoothly" a person can perform a physical action, often as a result of practice, familiarity, or natural ability (Wulf & Prinz, 2001). Motor fluency is relevant in various domains such as sports, music performance, and even everyday activities like walking or handwriting. High motor fluency generally indicates a level of mastery or expertise in a particular skill (Ericsson, Krampe, & Tesch-Römer, 1993). For instance, an expert pianist would play a complex piece with high motor fluency, making even the most intricate passages appear effortless. Similarly, an athlete may execute a series of movements with such fluidity that it seems almost automatic. In summary, while perceptual and motor fluency can be examined as separate phenomena, they are often closely linked in real-world tasks and activities. Effective performance in many domains, from sports and music to daily activities like driving, often requires both perceptual and motor systems to function well and in coordination with each other (Proteau, 1992).

Biological motion (BM), the pattern of movement unique to living organisms, serves as a crucial intersection between perceptual and motor fluency (Johansson, 1973). On the perceptual side, humans are adept at quickly and accurately recognizing biological motion, whether it's identifying the gait of a human walker from minimal visual cues (Blake &

Shiffrar, 2007) or interpreting subtle facial expressions (Ekman & Friesen, 1971). This ease of recognition is an aspect of perceptual fluency that extends to predicting future actions, such as the trajectory of a ball based on a player's kicking motion. On the motor side, the ability to execute complex movements, like dancing or athletic manoeuvres, requires high levels of motor fluency (Sevdalis & Keller, 2011). This motor fluency can be honed through a feedback loop that involves watching biological motion—either one's own or others'—and then practicing to improve. Interestingly, observing biological motion can activate neural pathways similar to those used when executing the same movements (Rizzolatti & Craighero, 2004), a phenomenon known as motor resonance (Calvo-Merino et al., 2005; 2006), which may facilitate the mutual enhancement of both perceptual and motor fluency. In summary, biological motion is intimately tied to both perceptual and motor fluency. Our ability to understand and interpret the movements of other living beings relies on perceptual fluency, a construct supported by research on visual perception and action recognition (Reber, Winkielman, & Schwarz, 1998; Blake & Shiffrar, 2007). On the other hand, our capacity to execute these movements ourselves depends on motor fluency, which is studied in the context of motor control and skill acquisition (Keller, 2012; Sevdalis & Keller, 2011). Both forms of fluency often work in tandem, influencing and enhancing each other in a variety of contexts. This interconnectedness is evident in studies that explore the relationship between observation and execution of action, such as research on mirror neurons and motor resonance (Rizzolatti & Craighero, 2004; Calvo-Merino et al., 2005; 2006).

In some cases, the fluency of biological motion itself serves as a form of communication. For example, the fluidity and grace of a dancer's movements can convey emotions or narratives.

Orgs et al. (2016) used an implicit measure of movement fluency to show that fluent apparent motion sequences produced subjectively longer durations than non-fluent sequences of the same objective duration. Moreover, fluent sequences were associated with increased connectivity between M1/SMA and right fusiform body area (FBA), showing that perceptual reconstruction of fluent movement involves cooperative recruitment of motor areas, beyond the ones traditionally associated with visual body processing (Orgs et al., 2016). Another study by Guo & Song (2019) demonstrated that as actions became more fluent perceptual sensitivity also improved. They found that action training prior to discrimination enhanced subsequent perceptual sensitivity, supporting the notion of a reciprocal relation between perception and action (Guo & Song, 2019).

The focus of the current study will be at the intersection between perceptual and motor fluency. Using an action prediction task, such as a temporal occlusion paradigm, and biological movement as the stimuli of interest, the role of perceptual fluency will be assessed through participants' ability to discriminate fluent movement. In this sense, action prediction is built on our internal sensorimotor representations of actions, which are constructed based on our own experiences of perceiving or performing them and can therefore be enhanced by both visual and motor familiarity with the specific movements. Using human dance movement as the stimuli of interest, allows us to explore a complex whole-body movement which participants might not have motor experience with, but still provides ecological validity.

This study was aimed as the first of a series of experiments, where a temporal occlusion paradigm could be tested at the right level of difficulty and make sure the fluent and non-fluent conditions could be discriminated against, showing enough range of variability in terms of sensitivity/discrimination across participants without any ceiling or floor effects. It was predicted that the fluent condition (congruent) would be easier to identify as the “correct” continuation of the movement compared to the non-fluent condition (incongruent) - in other words, P [Z(hits)] > P [Z(false alarms)], resulting in a positive measure of sensitivity (d’). Therefore, the directional experimental hypothesis was that participants would be more accurate and quicker in their responses to the congruent condition.

6.2. Methods

6.2.1. Participants

Twenty-nine participants (M=29, SD=9.27; 14 women), all right-handed, took part in the experiment. All participants had normal or corrected to normal vision, and reported no cognitive, attentional, or neurological impairment. Written and informed consent was obtained from all participants and the study was approved by the Psychology Research Ethics Committee at City, University London (ETH1819-1187), following guidelines and procedures established in the Declaration of Helsinki.

6.2.2. Materials

Twenty-four whole-body dance videos with scrambled faces were used in the experiment, half portrayed ballet movements and the other half capoeira moves (stimuli used with permission from Calvo-Merino et al., 2005). The videos were 3 seconds long, recorded in a neutral background, matching kinematic level in both styles according to speed, part of the body involved, body location in space and movement direction. For the occlusion paradigm, videos were decomposed in 75 frames (25 frames per second). The experimental task showed only the first half of the videos and created both congruent and incongruent conditions using a time-coherent or 400ms-delayed frame respectively as a target (see below full description on design section).

6.2.3. Design

This study used a within-participants experimental design. A temporal occlusion paradigm was implemented to create a congruent or incongruent continuation of the movement *(see Figure 15*). Participants saw the first 1500ms of the video, followed by a 350ms occlusion (black screen), and a target image that belongs within the same video. The target frame would represent a congruent continuation of the movement when the frame did follow temporal coherence from video's onset, corresponding with 1850ms from video onset (correctly synchronised timing, corresponding with frame 47). Incongruent trials would display a target frame corresponding to 2250ms since video onset (representing a 400ms delay, corresponding to frame 57). Both possible target frames correspond to the same video, portraying a continuation of the movement displayed in the video, but only one frame represents a fluent continuation and therefore congruent with the timing from onset. With a total of 24 videos (12 ballet, 12 capoeira), and two equally possible conditions per video (50% congruent / 50% incongruent), there were 48 different trials in total. Trials were randomly presented over 10 blocks, adding up to a total of 480 trials. The experiment measured participants' accuracy and reaction times when responding a 2AFC (yes/no) to the statement: "Indicate if that image is the correct continuation of the movement".

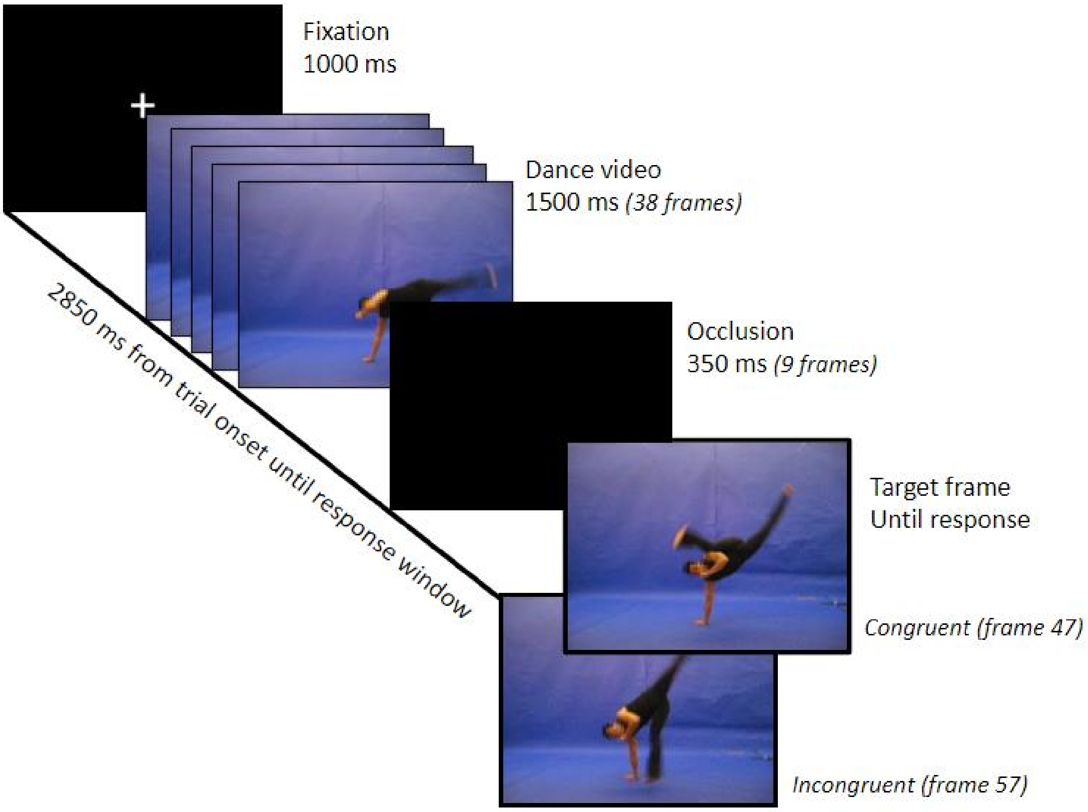


Figure 15. The experimental task used a temporal occlusion paradigm, showing 1500ms videos of different ballet and capoeira movements, followed by a 350ms occlusion and a target image from the same video which could be time-congruent with the video's onset or 400ms delayed. The participant had to respond yes or no to the question: "Indicate if that image is the correct continuation of the movement".

6.2.4. Procedure

Participants were brief and gave informed consent prior to the experiment. The task was to watch 24 different videos portraying dance movements followed by a short occlusion and a target image and respond whether the target image was a correct continuation of the video. Each trial started with a 1000 ms fixation cross, then 1500ms video portraying a ballet or capoeira movement followed by a 350ms occlusion and a target static image until response *(see figure 15).* Participants were instructed to respond yes or no to the statement: "Indicate if that image is the correct continuation of the movement". The response was recorded by participants' pressing on the keyboard with their right hand.

6.3. Results

Behavioural data was collected, using a temporal occlusion paradigm where participants had to respond as a 2AFC (yes/no) to whether the target image seemed to be the correct continuation from the video's movement. The experimental manipulation consisted in using a target frame that would be congruent (fluent) or incongruent (non-fluent) with the timing since video's onset. Two different dance styles were used, displaying ballet or capoeira moves. Accuracy (ACC) and reaction times (RTs) were measured for congruent and incongruent trials in both ballet and capoeira movements. It was hypothesized that participants would discriminate faster, and more accurately congruent trials compared to incongruent ones.

According to the experimental design, 2x2 RM ANOVA were used to analyse accuracy and reaction times separately, looking for evidence pointing towards the experimental hypothesis, but also a lack of effect on dance style. Effect sizes were also calculated and reported as Partial Eta Squared (PES). As a separate measure to evaluate the effectiveness of discrimination between conditions d' (d prime) and criterion c were calculated from the behavioural responses (see Chapter 1.4 Measurements & Methods).

Table 3. Descriptive statistics showing mean (M), standard deviation (SD) and standard error (SE) of accuracy (ACC) and reaction time (RT) measures, for both congruent and incongruent conditions & for each dance style separately (ballet/ capoeira).

			M	SD
ACC	Ballet	Congruent	0.65	0.13
		Incongruent	0.55	0.20
	Capoeira	Congruent	0.68	0.19
		Incongruent	0.54	0.19
RT	Ballet	Congruent	997.09	264.23
		Incongruent	1071.63	312.75
	Capoeira	Congruent	965.16	264.77
		Incongruent	1077.06	360.41

Sample means suggested participants discriminated better and faster congruent trials regardless of the dance style *(see table 3)*. Inferential tests showed those differences were only statistically significant for reaction times but not quite for accuracy measures. Statistical analysis using 2x2 RM ANOVA with factors dance style (ballet/capoeira) and fluency (congruent/incongruent target frame) was performed separately for accuracy (ACC) and reaction times (RTs) measurements. Results revealed a significant main effect of fluency on RTs (F(1,27)=4.549, p=0.042, PES = .144), and an almost-significant main effect of fluency on ACC (F(1,27)=4.125, p=0.050, PES = .135). There were no significant main effects of dance style or interactions (*see figure 16*).

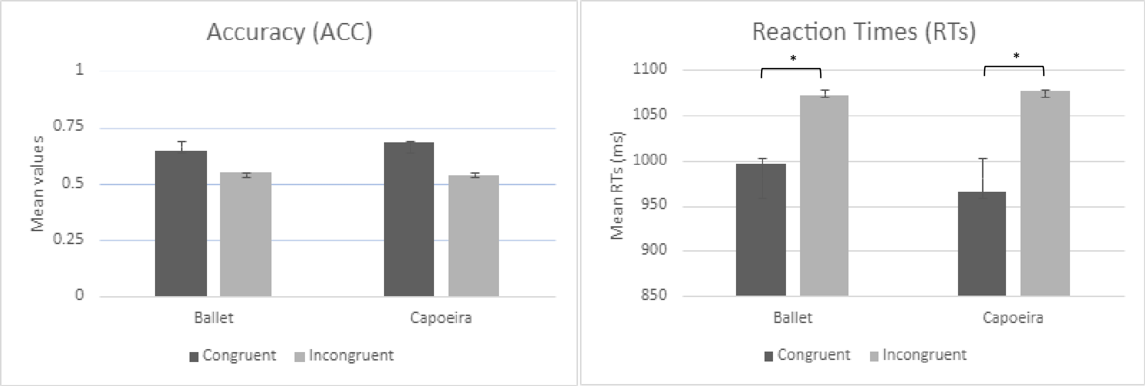


Figure 16. Graphs showing mean values for accuracy (left) and reaction times (right) for each experimental condition. Error bars represent standard error. Significant differences are represented by (*).

As a separate measure to evaluate the effectiveness of discrimination between conditions, sensitivity (d’) and response bias (criterion c) were calculated from the behavioural responses separately for each participant *(see figure 17).* Average across participants were calculated, showing appropriate levels of sensitivity (d’ = 1.27) meaning participants were in fact able to discriminate the fluent condition, and response bias (c = 0.26) meaning participants’ responses where not highly biased.

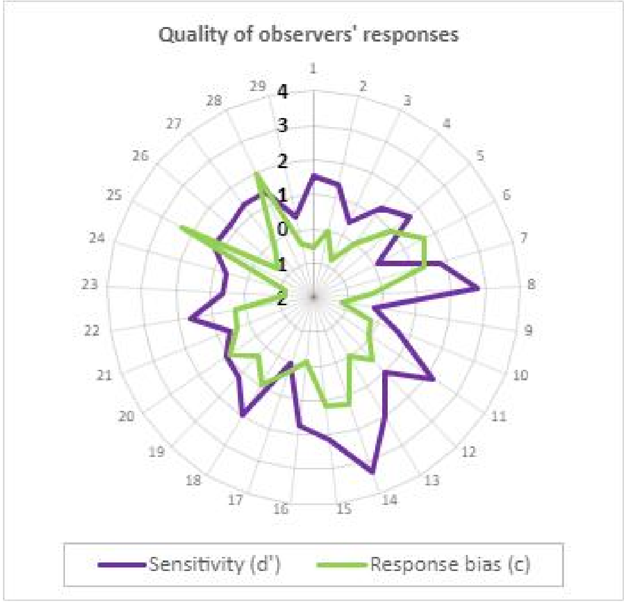


Figure 17. Graph shows values on sensitivity (d’) and response bias (c), as standardized measures to assess quality of responses per participant.

6.4. Discussion

These results support the experimental hypothesis that participants were significantly faster under the congruent condition, indicating that fluent continuations of the movement were easier to perceive, regardless of the dance style. Participants were able to discriminate quite well between congruent and incongruent trials, as shown by the sensitivity/discriminability index (d’), even though the difference in accuracy was found at threshold (p=.050) to be considered significant. There were no differences between the different dance styles, as predicted by the sample of participants not including dance experts.

The current study's findings lend significant empirical support to the interrelated nature of perceptual and motor fluency, specifically within the context of biological motion. In line with our directional experimental hypothesis, participants demonstrated notably faster reaction times and greater sensitivity (d') in responding to fluent (congruent) biological movements as opposed to non-fluent (incongruent) ones. Interestingly, this enhanced performance did not extend to accuracy, which did not show a significant difference between conditions. The lack of significant differences in accuracy warrants particular discussion and replication studies. While sensitivity and reaction time are pivotal indicators of perceptual fluency, the absence of an accuracy effect suggests that other cognitive factors may come into play in the discrimination task. It is possible that participants were equally "accurate" in discriminating both fluent and non-fluent movements but did so more "efficiently" and with greater "confidence" in the fluent condition, as indicated by the quicker reaction times and higher d' values.

These results are in harmony with existing literature positing the interconnectedness of perceptual and motor systems (Orgs et al., 2016; Guo & Song, 2019). Our data reinforce the idea that internal sensorimotor representations, built upon both visual and motor experiences, play a role in influencing perceptual tasks. This could be a manifestation of the concept of motor resonance (Calvo-Merino et al., 2005; 2006; Rizzolatti & Craighero, 2004), where the observation of biological motion activates neural pathways also used in the execution of those movements. The findings have important implications for real-world applications, from athletic training to clinical rehabilitation. The fact that reaction times and sensitivity were influenced by fluency could be instrumental in designing training regimens that aim to improve both the perception and execution of complex biological movements, such as dance, which served as the stimuli for our study.

Nonetheless, some limitations must be acknowledged. The study was constrained to a specific form of biological motion—dance— which is particularly complex as a stimulus and rich in terms of information (perceptual, motor, aesthetic, emotions). Further studies could employ motor experts as a population of interest to confirm or challenge our findings. The absence of an accuracy effect also calls for further investigation into the interplay between different cognitive and motoric components that contribute to action discrimination.

In conclusion, our study enriches our understanding of fluency as a multifaceted construct affecting both perception and action. While it confirmed some aspects of our experimental hypothesis, it also opened intriguing questions regarding the relative roles of sensitivity, reaction time, and accuracy in action prediction tasks, when using more ecologically valid stimuli. As the first in a series of planned experiments, these results provide a strong foundation for further exploration into the complex interrelationship between perceptual and motor fluency, laying out a fertile ground for future research.

Chapter 7: Aesthetic Preference & Action Prediction: an online study

7.1. Introduction

There is plenty of evidence on how perceptual and motor experience modulate our predictions about others' actions (*for a review, see Chapter 5*). Action features are inferred and predicted from movement kinematics (Saygin & Stadler, 2012). Neural studies looking at action prediction have shown higher AON activations during anticipation compared to observation for actions, using temporal occlusion paradigms (Balser et al, 2014). From the perspective of familiarity with the actions, some studies have also tested visual and motor mechanisms underpinning action prediction, finding that both visual and visuo-motor familiarity enhanced prediction performance (Mulligan et al., 2016). From the expertise perspective, the use of "high-skilled performers" has broadly been used, overall showing a motor enhancement when comparing perceptual to motor experts (Paull & Glencross, 1997; Ranganathan & Carlton, 2007; Aglioti et al., 2008; Rammami & Miall, 2004; Urgesi et al., 2012; Verfaillie & Daems, 2002). Sebanz and Shiffrar (2009) also found that experts made more accurate predictions than did novices, only when dynamic movement information was available, indicating that experts may have tapped into their own action repertoire.
But, importantly, other studies using dance movements have also shown stronger engagement of visual and sensorimotor regions when non-experts watched dance movements they like (Calvo-Merino et al, 2008; Cross et al, 2011).

Movement fluency has a profound impact on how we perceive and appreciate dance (Bläsing, Calvo-Merino, Cross, Jola, Honisch, & Stevens, 2012). In dance, fluency translates into smooth transitions, precise timing, and effortless execution of complicated moves, all of which contribute to the overall aesthetic appeal of a performance (Orgs et al., 2013). A dancer who demonstrates high levels of movement fluency can convey emotions more effectively, tell a story more convincingly, and captivate an audience more thoroughly (Davidson, 2012). Fluent movements are generally easier on the eyes and mind, adhering to principles of perceptual fluency by requiring less cognitive effort to interpret or predict (Reber, Winkielman, & Schwarz, 1998). Moreover, motor fluency often signifies mastery of technique and control, qualities highly prized in dance (Kaufman & Mann, 2019). When a dancer moves with fluidity and precision, we are more likely to perceive them as skilled and proficient, which can heighten our emotional engagement with the performance. Fluid movements are often intricately linked with the narrative or emotive aspects of dance, allowing the dancer to express nuances that might be lost if the movements were stiff or awkward (Hanna, 1987). The significance of movement fluency in dance can also be seen in different styles and cultural contexts (Kaeppler, 1978). Whether in ballet, contemporary dance, hip-hop, or traditional cultural dances, fluency serves as a universal yardstick for skill and artistic excellence (Orgs et al., 2016). In some dance forms, fluency is not just about the smoothness of movement but also the seamless integration of rhythmic patterns, body isolations, or intricate footwork, which adds layers of complexity and depth to the performance (Orgs et al., 2016).

Additionally, fluent movement in dance is often perceived as an indicator of a dancer's emotional and mental state (Davidson, 2012). A fluent dancer is often assumed to be more "in the zone", more connected to the music, and more in tune with their emotions, adding another layer to the audience's appreciation and understanding of the performance. In summary, movement fluency plays a pivotal role in our perception of dance, influencing how we assess skill, engage emotionally, and derive aesthetic pleasure from a performance. It serves as a universal metric for excellence, regardless of the specific style or cultural origin of the dance, enriching our experience as viewers and elevating the art form itself.

In recent times, the field of neuroaesthetics has laid the groundwork for understanding aesthetic preferences from both a physiological and methodological perspective (Chatterjee, 2011; Nadal and Pearce, 2011). This area of research typically investigates how the brain processes aesthetic experiences in real-time. Such studies are grounded in findings that correlate brain activity with experiences of aesthetic appreciation. Multiple brain regions involved in emotional evaluation, including the orbitofrontal and medial frontal cortex, ventral striatum, anterior cingulate gyrus, and insula, have been found to play a role in aesthetic experiences (Chatterjee and Vartanian, 2014).

In a seminal study by Calvo-Merino and colleagues (2010), they asked participants to make aesthetic judgements about body postures. They examined the effect of transcranial magnetic stimulation (TMS) on the participant’s preference between two postures and assessed whether this preference reflected the person’s own ratings when shown the postures individually. This subjectivist approach did not assume which postures would be aesthetically pleasing. They observed differential effects of repetitive TMS over the EBA and the ventral premotor area, with the latter tending to improve and the former tending to worsen judgements. This therefore suggests a role for these body-related brain areas in appreciating aesthetic aspects of the body in space.

In summary, movement fluency plays a pivotal role in our perception of dance, influencing how we assess skill, engage emotionally, and derive aesthetic pleasure from a performance. We know Aesthetic perception of human movement depends on “good” continuation of movement (Orgs et al 2013), and that body-related brain areas have a role in appreciating aesthetic aspects of the body in space (Calvo-Merino et al., 2010), in a way that movement fluency can have a modulatory effect on aesthetic appreciation.

There is plenty of evidence on how perceptual and motor experience modulate our predictions about others’ actions, from the familiarity with the action point of view, to a highly skilled performance enhancement. Action features can be inferred and predicted from movement kinematics. On the other hand, some studies using dance movements have shown stronger engagement of visual & sensorimotor regions when non-experts watch dance movements they like. Research studying aesthetics of dance, has also been proven that smooth and predictable movements are preferred in aesthetic evaluation. There is growing evidence to support that action prediction and aesthetic evaluation of dance movements share similar mechanisms evaluating visually the kinematics from biological motion.

With this experiment, we aimed to measure whether our aesthetic appreciation of dance movements could enhance our sensitivity predicting those actions. Based on recent evidence (Arslanova et al., 2023), we also aimed to explore whether predicting the correct continuation of body movements is associated with individual differences using different questionnaires exploring dance experience, alexithymia, empathy and interoceptive skills. It was hypothesized that participants would overall show better performance on congruent trials, based on previous empirical work *(see Chapter 6)*. A second hypothesis was to explore whether aesthetic appreciation or other individual differences could be related to, or even predict performance.

7.2. Methods

7.2.1. Participants

Forty-five participants (M=39.13, SD=10.89; 31 women) of which 6 were left-handed and 2 ambidextrous, took part in the experiment recruited via Prolific and paid at a rate of £9/hour. All participants had normal or corrected to normal vision, and reported no cognitive, attentional, or neurological impairment. Written and informed consent was obtained from all participants and the study was approved by the Psychology Research Ethics Committee at City, University London (ETH2223-0289), following guidelines and procedures established in the Declaration of Helsinki.

7.2.2. Materials

Twenty-four whole-body dance videos with scrambled faces were used in the experiment, half portrayed ballet movements and the other half capoeira moves (stimuli used with permission from Calvo-Merino et al., 2005). The videos were 3 seconds long, recorded in a neutral background, matching kinematic level in both styles according to speed, part of the body involved, body location in space and movement direction. For the occlusion paradigm, videos were decomposed in 75 frames (25 frames per second). The experimental task showed only the first half of the videos and created both congruent and incongruent conditions using a time-coherent or 400ms-delayed frame respectively as a target (same stimuli and paradigm as in previous experiment – *see Figure 18 or Chapter 6*). A key difference implemented from feedback in the previous study, was the clarity of the experimental question presented to participants during the temporal occlusion paradigm: *“Does the last frame follow a fluent continuation of the movement on the video?”.*

In the current experiment, a collection of standardized questionnaires was also used. These include the Multidimensional Assessment of Interoceptive Awareness-2 (MAIA-2), which assesses multiple dimensions of interoceptive awareness; the Goldsmiths Dance Sophistication Index (Gold-DSI), aimed at measuring an individual's engagement and sophistication in relation to dance; the Interpersonal Reactivity Index (IRI), which evaluates empathy and perspective-taking abilities in social interactions; and the Toronto Alexithymia Scale-20 (TAS-20), designed to measure emotional awareness and the capacity to articulate emotions. These instruments were selected for their strong psychometric qualities and their ability to yield detailed insights into the cognitive and emotional variables we aim to explore as individual differences *(see also section 1.4 Measurements & Methods)*.

7.2.3. Design

This study used a mixed-methods experimental design. The between-participants independent variable used was dance expertise, creating 2 groups for later comparison (experts and non-experts). All the other variables manipulated (IVs: dance style & congruency), measured (DVs: ACC, RTs), and calculated (quality of observers’ responses: d’ & c) in this study followed a within-participants design.

All participants had to watch and rate their preference for each of the dance videos to-be-used, prior to the main experimental task, and completed the Gold-DSI, IRI, MAIA-2 and

TAS-20 questionnaires. The same temporal occlusion paradigm implemented for the previous experiment was used as the action prediction task *(see Figure 18).* The first 1500ms of the video were displayed, followed by a 350ms occlusion and a target image. On the congruent trials, the target frame would represent a fluent continuation of the movement when the frame was following video’s onset, corresponding with correctly synchronised timing (frame 47 or 1850ms from onset). Incongruent trials would display a target frame representing a 400ms delay as the non-fluent continuation (corresponding to frame 57 or 2250ms since video onset). Both target frames were extracted within the same clip, portraying a continuation of the movement displayed, but only one frame represents a fluent continuation and therefore congruent with the timing from onset. With a total of 24 videos (12 ballet, 12 capoeira), and two equally possible conditions per video (50% congruent / 50% incongruent), trials were randomized within block and there were 4 blocks, adding up to a total of 192 trials. The experiment measured participants’ accuracy and reaction times when responding a 2AFC (yes/no) to the statement: “Does the last frame follow a fluent continuation of the movement on the video?”.

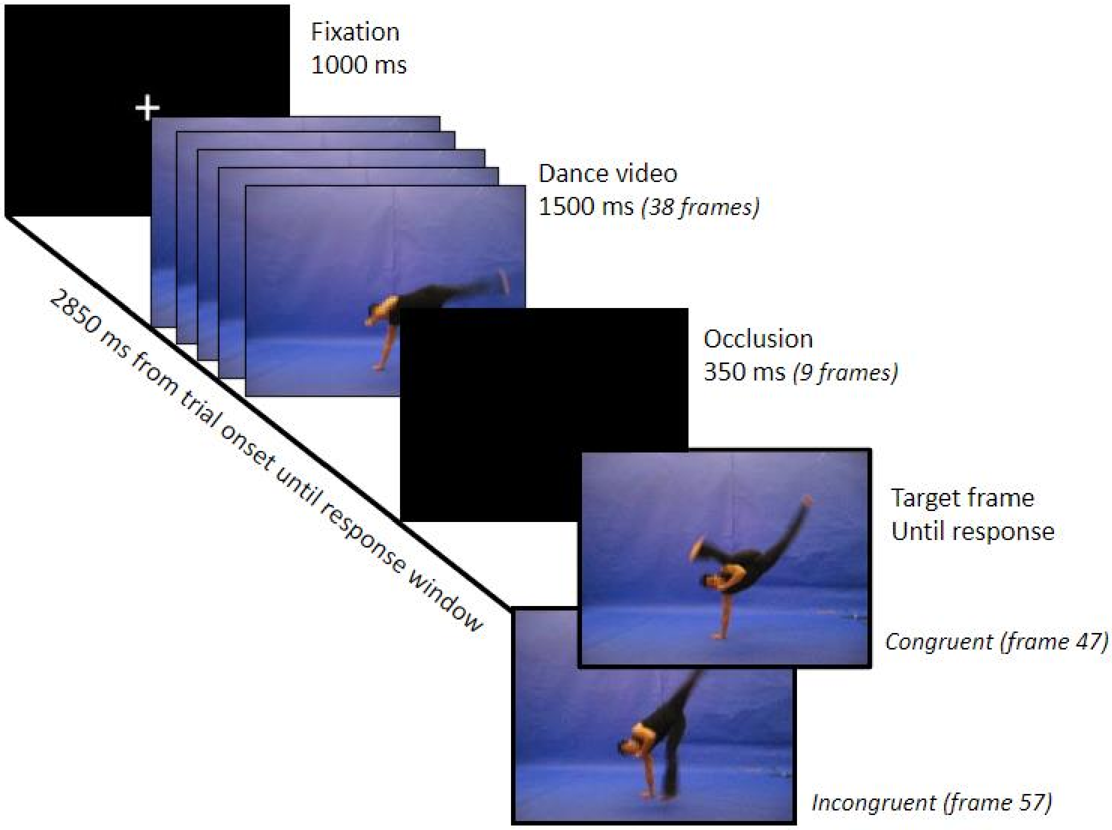


Figure 18. The experimental task used a temporal occlusion paradigm, showing 1500ms videos of different ballet and capoeira movements, followed by a 350ms occlusion and a target image from the same video which could be time-congruent with the video's onset or 400ms delayed. The participant had to respond yes or no to the question: "Does the last frame follow a fluent continuation of the movement on the video?".

7.2.4. Procedure

The study was implemented using the online platform *Testable* and advertised for online participation via *Prolific*. Participants were Prolific users, who signed up for the online study. Prior to participation, they had to read the information about the study and gave informed consent. There were 6 different tasks all participants had to complete in the same order.

The first task was to watch 24 different 3 seconds-long videos portraying 12 ballet and 12 capoeira movements, presented in randomised order, and give an aesthetic preference rating (0 to 100) for each. As the second task, they had to complete the Goldsmiths Dance Sophistication Index (Gold-DSI) questionnaire, which is focused on dance experience. Third task was the main experimental task. The same videos they had to rate on the first task, were used in the occlusion paradigm, were participants had to predict whether the target image was a correct continuation of the video. Each trial started with a 1000 ms fixation cross, then 1500ms video portraying a ballet or capoeira movement followed by a 350ms occlusion and a target static image until response *(see figure 18)*. Participants were instructed to respond yes or no to the statement: *"Does the last frame follow a fluent continuation of the movement on the video?"*. The response was recorded by participants' pressing "K" for yes or "L" for no, on their keyboard with their right hand. There were 48 different conditions, presented in randomised order and repeated over 4 blocks. The remaining tasks were to complete 3 more questionnaires in the following order: the Interpersonal Reactivity Index (IRI), which evaluates empathy and perspective-taking abilities in social interactions; the Multidimensional Assessment of Interoceptive Awareness-2 (MAIA-2), which assesses multiple dimensions of interoceptive awareness; and the Toronto Alexithymia Scale-20 (TAS-20). The experiment was paid at a rate of £9 per hour via Prolific, and on average it took 35 minutes to complete.

7.3. Results

Behavioural data (N=45) was collected through an online platform, using a temporal occlusion paradigm where participants had to respond as a 2AFC (yes/no) to whether the target image followed a fluent continuation from the video's movement. The main experimental manipulation consisted in using a target frame that would be congruent (fluent) or incongruent (non-fluent) with the timing since movement's onset. Two different dance styles were used, displaying ballet or capoeira moves. Accuracy (ACC) and reaction times (RTs) were measured for congruent and incongruent trials in both ballet and capoeira movements. Aesthetic ratings of the videos used in the paradigm, and a series of questionnaires were also collected during the experiment. It was hypothesized that participants would overall show better performance on congruent trials compared to incongruent ones. A second hypothesis was to explore whether aesthetic appreciation or related individual differences could influence or predict performance.

According to the experimental design, 2x2 RM ANOVA were used to analyse accuracy and reaction times separately, looking for evidence pointing towards the experimental hypothesis, but also a lack of effect on dance style. Effect sizes were also calculated and reported as Partial Eta Squared (PES). As a separate measure to evaluate the effectiveness of discrimination between conditions, sensitivity (d’) and response bias (criterion c) were calculated from the behavioural responses (see Chapter 1.4 Measurements & Methods). To explore the second experimental hypothesis, parametric correlations were used looking at the relationship between aesthetic preference and behavioural performance. To further investigate the role of individual differences in action prediction and fluency perception, a series of linear regression models were used. An index of “fluency enhancement” was computed as the difference between congruent and incongruent conditions for both behavioural measures, ACC, and RTs. For each dance style separately, this index representing the advantage of fluent perception (over non-fluent), was used to explore whether it could be related to aesthetic ratings or predicted by individual differences from the questionnaires’ data.

Table 4. Descriptive statistics showing mean (M) and standard deviation (SD) for all measures collected in the study. Accuracy and reaction time measures were specified per experimental condition: congruent and incongruent trials & for each dance style separately (ballet/ capoeira). Other measures included in this table: aesthetic preference ratings, response bias, sensitivity, Gold-DSI, IRI, MAIA-2 & TAS-20.

			M	SD
Accuracy	Ballet	Congruent	0.66	0.17
		Incongruent	0.55	0.20
	Capoeira	Congruent	0.76	0.17
		Incongruent	0.52	0.16
Reaction Times	Ballet	Congruent	986.69	347.09
		Incongruent	1010.90	346.67
	Capoeira	Congruent	989.81	362.52
		Incongruent	1052.86	391.21
Aesthetics Preference	Ballet		64.75	16.54
	Capoeira		44.78	17.12
response bias (c)			0.34	0.49
sensitivity (d')	Ballet		0.80	0.85
	Capoeira		0.89	0.66
Gold-DSI	Body Awareness		3.93	0.58
	Social Dancing		3.94	0.33
	Urge to Dance		4.97	0.77
	Dance Training		1.81	1.14
	Observational Dance Experience		3.62	0.58
IRI	Empathic Concern		3.14	0.33
	Fantasy		3.10	0.47
	Personal Distress		3.00	0.55
	Perspective Taking		3.36	0.33
MAIA-2	Attention Regulation		3.96	0.88
	Body Listening		3.67	1.00
	Emotional Awareness		4.35	1.00
	Not-Distracting		3.77	1.03
	Noticing		4.15	0.81
	Not-Worrying		3.67	0.57
	Self-Regulation		3.94	1.03
	Trust		4.39	1.00
	MAIA-2 Total		3.97	0.58
TAS-20	Difficulty Describing Feelings		13.13	2.73
	Difficulty Identifying Feeling		13.80	5.77
	Externally Oriented Thinking		27.24	2.15
	TAS-20 Total		54.18	7.71

Accuracy was analysed using a 2x2 RM ANOVA. Results showed a main effect of dance (F(1,44)=7.39, p=.009, PES=.144), with more accurate values in capoeira (M=.606, SE=.012) compared to ballet (M=.606, SE=.014); a main effect of congruency (F(1,44)=29.51, p<.001, PES=.401), with more accurate values in congruent (M=.709, SE=.019) compared to incongruent trials (M=.534, SE=.021); and a 2-way interaction (F(1,44)=4.87, p=.033, PES=.100), explained by significant differences between dance styles on congruent (Mean difference = .101, SE=.033, p=.003) but not on incongruent trials (Mean difference = .037, SE=.034, p=.290).

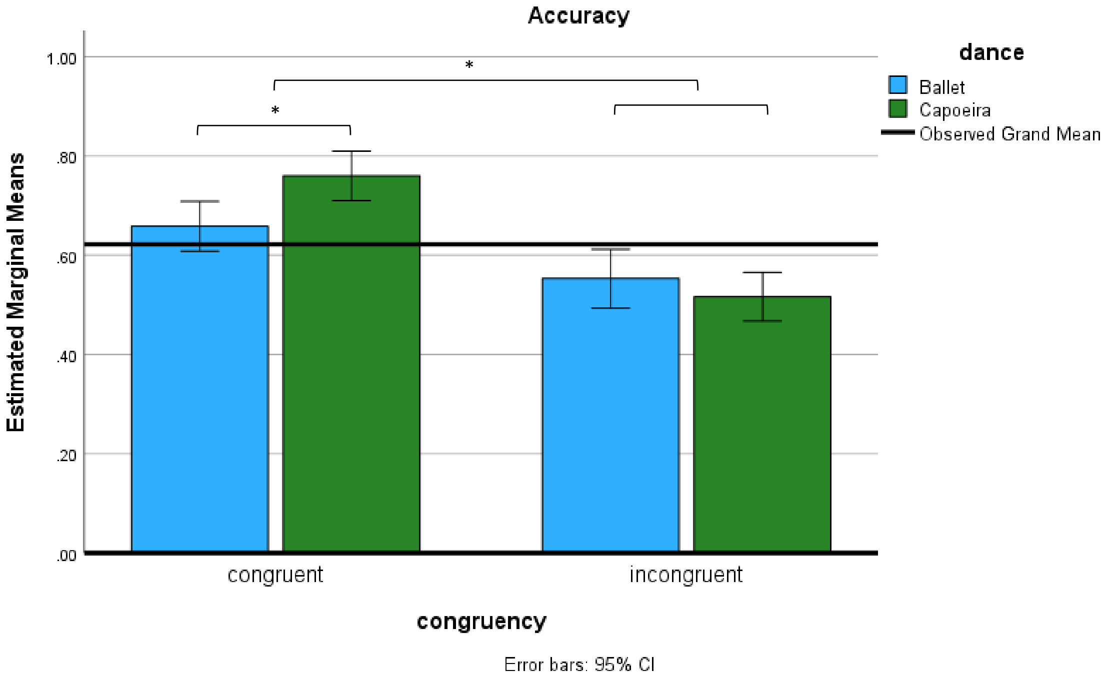


Figure 19. Bar graph represents mean values for accuracy measures across experimental conditions. Error bars represent 95% confidence intervals. Significant differences are represented as (*).

Reaction Times (RTs) were analysed using a 2x2 RM ANOVA. Results showed a main effect of congruency on reaction times (F(1,44)=11.74, p=.001, PES=.211), showing faster responses to congruent (M=988.23, SE=51.87) compared to incongruent trials (M=1031.88, SE=54.39).

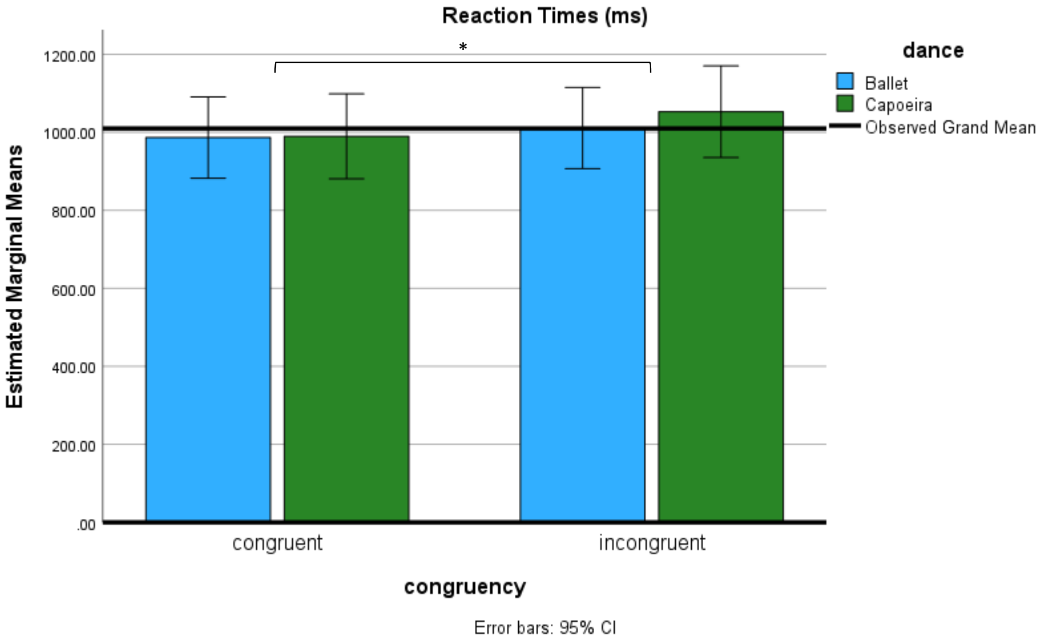


Figure 20. Bar graph represents mean values for RTs measures across experimental conditions. Error bars represent 95% confidence intervals.

Sensitivity (d’) was calculated for each dance style and analysed using a paired samples t-test. Results showed no significant difference on how well participants could discriminate the movements when comparing dance styles (t(44)=.71, p=.481). Instead, a strong positive correlation between sensitivity measures (d’) for ballet and capoeira movements was found (r(43)=.435, p=.003).

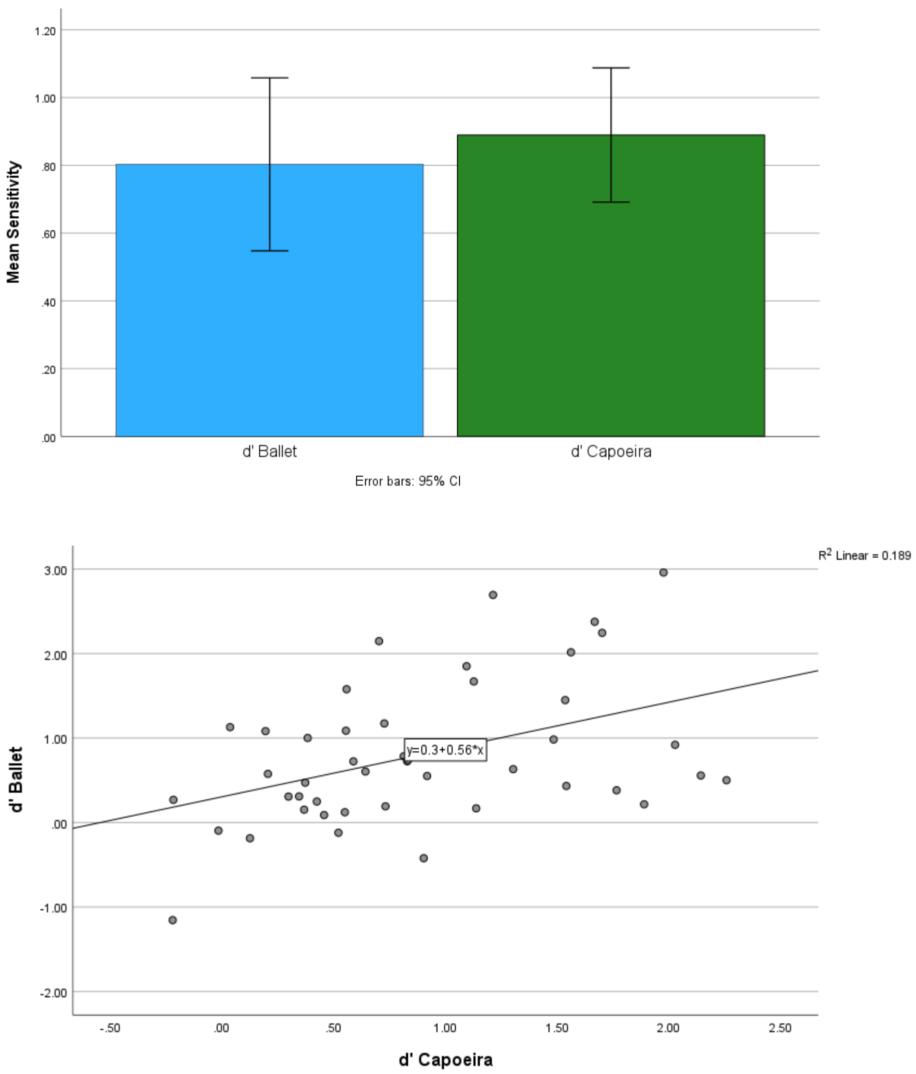


Figure 21. Top, bar graph represents mean values for sensitivity measures for ballet and capoeira. Error bars represent 95% confidence intervals. Bottom, plot represents correlation between sensitivity measures for both dance styles.

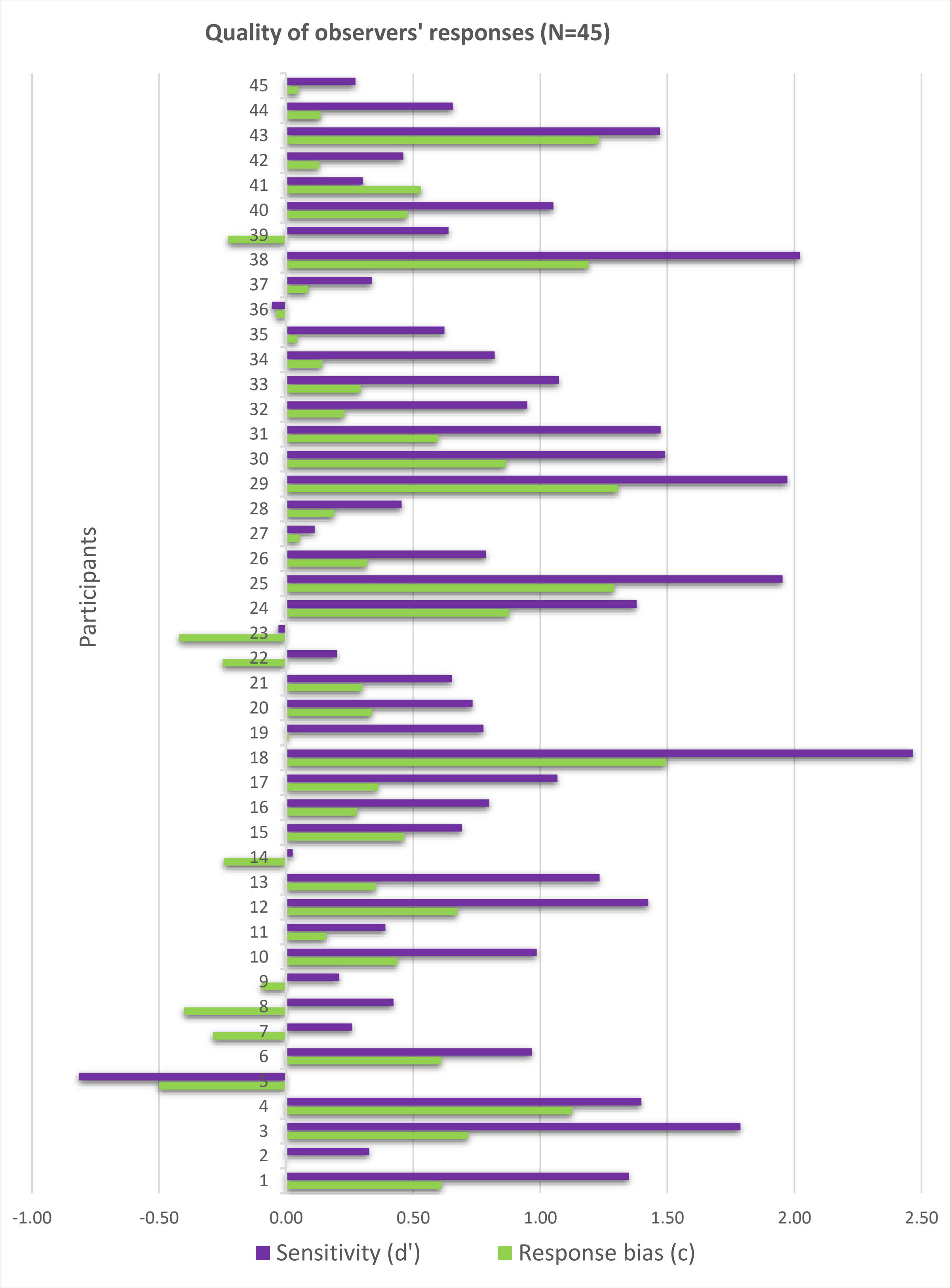


Figure 22. Graph represents mean values for sensitivity and response bias across experimental conditions for each participant.

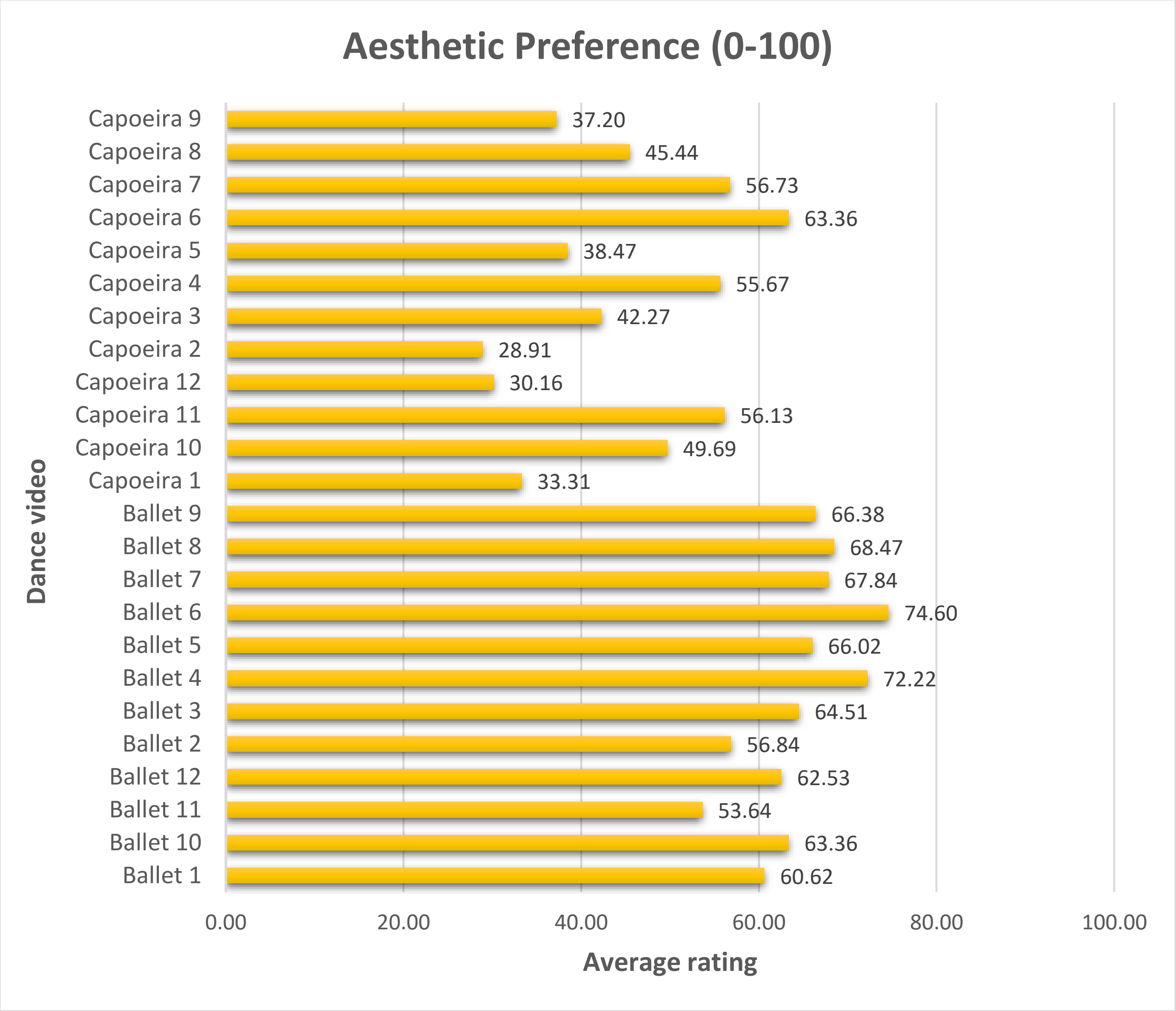


Figure 23. Mean aesthetic preference ratings (0-100), averaged across participants and displayed per dance video clip. A total of 24 videos were used in this experiment, displaying 12 different movements per style, previously matched for kinematic features across styles.

Aesthetic preference was analysed using a paired sample t-test to compare dance styles. Results showed a significant difference (t(44)=595, p<.001), meaning that on average participants liked more ballet (M=64.75, SD=16.55) compared to capoeira moves (M=44.78, SD=17.12).

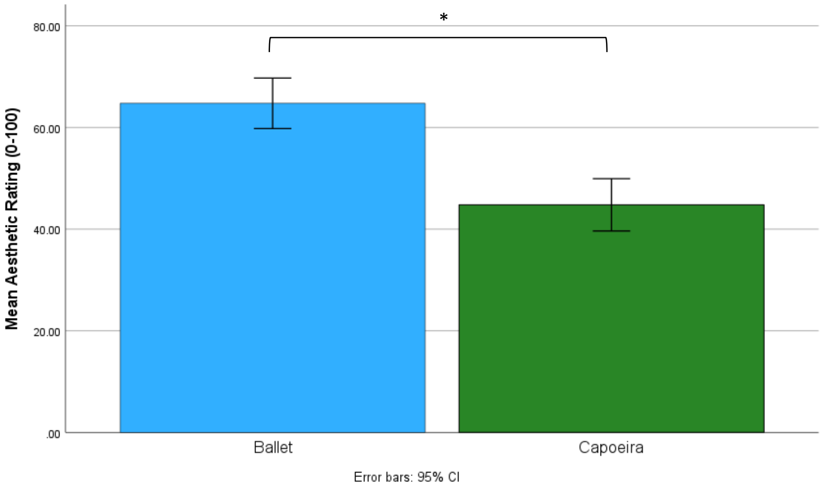


Figure 24. Bar graph represents mean values for aesthetic preference ratings, averaged by dance style. Error bars represent 95% confidence intervals. Significant differences are represented as (*).

To further investigate the role of individual differences, an index of fluency enhancement was computed as the difference between congruent and incongruent conditions for both behavioural measures (ACC and RTs). Pearson correlations were used to explore these potential relationships between behavioural measures and aesthetic preferences. Results showed no significant relationships between accuracy and aesthetic preference (p>.05). Although, RTs showed significant relationships expressing inverse patterns depending on the dance style. Aesthetic preference ratings for ballet were positively correlated with the difference between congruent and incongruent conditions when watching ballet moves (r=.364, p=.014). Aesthetic preference ratings for capoeira were negatively correlated with the difference between congruent and incongruent conditions when watching capoeira moves (r= -.5, p<.001).

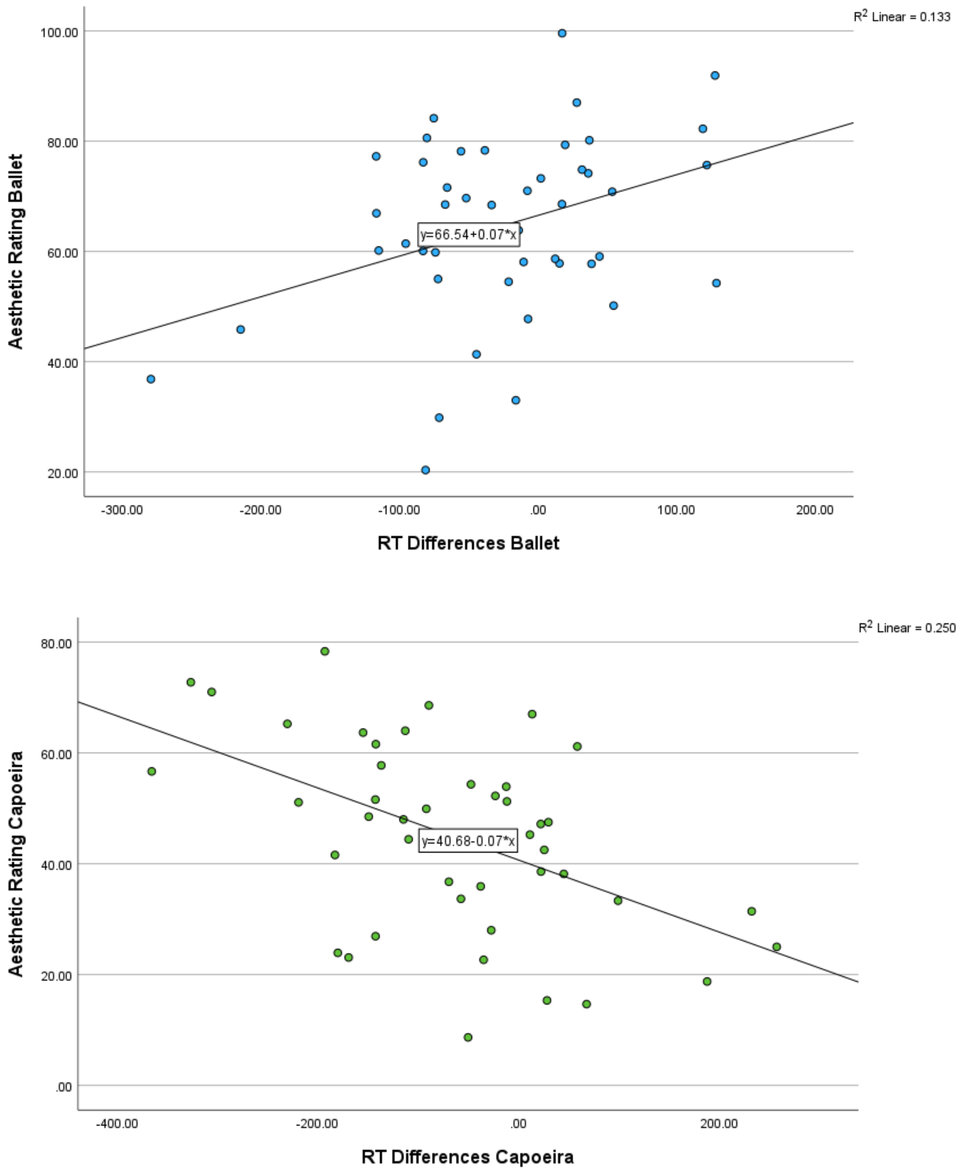


Figure 25. Top, plot represents correlation between aesthetic ratings RT differences in ballet. Bottom, plot represents correlation between aesthetic ratings RT differences in capoeira.

Further analysis aimed to explore whether individual differences, measured via questionnaires Gold-DSI, IRI, MAIA-2 and TAS-20 *(see Table 4),* could predict the effects observed in performance (i.e., significant dance style differences on ACC & RTs). A separate multiple linear regression models was computed for each measure to be predicted, using a stepwise selection model, meaning only significant predictors would enter the model from all the included as potential predictors (scales and subscales collected with the questionnaires).

The first multiple linear regression was conducted to examine how individual differences predicted the differences in ACC for ballet moves. The model was statistically significant, F(2, 43) = 6.341, p = .016, accounting for 12.9% of the variance (Adjusted R Square = .108). The results indicate that TAS-20 significantly predicted ACC differences in ballet, β = .358, t(22)=2.52 p = .016.

The second multiple linear regression was conducted to examine how individual differences predicted the differences in ACC for capoeira moves. The model was statistically significant, F(2, 43) = 5.67, p = .022, accounting for 11.7% of the variance (Adjusted R Square = .096). The results indicate that Difficulty Identifying Feelings (subscale of TAS-20) significantly predicted ACC differences in capoeira, β = -.341, t(22)=2.38 p = .022.

The third multiple linear regression was conducted to examine how individual differences predicted the differences in RTs for ballet moves. The model was statistically significant, F(2, 44) = 4.91, p = .012, accounting for 18.9% of the variance (Adjusted R Square = .151). The results indicate that Body Awareness (subscale of Gold-DSI) and Trust (subscale of MAIA-2) significantly predicted RT differences in ballet, β = -.415, t(22)=-2.86 p = .007 and β = .302, t(22)=2.07 p = .044 respectively.

The last multiple regression conducted to examine how individual differences predicted the differences in RTs for capoeira moves showed no significant predictors.

7.4. Discussion

Our study's findings strongly support the main hypothesis, revealing that participants performed better and faster when observing a fluent continuation of movements, in alignment with previous work on the importance of movement fluency in dance and aesthetic experiences (Bläsing et al., 2012; Orgs et al., 2013). This reaffirms the critical role of fluency in the perceptual evaluation of dance, augmenting our understanding of how smooth and predictable movements contribute to aesthetic appreciation.

Interestingly, ballet moves were associated with greater accuracy among participants. This specificity could be attributed to the wider cultural exposure and acceptance of ballet as a refined form of artistic expression (Kaufman & Mann, 2019). However, it's crucial to note that the sensitivity measures did not indicate a difference in discriminability based on dance style. This suggests that while cultural or aesthetic preferences may influence accuracy, they do not necessarily impact the fundamental perceptual mechanisms involved in action prediction and kinematics, as highlighted by Saygin & Stadler (2012). Furthermore, a strong correlation was observed between participants' sensitivity to both ballet and capoeira styles, supporting the idea that the underlying neural mechanisms responsible for action prediction and aesthetic evaluation may be generalized across different forms of dance (Mulligan et al., 2016; Sebanz & Shiffrar, 2009).

The study also found intriguing associations between aesthetic ratings and response times, particularly in the context of fluent and non-fluent conditions. Specifically, higher aesthetic ratings for ballet were directly related to higher differences in response times between the fluent and non-fluent conditions, while the opposite pattern was observed for capoeira. This could be interpreted as further evidence for the modulatory effect of aesthetic preference on cognitive processing speeds (Chatterjee, 2011; Nadal and Pearce, 2011).

Finally, our regression analysis illuminated the role of individual differences in participants' performance, as hypothesized. These individual variances could be predicted by self-reported measures from questionnaires probing into dance experience, alexithymia and interoceptive skills. This finding is congruent with recent evidence suggesting that both perceptual and emotional factors can influence the aesthetic evaluation of dance and action prediction (Arslanova et al., 2023).

In conclusion, our results offer compelling insights into the complex interplay between aesthetic appreciation, perceptual fluency, and action prediction, enriching our understanding of how these processes are integrated in the human experience of dance. Future research may delve deeper including experimental manipulations on kinematics (not just matched and controlled), and comparison with experts to offer more nuanced perspectives.

Chapter 8: General Discussion

8.1. Embodiment, attention & WM

The integration of action representation, body perception, working memory (WM), and attention can be framed as a synergistic confluence of cognitive and neural systems. Action representation is rooted in the dynamic intersection between cognitive models and neural architecture. The notion of embodiment offers a conceptual vantage point, asserting that our physical presence substantially impacts our cognitive processes related to action representation (Gallese & Sinigaglia, 2011).

The perceptual aspects of body and movement are intrinsically linked and share common neural substrates, aiding not only motor functions but also social interactions (Blake & Shiffrar, 2007). The interconnectedness of these perceptual domains serves evolutionary and adaptive purposes, underscoring the need for their integrated exploration in both cognitive psychology and neuroscience.

The concept of attention serves as an operational model for understanding how cognitive resources are allocated to process specific stimuli (Chun et al., 2011). This cognitive mechanism acts as a filter, reconciling bottom-up sensory inputs with top-down cognitive modulations. In other words, attention operates at the nexus of sensory input and cognitive expectations, influencing perceptual outcomes based on previous experiences, beliefs, and current mental states (Desimone & Duncan, 1995).

The multicomponent model of WM provides a framework for understanding the cognitive structures and processes underlying the storage and manipulation of information (Baddeley & Hitch, 1974). Recent models have extended this framework to propose a convergence between the neural systems implicated in perception and those used for the retention of information in WM (Postle, 2006). This intersection has significant evolutionary underpinnings, particularly when it comes to retaining information related to bodily movements and actions.

Recent developments in the study of attention have extended the framework to encompass both external and internal cognitive representations (Chun et al., 2011). Such a reconceptualization allows for a more nuanced understanding of how attention modulates both sensory inputs and internal cognitive schemas. Specifically, attention operates as a network of cognitive and neural systems that prioritize information based on salience and importance (Narhi-Martinez et al., 2022). This, in turn, has far-reaching implications for our understanding of visual working memory, particularly in the context of top-down modulatory processes that facilitate selective focus on relevant information.

To summarize, action representation, body perception, WM, and attention are not isolated constructs but rather interrelated elements of a cohesive cognitive-neural ecosystem. Each domain enriches and informs the other, providing a holistic understanding of human cognition that is greater than the sum of its individual parts. Future research endeavours should aim to explore these interrelationships in greater depth, utilizing interdisciplinary methodologies that can capture the complexity of these interconnected systems.

8.2. What could we learn from WM for bodies & attentional control?

In an effort to elucidate the intricate relationship between action representation, body perception, working memory (WM), and attention, it is imperative to delineate their respective roles and interplay. From the vantage point of embodiment, action representation epitomizes the synthesis of cognitive schemas and neural processes, advancing the notion that corporeal presence significantly impacts cognitive functioning, including action comprehension (Glenberg, 2010). This view aligns well with research on body and movement perception, which posits that these sensory processes, built on shared neural substrates, facilitating not only motor actions but also social interactions (Blake & Shiffrar, 2007).

Concurrently, attention serves as the regulatory mechanism that arbitrates the salience of sensory inputs and internal schemas (Chun et al., 2011). Attentional control therefore operates at the intersection of bottom-up sensory inputs and top-down cognitive processes, shaping perception into a coherent and adaptive experience (Desimone & Duncan, 1995). In parallel, evolving models of WM underscore its integrative capacity to encode and retain multi-modal information, implicating common neural pathways in perception and memory (Postle, 2006).

Current findings illuminate the nuanced interplay between WM, attentional modulation, and their impact on both behaviour and neural activity. Behavioural results indicate that participants perform better in tasks with congruent information and lower memory load. This appears to be in alignment with Oberauer's 'slots' model of WM, which suggests that attentional control is pivotal in determining the type and quantity of information that can be held in WM (Oberauer & Hein, 2012).

Visual evoked potentials (VEPs) and somatosensory evoked potentials (SEPs) revealed that both visual and somatosensory areas are actively engaged during the retention interval, and their activity is modulated by both the type of stimuli and attentional focus. These observations corroborate earlier work indicating that body-related cortices are specifically recruited during WM tasks involving bodily images (Galvez-Pol et al., 2018a; b). Similarly, the saliency effect of colour supports theories on the suppressive role of attention in improving data fidelity in WM (Gazzaley & Nobre, 2012). Such effects also lend empirical support to the study's hypothesis concerning attentional modulation in visual and somatosensory areas, reaffirming the fluidity of attentional allocation (Serences et al., 2009; Harrison & Tong, 2009).

In sum, the empirical data derived from this research deepens our comprehension of how attentional mechanisms—both top-down and bottom-up—jointly influence the encoding and retention processes in WM. This, in turn, contributes significantly to our growing understanding of bodily representation and attentional regulation. As avenues for future investigation, the inclusion of more diverse stimuli and exploration of additional neural markers may offer further insights into this complex cognitive network.

8.3. Why looking through the lens of prediction?

Action prediction represents a multidimensional cognitive process, demanding the orchestration of several intricate components. It encompasses the detection of agents, the utilization of internal representations, and the inference of intentions, all within a dynamic and ever-changing context. This cognitive feat leans on a rich tapestry of sensorimotor experiences, social knowledge, and contextual expectations. The ability to anticipate the actions of others is instrumental in our understanding of social cognition and empathy, greatly enhancing our capacity for effective social interactions and adaptive behaviours.

The journey of action prediction embarks with the perception of sensory cues signalling agency and animacy, such as faces and biological motion. These cues act as guiding stars, leading to the activation of internal sensorimotor models, finely tuned over time through personal experiences and self-awareness. These models not only empower us to predict physical movements but also serve as a foundation for deciphering the hidden intentions and goals that underlie those actions. Adding depth to this intricate process is the influence of contextual information, encompassing socio-affective factors and cultural backgrounds, which further refines our predictive abilities.

At the heart of this complex cognitive ballet lies Prediction Error, a pivotal mechanism that offers a window into the flexibility and adaptability of both sensorimotor predictions and the broader framework of social cognition. This mechanism continuously recalibrates and refines our anticipatory models, integrating different types of information and top-down influences.

In summary, the complex nature of action prediction necessitates an integrative perspective that combines attention, perception, sensorimotor functions, implicit memory, and context, including socio-affective elements and cultural norms. Failing to unite these domains may lead us to overlook the synergistic effects that these elements exert on our capacity to predict actions and intentions. Prediction Error stands as a testament to the adaptability of this integrated system, serving as a critical mechanism for the continuous improvement of our anticipatory models based on diverse information sources.

This holistic review provides a comprehensive lens through which we can delve into not only the mechanics of action prediction but also its profound implications for social cognition, empathy, and theory of mind. The future of research in this field should continue to explore the neural mechanisms underpinning action prediction and its far-reaching implications for diverse domains, including social neuroscience, developmental psychology, and clinical populations. A heightened understanding of action prediction promises to unlock deeper insights into human behaviour and the intricate dance of social interactions.

8.4. Empirical study of AP using fluency

From this study, we've gained valuable insights into the interplay of perceptual and motor fluency, particularly within the domain of biological motion. Our findings affirm our expectations that participants exhibited significantly faster reaction times and heightened sensitivity (d') when discriminating fluent (congruent) biological movements compared to non-fluent (incongruent) ones. These results underscore the close relationship between our perceptual and motor systems. They suggest that our internal sensorimotor representations, crafted through a blend of visual and motor experiences, play a role in shaping how we perceive and respond to complex actions. This concept aligns with the idea of motor resonance, where watching biological motion activates neural pathways used in executing those very movements.

Practically, these findings hold significance for fields ranging from sports training to clinical rehabilitation. The fact that reaction times and sensitivity are influenced by fluency could be pivotal for designing training programs aimed at improving both the perception and execution of intricate biological movements, like dance, which formed the core of our study stimuli.

Nevertheless, it's important to acknowledge certain limitations. Our study focused exclusively on one form of biological motion—dance—which is exceptionally complex in terms of information content, encompassing perceptual, motoric, aesthetic, and emotional aspects. Further research could explore the involvement of motor experts to validate or challenge our findings. Additionally, the absence of an accuracy effect prompts further inquiry into the nuanced interplay between various cognitive and motor components in action discrimination.

In summary, our study has deepened our understanding of fluency as a concept that impacts both perception and action. While it confirmed specific aspects of our expectations, it has also prompted questions about the roles of sensitivity, reaction time, and accuracy in action prediction tasks, especially when using more ecologically valid stimuli. As the first step in an ongoing series of experiments, these results provided a strong foundation for future research, shedding light on the complex relationship between perceptual and motor fluency.

8.5. What else could we learn from aesthetics & individual differences in AP?

This study provided robust support for our primary hypothesis, revealing that participants exhibited superior performance and quicker responses when observing fluent continuations of movements. This aligned harmoniously with prior research emphasizing the significance of movement fluency in dance and aesthetic encounters. It reaffirms the central role of fluency in the perceptual assessment of dance, deepening our comprehension of how smooth and predictable movements contribute to aesthetic enjoyment.

An intriguing observation emerged regarding ballet movements, which correlated with greater accuracy among participants. This specificity can be attributed to the broader cultural exposure and recognition of ballet as a refined artistic form. However, it's vital to emphasize that sensitivity measurements indicated no substantial differences in discriminability based on dance style. This implies that while cultural or aesthetic inclinations may influence accuracy, they don't necessarily impact the fundamental perceptual mechanisms involved in action prediction and kinematics. This is in line with prior research that underscores the common neural mechanisms underlying action prediction and aesthetic evaluation across various dance forms.

Additionally, our study unveiled noteworthy connections between aesthetic ratings and response times, particularly in the context of fluency. Notably, higher aesthetic ratings for ballet were associated with more pronounced differences in response times between fluent and non-fluent conditions, with an inverse pattern observed for capoeira. This implies that aesthetic preferences may interact with cognitive processing speeds, aligning with existing research in this domain.

Finally, our regression analysis shed light on the role of individual differences in participant performance, in accordance with our expectations. These individual variances could be predicted by self-reported measures related to dance experience, alexithymia, and interoceptive skills. This finding aligns with recent evidence highlighting the influence of both perceptual and emotional factors on the aesthetic evaluation of dance and action prediction.

In conclusion, our results provide compelling insights into the intricate interplay between aesthetic appreciation, perceptual fluency, and action prediction, enhancing our understanding of how these processes converge in the human experience of dance. Future research can further delve into this domain, considering experimental manipulations on kinematics and comparisons with motor experts to offer deeper and more nuanced perspectives.

References

Abernethy, B., & Zawi, K. (2007). Pickup of essential kinematics underpins expert perception of movement patterns. Journal of Motor Behavior, 39(5), 353-367.

Abernethy, Z., & Jackson, R. W. (2008). Expertise and the perception of kinematic and situational probability information. Perception, 37(8), 1120-1133.

Abreu, A. M., Balser, A., et al. (2012). Sensory-motor brain responses to action observation are influenced by expertise: Evidence from professional athletes. Behavioral Brain Research, 228(2), 437-444.

Adam, K. C. S., Rademaker, R. L., & Serences, J. (2021). Evidence for, and challenges to, sensory recruitment models of visual working memory. https://doi.org/10.31234/osf.io/wb5e6

Adam, K. C., Robison, M. K., & Vogel, E. K. (2021). Using anticipatory activity in early visual cortex to predict behavior in masked priming. Journal of Cognitive Neuroscience, 33(1), 146-160.

Adam, R., & Elsner, B. (2018). Age-related changes in action prediction during sports observation across different skill levels. Journal of Sport & Exercise Psychology, 40(2), 63-71.

Adam, R., Daum, M. M., et al. (2016). Action observation in the infant brain: The role of body form and motion. Social Cognitive and Affective Neuroscience, 11(9), 1444-1452.

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. Nature Neuroscience, 11(9), 1109–1116.

Ahlfors, S. P., Han, J., Belliveau, J. W., & Hämäläinen, M. S. (2010). Sensitivity of MEG and EEG to source orientation. Brain Topography, 23(3), 227-232.

Ahn, W. Y., & Vassileva, J. (2016). Machine-learning identifies substance-specific behavioral markers for opiate and stimulant dependence. Drug and alcohol dependence, 161, 247-257.

Aitchison, L., & Lengyel, M. (2017). With or without you: predictive coding and Bayesian inference in the brain. "Current Opinion in Neurobiology," 46, 219-227.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. Trends in Cognitive Sciences, 4(7), 267-278.

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. Personality and Social Psychology Review, 13(3), 219-235.

Amoruso, L., Panasiti, M. S., et al. (2014). Time course of motor resonance during action observation: Evidence for a facilitatory role. Journal of Cognitive Neuroscience, 26(9), 2086-2101.

Andersen, R. A., & Cui, H. (2009). Intention, action planning, and decision making in parietal-frontal circuits. Neuron, 63(5), 568-583.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2013). Value-driven attentional capture. Proceedings of the National Academy of Sciences, 108(25), 10367-10371.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2021). Learned value magnifies salience-based attentional capture. PLoS ONE, 6(11), e27026.

Angelaki, D. E., & Cullen, K. E. (2008). Vestibular system: The many facets of a multimodal sense. Annual Review of Neuroscience, 31, 125-150.

Arslanova, I., Kotsaris, V., & Tsakiris, M. (2023). Perceived time expands and contracts within each heartbeat. Current biology, 33(7), 1389-1395.

Astafiev, S. V., Shulman, G. L., Stanley, C. M., Snyder, A. Z., Van Essen, D. C., & Corbetta, M. (2004). Functional organization of human intraparietal and frontal cortex for attending, looking, and pointing. Journal of Neuroscience, 24(11), 4687-4696.

Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. Nature Neuroscience, 7(5), 542–548.

Atkinson RC, Shiffrin RM (1968). Spence KW, Spence JT (eds.). Human Memory: A Proposed System and its Control Processes. Psychology of Learning and Motivation. Vol. 2. Academic Press. pp. 89–195. doi:10.1016/S0079-7421(08)60422-3

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In The psychology of learning and motivation (Vol. 2, pp. 89-195). Academic Press.

Avenanti A, Candidi M, Urgesi C. 2013. Vicarious motor activation during action perception: beyond correlational evidence. Front Hum Neurosci 2013 May; 7 185.

Awh, E., & Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. Trends in Cognitive Sciences, 5(3), 119-126.

Awh, E., Vogel, E. K., & Oh, S.-H. (1998). Interactions between attention and working memory. Neuroscience, 139(1), 201-208.

Baars, B. J., & Gage, N. M. (2010). Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience. Elsevier.

Bach, P., Nicholson, T., & Hudson, M. (2014). The affordance-matching hypothesis: How objects guide action understanding and prediction. Frontiers in Human Neuroscience, 8, 254.

Baddeley A. 2003. Working memory: looking back and looking forward. Nat. Neurosci. Rev, 4, 829-839.

Baddeley, A. (1993). Working memory or working attention? In Attention: Selection, awareness, and control (pp. 351-371). Oxford: Clarendon Press.

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. Annual Review of Psychology, 63, 1-29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (Vol. 8, pp. 47–89). New York: Academic Press.

Badre, D., & D'esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical?. Nature Reviews Neuroscience, 10(9), 659-669.

Bagby, R. M., Taylor, G. J., & Parker, J. D. (1994). The Twenty-Item Toronto Alexithymia Scale—I. Item selection and cross-validation of the factor structure. Journal of Psychosomatic Research, 38(1), 23-32.

Balser, A., Abreu, A. M., et al. (2014). Dissociation of neural mechanisms underlying explicit judgments and implicit moves during observational learning. NeuroImage, 87, 232-243.

Balser, N., Lorey, B., Pilgramm, S., Naumann, T., Kindermann, S., Stark, R., ... & Zentgraf, K. (2014). Prediction of human actions: expertise and task-related effects on neural activation of the action observation network. Human Brain Mapping, 35(8), 4016-4034.

Balsters, J. H., Ramnani, N., & Berridge, K. C. (2016). Differential neural dynamics of reward and punishment processing in antisocial personality disorder. Brain, 139(12), 3336-3349.

Bar, M. (2004). Visual objects in context. Nature Reviews Neuroscience, 5(8), 617–629.

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. Trends in Cognitive Sciences, 11(7), 280-289.

Bar, M. (2009). Predictions: A universal principle in the operation of the human brain. Philosophical Transactions of the Royal Society B, 364(1521), 1181–1182.

Barsalou, L. W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22(4), 577-609.

Barsalou, L. W. (2008). Grounded cognition. Annual Review of Psychology, 59, 617-645.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. Science, 321(5890), 851-854.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1997). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1-3), 7-15.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. Science, 275(5304), 1293-1295.

Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. Vision Research, 49(10), 1154-1165.

Becke, A., Müller, N. G., Vellage, A. K., & Weber, B. (2015). Visuospatial imagery is a fruitful source of “remembering” in the blind. Neuropsychologia, 68, 104-121.

Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. Psychological Science, 13(4), 370-374.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. Science, 308(5721), 529-534.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. Annual Review of Neuroscience, 33, 1-21.

Blake, R., & Shiffrar, M. (2007). Perception of human motion. Annual Review of Psychology, 58, 47-73.

Blakemore, S. J., & Frith, C. D. (2005). The role of motor contagion in the prediction of action. Neuropsychologia, 43(2), 260-267.

Blakemore, S. J., Wolpert, D., & Frith, C. (2001). The cerebellum contributes to somatosensory cortical activity during self-produced tactile stimulation. NeuroImage, 10(4), 448-459.

Bläsing, B., Calvo-Merino, B., Cross, E. S., Jola, C., Honisch, J., & Stevens, C. J. (2012). Neurocognitive control in dance perception and performance. Acta Psychologica, 139(2), 300-308.

Bolis, D., & Schilbach, L. (2018). Observing and participating in social interactions: Action perception in individuals with autism spectrum disorders. Frontiers in Integrative Neuroscience, 12, 3.

Born, R. T., & Bradley, D. C. (2005). Structure and function of visual area MT. Annual Review of Neuroscience, 28, 157-189.

Brattan, V., Roberts, J. W., & Krekelberg, B. (2015). The relationship between movement speed and time-to-collision in the prediction of observed actions. PLOS ONE, 10(9), e0136132.

Brault, S., Dessing, J. C., & Craig, C. M. (2012). From a ballistic movement to catching a ball at top speed: The role of vision in anticipation. Brain Research, 1478, 47-58.

Brewer, J. A., Worhunsky, P. D., Gray, J. R., Tang, Y. Y., Weber, J., & Kober, H. (2011). Meditation experience is associated with differences in default mode network activity and connectivity. Proceedings of the National Academy of Sciences, 108(50), 20254-20259.

Broadbent, D. E. (1958). Perception and communication. Pergamon Press.

Brown, H., Friston, K. J., & Bestmann, S. (2011). Active inference, attention, and motor preparation. Frontiers in Psychology, 2, 218.

Burk, D., Ingram, J. N., et al. (2014). The action-specific effect of action observation. Brain and Cognition, 87, 108-115.

Butler, J. S., Smith, S. T., Campos, J. L., & Bülthoff, H. H. (2011). Bayesian integration of visual and vestibular signals for heading. Journal of Vision, 11(13), 23.

Calvo-Merino B, Grèzes J, Glaser DE, Passingham RE, Haggard P. 2006. Seeing or doing? Influence of visual and motor familiarity in action observation. Current Biology 16:1905-1910. Calvo-Merino, B,

Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: configural effects in action observation. Psychological Research PRPF, 74, 400-406.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. Cerebral Cortex, 15(8), 1243-1249.

Calvo-Merino, B., Grezes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or doing? Influence of visual and motor familiarity in action observation. Current Biology, 15(3), 190-193.

Calvo-Merino, B., Jola, C., Glaser, D. E., & Haggard, P. (2008). Towards a sensorimotor aesthetics of performing art. Consciousness and cognition, 17(3), 911-922.

Calvo-Merino, B., Urgesi, C., Orgs, G., Aglioti, S. M., & Haggard, P. (2010). Extrastriate body area underlies aesthetic evaluation of body stimuli. Experimental brain research, 204, 447-456.

Cañal-Bruland, R., & Schmidt, M. (2009). Response bias in judging deceptive movements. Attention, Perception, & Psychophysics, 71(5), 1124-1130.

Cañal-Bruland, R., van der Kamp, J., & van Kesteren, J. (2010). An examination of motor and perceptual contributions to the recognition of deception from others' actions. Human movement science, 29(1), 94-102.

Cannon, J., Ferguson, H. J., & Fulceri, F. (2021). Predictive coding and autism: A review and synthesis of evidence. Behavioral and Brain Sciences, 44, e142.

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., ... & Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. Science, 313(5793), 1626-1628.

Carhart-Harris, R. L., & Nutt, D. J. (2017). Serotonin and brain function: a tale of two receptors. Journal of Psychopharmacology, 31(9), 1091-1120.

Carlisle, N. B., Arita, J. T., Pardo, D., & Woodman, G. F. (2011). Attentional templates in visual working memory. Journal of Neuroscience, 31(25), 9315-9322.

Carrasco, M. (2011). Visual attention: The past 25 years. Vision Research, 51(13), 1484-1525.

Case, L. K., Wilson, R. C., & Ramachandran, V. S. (2012). Diminished size-weight illusion in anorexia nervosa: Evidence for visuo-proprioceptive integration deficit. Experimental Brain Research, 217(1), 79–87.

Casjens, S., Eckert, T., Woitalla, D., Ellrichmann, G., & Tönges, L. (2013). Impaired perception of biological motion in Parkinson’s disease. "European Journal of Neuroscience," 37(12), 1-7.

Caspers, S., et al. (2010). The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. NeuroImage, 50(4), 1517–1530.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. NeuroImage, 50(3), 1148-1167.

Chatterjee, A. (2011). Neuroaesthetics: A coming of age story. Journal of Cognitive Neuroscience, 23(1), 53-62.

Chatterjee, A., & Vartanian, O. (2014). Neuroaesthetics. Trends in Cognitive Sciences, 18(7), 370-375.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. Annual Review of Psychology, 62, 73-101.

Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1485), 1585-1599.

Cisek, P., & Kalaska, J. F. (2004). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. Neuron, 45(5), 801–814.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. Annual review of neuroscience, 33, 269-298.

Clark, A. (1997). Being there: Putting brain, body, and world together again. MIT press.

Clark, A. (1999). An embodied cognitive science?. Trends in cognitive sciences, 3(9), 345-351.

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

Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7-19.

Claussi, A., Karunakaran, K. K., et al. (2019). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. Journal of Neurophysiology, 121(4), 1133-1145.

Cohen, M. X. (2014). Analyzing neural time series data: Theory and practice. MIT Press.

Colling, L. J., Mulligan, D., et al. (2016). The neural signature of expertise in action video game players. Journal of Cognitive Neuroscience, 28(3), 330-342.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. Current Biology, 14(19), R850-R852.

Cook, J., Saygin, A. P., Swain, R., Blakemore, S. J., & Press, C. (2009). Reduced sensitivity to minimum-jerk biological motion in autism spectrum conditions. Neuropsychologia, 47(14), 3275-3278.

Cook, R., Bird, G., Catmur, C., Press, C., & Heyes, C. (2014). Mirror neurons: from origin to function. Behavioral and brain sciences, 37(2), 177-192.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews Neuroscience, 3(3), 201-215.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2000). Neural systems for visual orienting and their relationships to spatial working memory. Journal of Cognitive Neuroscience, 12(3), 508-523.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences, 24(1), 87-114.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? Current Directions in Psychological Science, 19(1), 51-57.

Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. Nature Reviews Neuroscience, 3(8), 655-666.

Craig, A. D. (2009). How do you feel–now? The anterior insula and human awareness. Nature Reviews Neuroscience, 10(1), 59-70.

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: A motor-visual attentional effect. Journal of Experimental Psychology: Human Perception and Performance, 25(6), 1673-1692.

Cross ES, Kraemer DJ, Hamilton AFC, Kelley WM, Grafton ST. 2009. Sensitivity of the action observation network to physical and observational learning. Cerebral Cortex 19:315-326.

Cross, E. S., Hamilton, A. F. C., et al. (2011). The influence of aesthetic appreciation on emotionally related judgments. Emotion, 11(4), 965-976.

Csibra, G. (2007). Action mirroring and action understanding: An alternative account. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), Joint Attention: Communication and Other Minds (pp. 139-157). Oxford University Press.

Curtis, C. E., & D’Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Sciences, 7(9), 415-423.

Curtis, C. E., & D’Esposito, M. (2004). The effects of prefrontal lesions on working memory performance and theory. Cognitive, Affective, & Behavioral Neuroscience, 4(4), 528-539.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. Annual Review of Psychology, 66, 115–142.

Damasio, A. (1994). Descartes' Error: Emotion, Reason, and the Human Brain. Putnam Publishing.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19(4), 450-466.

Daum, M. M., Adam, R., et al. (2016). The time course of visuo-motor neural coupling reveals the impact of vision on motor performance in virtual reality. NeuroImage, 126, 155-162.

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. Psychology of Music, 40(5), 595-633.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of Personality and Social Psychology, 44(1), 113-126.

De Beukelaar, T. T., Jacquet, P. O., et al. (2015). Single-trial EEG-fMRI reveals how brain performs predefined motor sequences under adverse conditions. NeuroImage, 105, 151-160.

De Brujin, O., Aglioti, S. M., et al. (2007). Predicting the outcome of the actions of others: A neural model of the perception of human actions. Psychological Research, 71(5), 525-541.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. Trends in cognitive sciences, 10(5), 204-211.

Dennett, D. C. (1993). The embodied mind: Cognitive science and human experience.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. Annual review of neuroscience, 18(1), 193-222.

Dessing, J. C., & Craig, C. M. (2010). Bimanual visuomotor coordination: What can we learn from studies in expert and preadolescent players? In M. Raab, J. G. Johnson, & H. Heekeren (Eds.), Progress in Brain Research (Vol. 174, pp. 61-74). Elsevier.

Di Noto, P. M., Newman, L., Wall, S., & Einstein, G. (2013). The hermunculus: What is known about the representation of the female body in the brain? Cerebral Cortex, 23(5), 1005-1013.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. Experimental Brain Research, 91(1), 176-180.

Dien, J. (2010). The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. Journal of Neuroscience Methods, 187(1), 138-145.

Dima, D., Corradi-Dell'Acqua, C., Sweeney-Reed, C. M., & Beste, C. (2022). Social-affective features predict action similarity judgments better than visual and action-related characteristics. Journal of Neuroscience, 42(1), 138-150.

Dima, D., Perry, G., Singh, K. D., & Friston, K. J. (2022). Understanding Action Prediction: A Conceptual Review and Theoretical Framework. Psychological Bulletin, 148(1), 30-52.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., ... & Petersen, S. E. (2007). Distinct brain networks for adaptive and stable task control in humans. Proceedings of the National Academy of Sciences, 104(26), 11073-11078.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2001). The effect of task relevance on the cortical response to changes in visual and auditory stimuli: An event-related fMRI study. NeuroImage, 14(6), 1256-1267.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. Science, 293(5539), 2470-2473.

Doyon, J., Bellec, P., Amsel, R., Penhune, V., Monchi, O., Carrier, J., ... & Benali, H. (2009). Contributions of the basal ganglia and functionally related brain structures to motor learning. Behavioural brain research, 199(1), 61-75.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. Psychological Review, 96(3), 433-458.

Ehrsson, H. H. (2012). The concept of body ownership and its relation to multisensory integration. In M. Murray & M. T. Wallace (Eds.), The Neural Bases of Multisensory Processes (pp. 775-788). CRC Press/Taylor & Francis.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), 124-129.

Endress, A. D., & Wood, J. N. (2011). From movements to actions: Two mechanisms for learning action sequences. Cognitive Psychology, 63(3), 141–171.

Ericsson, K. A. (2014). The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games. Psychology Press.

Ericsson, K. A., & Pool, R. (2016). Peak: Secrets from the new science of expertise. Houghton Mifflin Harcourt.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. Psychological Review, 100(3), 363-406.

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

Ferri, S., Rizzolatti, G., & Orban, G. A. (2015). The organization of the posterior parietal cortex devoted to upper limb actions: An fMRI study. Human Brain Mapping, 36(10), 3845–3866.

Ferrier, D. (1876). Experiments on the brain of monkeys. The Journal of Physiology, 3(3), 232-256.

Feusner, J. D., Hembacher, E., Moller, H., & Moody, T. D. (2010). Abnormalities of object visual processing in body dysmorphic disorder. Psychological Medicine, 40(12), 2011-2021.

Fiehler, K., Burke, M., Bien, S., Röder, B., & Rösler, F. (2008). The human dorsal action control system develops in the absence of vision. Cerebral Cortex, 19(1), 1–12.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. Journal of Experimental Psychology: Human Perception and Performance, 18(4), 1030-1044.

Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. Science, 218(4571), 486-487.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. Journal of Memory and Language, 29(2), 181-200.

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science, 291(5508), 1560-1563.

Friston, K. J., Sengupta, B., & Auletta, G. (2014). Cognitive dynamics: From attractors to active inference. Proceedings of the IEEE, 102(4), 427-445.

Friston, K., & Frith, C. (2015). A Duet for One. "Consciousness and Cognition," 36, 390-405.

Frith, C. (2007). Making up the mind: How the brain creates our mental world. John Wiley & Sons.

Fuster, J. M. (2001). The prefrontal cortex—an update: Time is of the essence. Neuron, 30(2), 319-333.

Gallagher, S. (2005). How the body shapes the mind. Oxford University Press.

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

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. Cognitive Neuropsychology, 22(3-4), 455-479.

Gallese, V., & Sinigaglia, C. (2011). What is so special about embodied simulation?. Trends in cognitive sciences, 15(11), 512-519.

Galvez-Pol, A., Calvo-Merino, B., Capilla, A., & Forster, B. (2018a). Persistent recruitment of somatosensory cortex during active maintenance of hand images in working memory. NeuroImage, 174, 153–163.

Galvez-Pol, A., Calvo-Merino, B., Capilla, A., & Forster, B. (2020). Persistent recruitment of somatosensory cortex during active maintenance of hand images in working memory. NeuroImage, 206, 116309.

Galvez-Pol, A., Calvo-Merino, B., Forster, B., 2020a. Revealing the body in the brain: an ERP method to examine sensorimotor activity during visual perception of body-related information. Cortex 125, 332–344.

Galvez-Pol, A., Forster, B., & Calvo-Merino, B. (2018b). Modulation of motor cortex activity in a visual working memory task of hand images. Neuropsychologia, 117, 75-83.

Galvez-Pol, A., Forster, B., Calvo-Merino, B., 2020b. Beyond action observation: Neurobehavioral mechanisms of memory for visually perceived bodies and actions. Neuroscience and Biobehavioral Reviews, 116, 508-518.

Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. Trends in Cognitive Sciences, 16(2), 129-135.

Gazzaniga, M. S. (2005). The ethical brain: The science of our moral dilemmas. Harper Perennial.

Gehring, W. J., Bryck, R. L., Jonides, J., Albin, R. L., & Badre, D. (2003). The mind's eye, looking inward? In search of executive control in internal attention shifting. Psychophysiology, 40(4), 572-585.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. Perception & Psychophysics, 67(7), 1252-1268.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. Science, 317(5843), 1351-1354.

Glenberg, A. M. (1997). What memory is for. Behavioral and Brain Sciences, 20(1), 1-19.

Glenberg, A. M., & Gallese, V. (2012). Action-based language: A theory of language acquisition, comprehension, and production. Cortex, 48(7), 905-922.

Gonzalez-Gadea, M. L., Chennu, S., Bekinschtein, T. A., et al. (2015). Predictive coding in autism spectrum disorder and attention deficit hyperactivity disorder. Journal of Neurophysiology, 113(12), 1989-2002.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. Trends in neurosciences, 15(1), 20-25.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. Nature, 349(6305), 154-156.

Goodale, M. A., Westwood, D. A., & Milner, A. D. (2004). Two distinct modes of control for object-directed action. Progress in brain research, 144, 131-144.

Gottlieb, J., Balan, P., Oristaglio, J., & Schneider, D. (2009). Task-specific computations in attentional maps. Vision Research, 49(10), 1216-1226.

Grafton, S. T. (2010). The cognitive neuroscience of prehension: Recent developments. Experimental Brain Research, 204(4), 475-491.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. Electroencephalogrography and Clinical Neurophysiology, 55(4), 468-484. https://doi.org/10.1016/0013-4694(83)90135-9

Gray, J. A., & McNaughton, N. (2000). The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system. Oxford University Press.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci., 31, 359-387.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. Annual Review of Neuroscience, 31, 359-387.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. Journal of Cognitive Neuroscience, 15(8), 1176-1194.

Grossman, E. D., Battelli, L., & Pascual-Leone, A. (2005). Repetitive TMS over posterior STS disrupts perception of biological motion. Vision Research, 45(22), 2847-2853.

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. Journal of Cognitive Neuroscience, 12(5), 711-720.

Gutteling, T. P., Kensinger, E. A., et al. (2013). A causal role for the right superior frontal gyrus in action-based response inhibition. Journal of Neuroscience, 33(5), 1986-1998.

Hagura, N., Kanai, R., et al. (2017). Action changes predictions about sensory outcomes. Journal of Neuroscience, 37(18), 4740-4751.

Hamilton, A. F. D. C. (2013). Reflecting on the mirror neuron system in autism: a systematic review of current theories. Developmental cognitive neuroscience, 3, 91-105.

Hanna, J. L. (1987). Dance and stress: Resistance to stress and perceptions of its effects. Dance Research Journal, 19(2), 49-60.

Harris, J. A., Miniussi, C., Harris, I. M., & Diamond, M. E. (2002). Transient storage of a tactile memory trace in primary somatosensory cortex. Journal of Neuroscience, 22(19), 8720-8725.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. Nature, 458(7238), 632–635.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General, 108(3), 356-388.

Hebb, D. O. (1949). The organization of behavior: A neuropsychological theory. John Wiley & Sons.

Heberlein, A. S., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. Journal of Cognitive Neuroscience, 16(7), 1143-1158.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. Proceedings of the National Academy of Sciences, 95(3), 781-787.

Hickok G. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. J Cogn Neurosci. 2009 Jul;21(7):1229-43. doi: 10.1162/jocn.2009.21189. PMID: 19199415; PMCID: PMC2773693.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. Behavioral and brain sciences, 24(5), 849-878.

Hooker, D. (1960), Plans and the structure of behavior. By George A. Miller, Eugene Galanter and Karl H. Pribram 1960. Henry Holt and company, New York. 226 pp. J. Comp. Neurol., 115: 217-217. https://doi.org/10.1002/cne.901150208

Hudson, M., Nicholson, T., & Bach, P. (2016). A direct examination of the effect of action intentions on movement kinematics. Cognition, 155, 32-42.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., et al. (2005). Grasping the intentions of others with one's own mirror neuron system. PLOS Biology, 3(3), e79.

Iacoboni, M., Woods, R. P., Brass, M., et al. (2001). Cortical mechanisms of human imitation. Science, 286(5449), 2526-2528.

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., PuÈtz, B., ... & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. Nature, 403(6766), 192-195.

Irimia, A., Erhart, M. J., & Brown, T. T. (2012). Variational Bayesian Independent Component Analysis for MEG/EEG Source Reconstruction. NeuroImage, 63(1), 387-403.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. Nature Reviews Neuroscience, 2(3), 194-203.

Jackson, R. C., Mori, S., & Shimada, S. (2006). The time–movement coupling function in the coordination of interception: Human movement timing as a function of ball speed. Experimental Brain Research, 175(2), 372-384.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. Journal of Experimental Psychology: General, 110(3), 306-340.

Jacquet, P. O., Springer, A., et al. (2016). The impact of expertise on brain activation of the action observation network during anticipation of tennis and volleyball serves. Frontiers in Human Neuroscience, 10, 568.

James, W. (1890). The Principles Of Psychology Volume II By William James (1890).

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Perception & Psychophysics, 14(2), 201-211.

Kaeppler, A. L. (1978). Dance in anthropological perspective. Annual Review of Anthropology, 7(1), 31-49.

Kahneman, D. (1973). Attention and effort. Prentice-Hall.

Kahneman, D. (2011). Thinking, Fast and Slow. Farrar, Straus and Giroux.

Kaiser, M. D., & Shiffrar, M. (2009). The visual perception of motion by observers with autism spectrum disorders: A review and synthesis. Psychonomic Bulletin & Review, 16(5), 761-777.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nature Neuroscience, 8(5), 679–685.

Kanakogi, Y., & Itakura, S. (2011). Developmental correspondence between action prediction and motor ability in early infancy. Nature Communications, 2, 341.

Kanwisher N. (2010) Functional specificity in the human brain: a window into the functional architecture of the mind. Proc. Natl Acad. Sci, 107, 11163-11170.

Karakose-Akbiyik, S., Caramazza, A. & Wurm, M.F. A shared neural code for the physics of actions and object events. Nat Commun 14, 3316 (2023). https://doi.org/10.1038/s41467-023-39062-8

Katja Fiehler, Michael Burke, Annerose Engel, Siegfried Bien, Frank Rösler, Kinesthetic Working Memory and Action Control within the Dorsal Stream (2008), Cerebral Cortex, Volume 18, Issue 2, Pages 243–253, https://doi.org/10.1093/cercor/bhm071

Katsuki, F., & Constantinidis, C. (2012b). Bottom-up and top-down attention: Different processes and overlapping neural systems. The Neuroscientist, 18(5), 436-443.

Katsuki, F., & Constantinidis, C. (2013). Early involvement of prefrontal cortex in visual bottom-up attention. Nature Neuroscience, 16(3), 291-297.

Katus, T., & Eimer, M. (2015). Lateralized delay period activity marks the focus of spatial attention in working memory: Evidence from somatosensory event-related brain potentials. Journal of Neuroscience, 35(17), 6689-6695.

Kaufman, J. C., & Mann, L. (2019). Understanding and promoting creativity in the arts. In J. C. Kaufman & R. J. Sternberg (Eds.), The Cambridge Handbook of Creativity (pp. 465-481). Cambridge University Press.

Kawato, M. (1999). Internal models for motor control and trajectory planning. Current opinion in neurobiology, 9(6), 718-727.

Keller, P. E. (2012). Mental imagery in music performance: Underlying mechanisms and potential benefits. Annals of the New York Academy of Sciences, 1252(1), 206-213.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. Journal of Verbal Learning and Verbal Behavior, 1(3), 153-161.

Keysers, C., & Gazzola, V. (2014). Dissociating the ability and propensity for empathy. Trends in cognitive sciences, 18(4), 163-166.

Kilner, J. M. (2011). More than one pathway to action understanding. Trends in Cognitive Sciences, 15(8), 352-357.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. Cognitive Processing, 8(3), 159-166.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). The mirror-neuron system: A Bayesian perspective. Neuroreport, 18(6), 619-623.

Kilner, J. M., Williams, A. M., et al. (2004). Modulating the human mirror system by attentional focus. Experimental Brain Research, 154(4), 385-390.

Kiyonaga, A., & Egner, T. (2013). Working memory as internal attention: Toward an integrative account of internal and external selection processes. Psychonomic Bulletin & Review, 20(2), 228-242.

Klaus, A., Alves da Silva, J., & Costa, R. M. (2019). What, if, and when to move: basal ganglia circuits and self-paced action initiation. Annual review of neuroscience, 42, 459-483.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. Human Neurobiology, 4(4), 219-227.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. Nature Reviews Neuroscience, 2(9), 635-642.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. Journal of Experimental Psychology: General, 133(3), 339-354.

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. Science, 317(5841), 1096-1099.

Lepsien, J., & Nobre, A. C. (2007). Attentional modulation of object representations in working memory. Cerebral Cortex, 17(9), 2072-2083.

Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. Behavioral and Brain Sciences, 8(4), 529-539.

Lingnau A, Gesierich B, Caramazza A. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. Proc Natl Acad Sci U S A. 2009 Jun 16;106(24):9925-30. doi: 10.1073/pnas.0902262106. Epub 2009 Jun 2. PMID: 19497880; PMCID: PMC2701024.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. Nature, 390(6657), 279-281.

Luria, R., Balaban, H., Awh, E., & Vogel, E. K. (2016). The contralateral delay activity as a neural measure of visual working memory. Neuroscience and Biobehavioral Reviews, 62, 100-108.

Ma, Q., Pu, M., Haihambo, N. P., Baetens, K., Heleven, E., Deroost, N., ... & Van Overwalle, F. (2022). The posterior cerebellum and temporoparietal junction support explicit learning of social belief sequences. Cognitive, Affective, & Behavioral Neuroscience, 1-25.

Macmillan, N. A., & Creelman, C. D. (2004). Detection Theory: A User's Guide. Psychology Press.

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. Proceedings of the National Academy of Sciences, 97(8), 4398-4403.

Makris, S., & Urgesi, C. (2015). Neural underpinnings of superior action prediction abilities in soccer players. Social Cognitive and Affective Neuroscience, 10(3), 342-351.

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. Current Biology, 14(9), 744–751.

Mason, A., Prodoehl, J., et al. (2020). Cerebellar lobules and dentate nuclei mirror cortical force–time profiles: Evidence for feedback control of human muscle co-contraction. Journal of Neurophysiology, 124(3), 838-851.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. Nature Neuroscience, 11(1), 103-107.

Meyer K, Kaplan JT, Essex R, Damasio H, Damasio A. 2011. Seeing touch is correlated with content-specific activity in primary somatosensory cortex. Cereb. Cortex. 21, 2113-2121.

Mikulan EP, Reynaldo L, Ibáñez A. Homuncular mirrors: misunderstanding causality in embodied cognition. Front Hum Neurosci. 2014 May 13;8:299. doi: 10.3389/fnhum.2014.00299. PMID: 24860479; PMCID: PMC4026706.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. Annual Review of Neuroscience,

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). Plans and the structure of behavior. Holt, Rinehart and Winston.

Milner, A. D., & Goodale, M. A. (1995). The visual brain in action. Oxford University Press

Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., ... & Davidson, D. L. W. (1991). Perception and action in 'visual form agnosia'. Brain, 114(1), 405-428.

Milner, B. (1962). Laterality effects in audition. In The Brain and Human Behavior (pp. 133-146). Springer.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In The human brain (pp. 373-394). Springer.

Minsky, M. (1988). The society of mind. Simon and Schuster.

Mishkin, M., & Ungerleider, L. G. (1982). Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. Behavioural brain research, 6(1), 57-77.

Moher, J., & Song, J. H. (2014). Perceptual decision processes flexibly adapt to avoid change-of-mind motor costs. Journal of Neurophysiology, 112(10), 2682-2695.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 125 human fMRI studies. Neuroscience & biobehavioral reviews, 36(1), 341-349.

Molenberghs, P., Cunnington, R., & Mattingley, J. B. (2012). Is the mirror neuron system involved in imitation? A short review and meta-analysis. Neuroscience & Biobehavioral Reviews, 36(3), 896-904.

Molenberghs, P., et al. (2012). The role of anterior parietal cortex in hand–object interactions. Frontiers in Psychology, 3, 1–6.

Mori, S., & Shimada, S. (2013). Expert anticipation from deceptive action. International Journal of Sport Psychology, 44(4), 356-367.

Mulligan, D., Colling, L. J., et al. (2016). Expert action video game players outperform novices in an action anticipation task. Journal of Experimental Psychology: Human Perception and Performance, 42(5), 743-753.

Murakami, S., Okada, Y., & Yagyu, K. (2002). Reduction of magnetoencephalographic interference in the electroencephalogram during simultaneous recording made inside a magnetically shielded room. Electroencephalography and Clinical Neurophysiology, 102(5), 569-576.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. Nature Reviews Neuroscience, 9(11), 856-869.

Nadal, M., & Pearce, M. T. (2011). The Copenhagen Neuroaesthetics Conference: Prospects and pitfalls for an emerging field. Brain and Cognition, 76(2), 172-183.

Narhi-Martinez, K., Moore, T., & Reynolds, J. H. (2022). The neurophysiology of attention. Journal of Neuroscience, 42(1), 53-64.

Narhi-Martinez, W., Dube, B., & Golomb, J. D. (2023). Attention as a multi-level system of weights and balances. Wiley Interdisciplinary Reviews: Cognitive Science, 14(1), e1633.

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

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. Personality and Social Psychology Review, 9(3), 184-211.

Niedermeyer, E., & da Silva, F. L. (2004). Electroencephalography: Basic principles, clinical applications, and related fields. Lippincott Williams & Wilkins.

Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., & Mesulam, M. M. (2004). Orienting attention to locations in perceptual versus mental representations. Journal of Cognitive Neuroscience, 16(3), 363-373.

Nobre, A. C., Griffin, I. C., & Rao, A. (2004). Spatial attention can bias search in visual short-term memory. Frontiers in Human Neuroscience, 1, 4.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation (Vol. 4, pp. 1-18). Springer.

Nowak, D. A., Hermsdörfer, J., et al. (2007b). The planning and control of prehension movements in the ipsilateral and contralateral hand. Cerebral Cortex, 17(3), 639-646.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. Behavioral and Brain Sciences, 24(5), 939-1031.

Oberauer, K. (2009). Design for a working memory. Psychology of Learning and Motivation, 51, 45-100.

Oberauer, K. (2019). The fourth element of working memory: A critical review of the evidence. Psychological Bulletin, 145(10), 1127-1153.

Oberauer, K., & Hein, L. (2012). Attention to information in working memory. Current Directions in Psychological Science, 21(3), 164-169.

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2015). Modeling working memory: An interference model of complex span. Psychonomic Bulletin & Review, 22(1), 32-48.

Olivers, C. N. L., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. Trends in Cognitive Sciences, 15(7), 327-334.

Orgs, G., Caspersen, D., & Haggard, P. (2016). You move, I watch, it matters: Aesthetic communication in dance.

Orgs, G., Dovern, A., Hagura, N., Haggard, P., Fink, G. R., & Weiss, P. H. (2016). Constructing visual perception of body movement with the motor cortex. Cerebral Cortex, 26(1), 440-449.

Orgs, G., Hagura, N., & Haggard, P. (2013). Learning to like it: aesthetic perception of bodies, movements and choreographic structure. Consciousness and Cognition, 22(2), 603-612.

Orgs, G., Kirsch, L., & Haggard, P. (2013). Time perception during apparent biological motion reflects subjective speed of movement, not objective rate of visual stimulation. Experimental brain research, 227, 223-229.

Overwalle, F., & Mariën, P. (2020). Dynamic causal modeling for ERP and fMRI: A new approach to understanding the neural control of complex motor behavior with tool use. NeuroImage, 211, 116607.

Palmer, C. (1997). Music performance. Annual Review of Psychology, 48(1), 115-138.

Palmer, C. E., Auksztulewicz, R., Ondobaka, S., & Kilner, J. M. (2020). Sensorimotor beta power reflects the precision-weighting afforded to sensory prediction errors. NeuroImage, 220, 117067.

Panasiti, M. S., Amoruso, L., et al. (2016). Developmental improvements in motor execution and interference control: Evidence from developmental and behavioral studies. Journal of Cognitive Neuroscience, 28(6), 907-922.

Parr, T., & Friston, K. J. (2018). The anatomy of inference: Generative models and brain structure. "Frontiers in Computational Neuroscience," 12, 90.

Passingham, R. E., Toni, I., & Rushworth, M. F. (2002). Specialisation within the prefrontal cortex: the ventral prefrontal cortex and associative learning. Experimental Brain Research, 146(4), 472-480.

Paull, G. C., & Glencross, D. J. (1997). Expert perception and decision making in baseball. International Journal of Sport Psychology, 28(3), 267-287.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain, 60(4), 389-443.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain, 60(4), 389-443.

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2011). The mechanics of embodiment: A dialog on embodiment and computational modeling. Frontiers in Psychology, 2, 5.

Pezzulo, G., Barsalou, L. W., Cangelosi, A., Fischer, M. H., McRae, K., & Spivey, M. J. (2012). Computational grounded cognition: a new alliance between grounded cognition and computational modeling. Frontiers in psychology, 3, 612.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson Jr, R., ... & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. Psychophysiology, 37(2), 127-152.

Pitcher D, CharlesL, Devlin JT, WalshV, Duchaine B. 2009. Triple dissociation of faces, bodies, and objects in extrastriate cortex. Curr. Biol. 19, 319-324.

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a third visual pathway specialized for social perception. Trends in Cognitive Sciences, 25(2), 100-110.

Pitcher, D., Charles, L., et al. (2009). Neural correlates of high-level expertise in skilled object recognition. Brain and Cognition, 69(2), 306-315.

Posner, M. I. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32(1), 3-25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In Attention and performance X: Control of language processes (pp. 531-556). Hillsdale, NJ: Erlbaum.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. Neuroscience, 139(1), 23-38.

Prinz, W. (1997). Perception and action planning. European journal of cognitive psychology, 9(2), 129-154.

Prodoehl, J., Weiss, P. H., et al. (2009). Increased brain activation during dual-task processing in Parkinson's disease: An fMRI study. Journal of Cognitive Neuroscience, 21(5), 1035-1046.

Proffitt, D. R. (2006). Embodied perception and the economy of action. Perspectives on Psychological Science, 1(2), 110-122.

Proske, U., & Gandevia, S. C. (2012). The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. Physiological Reviews, 92(4), 1651-1697.

Proteau, L. (1992). On the specificity of learning and the role of visual information for movement control. In M. Jeannerod (Ed.), Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience (pp. 507-530). MIT Press.

Pulvermüller, F. (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends in cognitive sciences, 17(9), 458-470.

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs: The D.O. Hebb lecture. Brain, 121(9), 1603-1630.

Rammami, A., & Miall, R. C. (2004). Musical expertise and the relative timing of stimulus occurrence in perception and action. Perception & Psychophysics, 66(3), 545-561.

Ranganathan, R., & Carlton, L. G. (2007). Perception-action coupling and anticipatory performance in baseball batting. Journal of Motor Behavior, 39(5), 369-380.

Reber, R., Schwarz, N., & Winkielman, P. (1998). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? Personality and Social Psychology Bulletin, 24(8), 1031-1042.

Redick, T. S., & Engle, R. W. (2006). Working memory capacity and attention network test performance. Applied Cognitive Psychology, 20(5), 713-721.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. Psychological Science, 14(4), 302–308.

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.

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. Journal of Experimental Psychology: General, 137(4), 615-625.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. Annual Review of Neuroscience, 27, 169–192.

Rizzolatti, G., & Sinigaglia, C. (2008). Mirrors in the brain: How our minds share actions and emotions. Oxford University Press.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. Nature reviews neuroscience, 11(4), 264-274.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. Cognitive Brain Research, 3(2), 131-141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. Nature reviews neuroscience, 2(9), 661-670.

Robbins, P., & Aydede, M. (Eds.). (2009). The Cambridge handbook of situated cognition. Cambridge University Press.

Rose, G. (2022). Visual methodologies: An introduction to researching with visual materials. Visual methodologies, 1-100.

Rossi, S., Hallett, M., Rossini, P. M., & Pascual-Leone, A. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clinical Neurophysiology, 120(12), 2008-2039.

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., & Marsden, C. D. (1982). Manual motor performance in a deafferented man. Brain, 105(3), 515-542.

Rugg, M. D., & Coles, M. G. (1995). Electrophysiology of mind: Event-related brain potentials and cognition. Oxford University Press.

Saalmann, Y. B., Pigarev, I. N., & Vidyasagar, T. R. (2007). Neural mechanisms of visual attention: How top-down feedback highlights relevant locations. Science, 316(5831), 1612-1615.

Sack, A. T. (2006). Transcranial magnetic stimulation, causal structure–function mapping and networks of functional relevance. Current Opinion in Neurology, 19(6), 593-599.

Saito, T., Motoki, K., Nouchi, R., & Sugiura, M. (2023). Facilitating animacy perception by manipulating stimuli exposure time. Frontiers in Psychology, 13, 1017685.

Saygin, A. P., & Stadler, W. (2012). The role of appearance and motion in action prediction. Psychological Research, 76(4), 388-394.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. Psychological review, 82(4), 225.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. Trends in Cognitive Sciences, 11(5), 211–218.

Schultz, J., & Frith, C. D. (2022). Animacy and the prediction of behaviour. Neuroscience & Biobehavioral Reviews, 104766.

Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. Physiological Reviews, 95(3), 853-951.

Schwartz, B. (2004). The paradox of choice: Why more is less. HarperCollins.

Searle, J. R. (1992). The Rediscovery of the Mind. MIT Press.

Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. Psychonomic Bulletin & Review, 16(1), 170-175.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), Handbook of bilingualism: Psycholinguistic approaches (pp. 371-388). Oxford University Press.

Sel, A., Forster, B., & Calvo-Merino, B. 2014 The Emotional Homunculus: ERP Evidence for Independent Somatosensory Responses during Facial Emotional Processing. J. Neurosci. 34, 3263-3267.

Serences, J. T., & Kastner, S. (2014). A multi-level account of selective attention. In The Oxford Handbook of Attention (pp. 76-104). Oxford: Oxford University Press.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. Trends in Cognitive Sciences, 10(1), 38-45.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. Psychological Science, 20(2), 207-214.

Sevdalis, V., & Keller, P. E. (2011). Perceiving performer identity and intended expression intensity in point-light displays of dance. Psychological Research, 75(5), 399-409.

Shadmehr, R., & Krakauer, J. W. (2008). A computational neuroanatomy for motor control. Experimental Brain Research, 185(3), 359-381.

Shen, J., et al. (2014). Amodal completion in visual working memory. Journal of Experimental Psychology: Human Perception and Performance, 40(4), 1513–1525.

Shen, K., Humphreys, G. W., & Shui, R. (2014). Memory for grasping is influenced by the perception of illusory size. Acta Psychologica, 145, 46-55.

Sherrington, C. S. (1907). On the proprio-ceptive system, especially in its reflex aspect. Brain, 29(4), 467-482.

Shomstein, S. (2012). Cognitive functions of the posterior parietal cortex: Top-down and bottom-up attentional control. Frontiers in Integrative Neuroscience, 6, 38.

Shomstein, S., & Behrmann, M. (2008). Object-based attention: Strength of object representation and attentional guidance. Perception & Psychophysics, 70(1), 132-144.

Shomstein, S., & Behrmann, M. (2010). Object-based attention: Strength of object representation and attentional guidance. Perception & Psychophysics, 72(5), 1327-1343.

Siciliano, L., Binkofski, F., & Wittmann, M. (2023). Motor resonance and the prediction of observed actions in the human cerebellum. NeuroImage, 294, 118278.

Siebner, H. R., Hartwigsen, G., Kassuba, T., & Rothwell, J. C. (2009). How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. Cortex, 45(9), 1035-1042.

Siman-Tov, T., Bosak, N., & Sprecher, E. (2019). The Brain as a Prediction Machine: An Integrative Review. Current Psychiatry Reports, 21(12), 136.

Smyth, M. M., & Pendleton, L. R. (1990, 1989). Working memory for movements. The Quarterly Journal of Experimental Psychology, 42(2), 295–316.

Smyth, M. M., et al. (1988). Interference in immediate spatial memory. Memory & Cognition, 16(1), 1–13.

Sokolov, A. A., Miall, R. C., & Ivry, R. B. (2017). The cerebellum: adaptive prediction for movement and cognition. Trends in cognitive sciences, 21(5), 313-332.

Stadler, W., Ott, D. V., Springer, A., Schubotz, R. I., Schütz-Bosbach, S., & Prinz, W. (2012). Repetitive TMS suggests a role of the human dorsal premotor cortex in action prediction. Frontiers in Human Neuroscience, 6, 20.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. Behavior Research Methods, Instruments, & Computers, 31(1), 137-149.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. Nature Reviews Neuroscience, 9(4), 255-266.

Tamè, L., & Longo, M. R. (2023). Emerging principles in functional representations of touch. Nature Reviews Psychology, 1-13.

Tarhan, L., De Freitas, J., & Konkle, T. (2021). Behavioral and neural representations en route to intuitive action understanding. Neuropsychologia, 163, 108048.

Taylor, K., Mandon, S., Freiwald, W. A., & Kreiter, A. K. (2005). Coherent oscillatory activity in monkey area V4 predicts successful allocation of attention. Cerebral Cortex, 15(9), 1424-1437.

Tenke, C. E., & Kayser, J. (2012). Surface Laplacians and phase–amplitude coupling. Journal of Clinical Neurophysiology, 29(3), 162-169.

Theeuwes, J. (1992). Perceptual selectivity for color and form. Perception & Psychophysics, 51(6), 599-606.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. Quarterly Journal of Experimental Psychology, 37A, 571-590.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. Nature, 428(6984), 751-754.

Toga, A. W., & Mazziotta, J. C. (Eds.). (2002). Brain Mapping: The Methods. Academic Press.

Tononi, G., & Koch, C. (2008). The neural correlates of consciousness: an update. Annals of the New York Academy of Sciences, 1124, 239-261.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12(1), 97-136.

Treue, S. (2003). Visual attention: the where, what, how and why of saliency. Current Opinion in Neurobiology, 13(4), 428-432.

Troje, N. F., & Chang, D. H. (2023). Life detection from biological motion. Current Directions in Psychological Science, 32(1), 26-32.

Tsubomi H, Fukuda K, Watanabe K, Vogel EK. 2013. Neural limits to representing objects still within view. J. Neurosci, 33, 8257-8263.

Tsubomi, H., Fukuda, K., et al. (2014). Sustained visuo-spatial attention reveals object-based representations in visual working memory. Neuropsychologia, 55, 105-113.

Tunik, E., Frey, S. H., et al. (2007). Doing the right thing: Brain correlates of feeling in control of our actions. Journal of Cognitive Neuroscience, 19(9), 1535-1548.

Umiltà, M. A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., & Rizzolatti, G. (2001). I know what you are doing: A neurophysiological study. Neuron, 32(1), 91–101.

Urgesi C, Calvo-Merino B, Haggard P, Aglioti SM. 2007a. Transcranial magnetic stimulation reveals two cortical pathways for visual body processing. The Journal of Neuroscience 27:8023-8030.

Urgesi CM, Candidi M, Ionta S, Aglioti SM. 2007b. Representation of body identity and body actions in extrastriate body area and ventral premotor cortex. Nature Neuroscience 10(1): 30–31.

Urgesi, C., Aglioti, S. M., Skrap, M., & Fabbro, F. (2012). The spiritual brain: selective cortical lesions modulate human self-transcendence. Neuron, 65(3), 309-319.

Urgesi, C., Candidi, M., & Avenanti, A. (2014). Neuroanatomical substrates of action perception and understanding: An anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. Frontiers in Human Neuroscience, 8, 344.

Urgesi, C., Candidi, M., Avenanti, A. (2010). Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. Frontiers in Human Neuroscience, 8, 344.

Urgesi, C., Moro, V., Candidi, M., & Aglioti, S. M. (2006). Mapping implied body actions in the human motor system. Journal of Neuroscience, 26(30), 7942–7949.

Vallar, G., & Ronchi, R. (2009). Somatoparaphrenia: A body delusion. A review of the neuropsychological literature. Experimental Brain Research, 192(3), 533-551.

van Berkum, J. J. A., Brown, C. M., Hagoort, P., et al. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. Journal of Memory and Language, 41(2), 147-182.

van Overwalle, F., Van de Steen, F., et al. (2022). The role of the cerebellum in predicting social events: A pilot study. Social Neuroscience, 17(1), 56-67.

Varela, F. J., Thompson, E., & Rosch, E. (1991). The embodied mind: Cognitive science and human experience. MIT press Cambridge, MA.

Verfaillie, K., & Daems, A. (2002). Representations of vision and knowledge about pictures and reality in visual memory: Investigations with resemble and identity matching tasks. Visual Cognition, 9(5), 659-700.

Versace, R., Vallet, G. T., Riou, B., Lesourd, M., Labeye, É., & Brunel, L. (2014). Act-In: An integrated view of memory mechanisms. Journal of Cognitive Psychology, 26(3), 280-306.

Vesia, M., & Crawford, J. D. (2012). Specialization of reach function in human posterior parietal cortex. Experimental Brain Research, 221(1), 1–18.

Vicary, S. A., Robbins, R. A., Calvo-Merino, B., & Stevens, C. J. (2014). Recognition of dance-like actions: Memory for static posture or dynamic movement? Memory & Cognition, 42(5), 755–767.

Vogel, E. K., & Luck, S. J. (2004). The visual N1 component as an index of a discrimination process. Psychophysiology, 41(2), 208–215.

Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. Nature, 428(6984), 748-751.

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. Nature, 438(7067), 500–503.

Wadsworth, H. M., & Kana, R. K. (2011). Brain mechanisms of perceiving tools and imagining tool use acts: A functional MRI study. Neuropsychologia, 49(7), 1863–1869.

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. Neuron, 59(6), 1037-1050.

Whitford, T. J., Jack, B. N., Pearson, D., Griffiths, O., Luque, D., Harris, A. W., ... & Ford, J. M. (2017). Neurophysiological evidence of efference copies to inner speech. eLife, 6, e28197.

Whittlesea, B. W., Jacoby, L. L., & Girard, K. (1990). Illusions of immediate memory: Evidence of an attributional basis for feelings of familiarity and perceptual quality. Journal of Memory and Language, 29(6), 716-732.

Wilson, M. (2002). Six views of embodied cognition. Psychonomic Bulletin & Review, 9(4), 625-636.

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation increases positive affect. Journal of Personality and Social Psychology, 81(6), 989-1000.

Witt, J. K. (2011). Action's effect on perception. Current Directions in Psychological Science, 20(3), 201-206.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. Psychonomic Bulletin & Review, 1(2), 202-238.

Wolfe, J. M. (2010). Visual search. In S. Martinez-Conde, S. L. Macknik, & S. Martinez-Conde (Eds.), Oxford Compendium of Visual Illusions (pp. 34-38). Oxford University Press.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? Nature Reviews Neuroscience, 5(6), 495-501.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. Current Biology, 11(18), R729-R732.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. Nature neuroscience, 3(11), 1212-1217.

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. Neural networks, 11(7-8), 1317-1329.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358(1431), 593-602.

Wolpert, D. M., Goodbody, S. J., & Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. Nature Neuroscience, 1(6), 529-533.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. Trends in cognitive sciences, 2(9), 338-347.

Woody, C. D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. Medical & Biological Engineering, 5(5), 539-554.

Wulf, G., & Prinz, W. (2001). Directing attention to movement effects enhances learning: A review. Psychonomic Bulletin & Review, 8(4), 648-660.

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: A review of influential factors. Medical Education, 44(1), 75-84.

Wurm, M. F., & Schubotz, R. I. (2012). Sustaining the interpretation of ambiguous scenes: The interplay of scene semantics, visual exploration, and semantic memory. Frontiers in Psychology, 3, 525.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. Vision Research, 48(20), 2070-2089.

Yang, J., Kanai, R., & Rees, G. (2015). An invisible hand: Investigating the proprioceptive control of the "rubber hand" illusion. PLOS ONE, 10(2), e0117159.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. Journal of Experimental Psychology: Human Perception and Performance, 10(5), 601-621.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. Nature Neuroscience, 5(10), 995-1002.

Zacks, J. M. (2008). Neuroimaging studies of mental rotation: A meta-analysis and review. Journal of Cognitive Neuroscience, 20(1), 1–19.

Zeki, S. (1991). Cerebral akinetopsia (visual motion blindness). Brain, 114(2), 811-824.

Zhao, Z., Piazza, E. A., Zhan, Y., & Gao, T. (2021). Human action recognition in point-light displays: A computational investigation using recurrent neural networks. PLoS Computational Biology, 17(9), e1009495.

Zwaan, R. A. (2014). Embodiment and language comprehension: Reframing the discussion. Trends in cognitive sciences, 18(5), 229-234.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: motor resonance in language comprehension. Journal of Experimental Psychology: General, 135(1), 1.