Spatial encoding of reach action target: An investigation into its body-part centric multi-sensory integration mechanism

A thesis submitted in fulfilment of the requirements

for the degree of Master of Science (MS) by Research

by

Rasika Lokhande

20128405

DEPARTMENT OF COGNITIVE SCIENCE

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

May 2023

Certificate

It is certified that the work contained in this thesis entitled **"Spatial encoding of reach action target: An investigation into its body-part centric multi-sensory integration mechanism "** by **Rasika Lokhande** has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

Devpriya Kumar

Prof. Devpriya Kumar

Associate Professor

CGS Department

Indian Institute of Technology Kanpur

May 2023

Declaration

This is to certify that the thesis titled **"Spatial encoding of reach action target: An investigation into its body-part centric multi-sensory integration mechanism "** has been authored by me. It presents the research conducted by me under the supervision of **Prof. Devpriya Kumar**.

To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted elsewhere, in part or in full, for a degree. Further, due credit has been attributed to the relevant state-of-the-art and collaborations with appropriate citations and acknowledgments, in line with established norms and practices.

Rasika

Rasika Lokhande

Roll No. 20128405

CGS Department

Indian Institute of Technology Kanpur

Abstract

Name of the student: **Rasika Lokhande**   Roll No: **20128405**

Degree for which submitted: **MSR**   Department: **CGS Department**

Thesis title: **Spatial encoding of reach action target: An investigation into its body-part centric multi-sensory integration mechanism**

Thesis supervisor: **Prof. Devpriya Kumar**

Month and year of thesis submission: **May 2023**

One of the fundamental human abilities involved in human-environment interaction is extending one's arm to make physical contact with an object. To perform a successful reaching action, the agent's cognitive system must encode the spatial location of the target. Peri-personal space is the region of space surrounding the agent's body where reach targets are located. One of the characteristics of peri-personal space is that the objects in this region of space are encoded with respect to the region of the body-surface that is in close proximity to the object. In addition, action-relevant body parameters, such as the position and structure of the limbs, are encoded in a body representation which is constructed by integrating visual and proprioceptive sensory information about the body.

Considering the above, this thesis investigates the following two questions. Firstly, is the spatial information of the reach target encoded relative to the body part located proximally, even if that body part is not the effector performing the reach action? Secondly, if the answer to the first question is affirmative, how does the integration of visuo-proprioceptive signals regarding the proximal body part contribute to the spatial encoding of the target? To investigate these questions, two experiments were conducted in an immersive virtual reality set-up, in which a reach task was performed while the action hand was rendered

invisible. Visual and proprioceptive information of the proximal body part was manipulated as the experimental conditions, and the accuracy of reaches was measured. The results of the two experiments indicate that the spatial encoding of the reach target occurs relative to the proximal body-part, and that the integration of the proximal body-part's visual and proprioceptive sensory signals underlies the spatial encoding of the reach target's location. However, the results further suggest that the spatial encoding of the target occurs relative to the body-part location indicated by its visual signal, as opposed to the body-part position inferred after the integration of multi-sensory signals.

Acknowledgements

Contents

List of Figures

List of Tables

Dedicated to my Guru, Dr. Vishnu Ramachandra Parnerkar

Chapter 1

Introduction

You extend your arm to grab a cup of coffee, grasp the book you're holding, or swat a mosquito buzzing around your ears. One of the most common ways we interact with our surroundings is by reaching out with our hands to touch the objects in our surroundings. Reaching is one of the fundamental human actions involved in human-environment interaction, underpinning diverse abilities such as feeding, object manipulation, and self-defense. A characteristic feature of reaching action is that it is goal-directed towards a specific external target. Once the target is specified during the action selection stage, it must be translated into a movement plan so that the effector (the hand) can make contact with the target(eg. the coffee mug). To accomplish this, *the spatial location* of the coffee mug must be encoded in a format suitable for motor control. Information from various sensory modalities about the spatial location of the target have to be integrated with the spatial information of the effector to perform an accurate reach. How is the spatial information of the reach target encoded? In particular, what is the reference frame in which the reach target is spatially encoded?

The targets of reaching action are situated in the region of in close proximity to the body, which is termed in the literature as *peri-personal space*. The characteristic feature of representation of spatial information of objects in peri-personal space is that it is encoded with respect to the area of body which is in proximity to the object (Serino, 2019). This notion that the representation of spatial information of the object is "anchored" to the proximal body-part is supported by both neuro-physiological and behavioural evidence (See Chapter 2). For instance, visuo-tactile multi-sensory interaction effects such as the *cross-modal congruency effect* has been observed in certain experiments, wherein subjects localize tactile stimuli faster when a task-irrelevant visual cue is presented congruent to

that body part, as compared to the condition where the task irrelevant cue is presented incongruent to the tactile stimulated body part (Spence, Pavani, Maravita, & Holmes, 2004). The magnitude of cross-modal congruency effect (and other similar multi-sensory interaction effects) is also indicative of the distance between the stimuli and the body (Spence, Pavani, & Driver, 2004; Iachini, Ruggiero, Ruotolo, & Vinciguerra, 2014), that is, these multi-sensory responses increase as the distance between the visual stimuli and the body decreases.

These multi-sensory interaction effects are used in literature as a proxy for understanding how peri-personal space in implicated in various cognitive processes. For instance, a functional link has been demonstrated between action and multi-sensory behavioural responses evoked by objects in peri-personal space (Ladavas & Serino, 2008; Patané et al., 2019; Brozzoli, Cardinali, Pavani, & Farnè, 2010), suggesting that the peri-personal space representation serves as a perceptual-motor interface for guiding goal directed actions. Furthermore, the multi-sensory interaction effects are also modulated by action possibilities (Lohmann, Belardinelli, & Butz, 2019; Iriki, Tanaka, & Iwamura, 1996; Senna, Cardinali, Farnè, & Brozzoli, 2019). It has thus been theorized that a generative mechanism underlies representation of peri-personal space, which predicts the likelihood of future sensory outcomes (Lohmann et al., 2019; Noel et al., 2018). In this light, the representation of spatial information of objects in peri-personal space could be interpreted as a representation of space in terms of a constantly updated set of probabilities that carry information about the likelihood that stimuli in the space around the body will make contact with a specific area of the body. If this is indeed the case, modulating parameters that affect the representation of peri-personal space should affect the outcome of actions involving contact, that is, reaching actions. However, one of the gaps in the literature is that the link between goal directed action and peri-personal space- where the reach action targets are present, has only been investigated through the multi-sensory interaction effects paradigm, and not through the outcome of reach action- that is, how accurate were the subjects in estimating the location of the target, when factors affecting peri-personal space representations are modulated.

If the reach targets are spatially encoded with respect to body-part proximal to them, it is plausible that the representation of body parameters is also involved in the reach targets' spatial encoding. One of the views in the literature is that the cognitive system represents body parameters, particularly those necessary for motor control- including the spatial information of the effector, in an integrated representation termed as *body schema*

([Ataria, Tanaka, & Gallagher](), [2021]). Notably, the construction of body schema is commonly conceptualized as an inference problem, where most probable estimates of the body parameters (here, the spatial position of the body-part) are dynamically computed from the inflow of sensory inputs from various sensory modalities, like vision and proprioception ([Van Beers, Sittig, & Gon](), [1999]). In other words, a visuo-proprioceptive integrative mechanism underlies the specification of spatial information in body schema. Could the same visuo-proprioceptive integration mechanism which specifies the spatial information of the body-part, also underlie the spatial encoding of the reach target which lies in proximity to this body-part?

To investigate this question, two experiments were conducted in which a reaching action task paradigm was implemented in an immersive virtual reality environment. With their action hands rendered invisible, the subjects were asked to make contact with a visual target, while visual and proprioceptive information of the action-irrelevant body part proximal to the target was manipulated, and accuracy of target location estimated by the participant was measured. In the first experiment, we investigated whether the reach target location estimation is indeed affected by the presence or absence of a proximal body-part, and if so, how does the uni-sensory and multi-sensory information of the proximal body-part affect the reach target location estimation. In the second experiment, we attempted to probe further into how visuo-proprioceptive integration mechanism underlies the spatial encoding of reach target, by manipulating the spatial discrepancies between the visual and proprioceptive information of the proximal body-part. The results of the two experiments support the hypothesis that reach target is encoded with respect to the proximal body-part, and that its multi-sensory integration mechanism underlies the reach action target encoding. However, the results suggest that the spatial encoding of reach target occurs relative to the location specified by the visual information of the proximal body-part, as opposed to the location of the body-part inferred after the multi-sensory integration process.

Chapter 2

Literature review

One of the classic questions in the philosophy of cognition is - What is the relationship between perception and action? One of the dominant ideas is that the purpose of perception is to gather information for guiding action, while action is a means to acquire perceptual information regarding the environment. In reaching action, action is directed towards objects of interest in the external environment. The body is central to the process of reaching action. Not only is the reach carried out by bodily movements, but the targets of reaching actions also lie in space close to the body(peri-personal space). Thus, for a reaching action, the cognitive system of the agent requires i) knowledge of body for computing the state of the effector (the body entity with which the action is carried out), and ii) information regarding the object of interest (the target entity on which the action is carried upon). Perceptual processes provide knowledge to the system regarding both these entities.

The aim of this chapter is to provide relevant background and discuss key concepts, theories and empirical findings that led to the development of the conjecture that we have investigated in this thesis. This chapter will thus be divided into two parts - the body and the target. Part I consists of a brief discussion on body representation, and its role in motor control. In Part II, we delve into theories and findings regarding peri-personal space, where the reach action target lies, and its implications in unfolding of action. I will end with connecting these distinct strands of literature and stating the conjecture.

2.1 Body representation and Action

Goal-directed action require information about properties of one's body, in order to plan movements, and guide the action. These properties include both structural and positional parameters of the body. Precise position of one's limbs, not only during action initiation (t = 0), but also during the action execution, is an important parameter for a successful action. Furthermore, information regarding structural properties of one's body, such as body configuration, body size, flexibility of joints, muscle strength, are necessary so that movements do not occur in biologically impossible or painful manner. According to the review by Ataria et al. (2021), one of the claims in the literature is that the cognitive system of the agent represents the above mentioned properties in an integrated representation of the body, termed as *body schema*. Importantly, body schema is conceptually distinguished from other body representations like body image, by the notion that the parameters encoded in body schema are in a specific format. Evidence from Schwoebel and Coslett (2005)'s study suggests that body parameters are encoded in three distinct representational formats - sensorimotor, visuo-spatial, and linguistic. Body schema encodes the bodily parameters in a sensorimotor format, which the motor system can directly exploit.

In the control-theoretic model of action, two types of process models have been proposed- the inverse model and the forward model, both of which function parallelly. The claim is that body schema plays a role in both these processes. The inverse model computes the motor commands required to achieve the goal, or the final state of the action, constrained by the current body capacities, which are represented in the body schema. The forward model predicts sensorimotor consequences of the action (Wolpert & Flanagan, 2001). Here, the execution of action is guided by matching the anticipatory model of the body, with the current model of the body, as represented by the body schema. Ataria et al. (2021) thus proposes a dual function of body schema - description of the body at a spatial resolution and format which can be utilized by the motor system, and for the anticipatory control of action.

2.1.1 Construction of body schema

One of the dominant questions regarding the body schema is - what are the computations underlying the construction of body schema. Ataria et al. (2021) suggests that construction of body schema is a dynamic process, that is, the body representation is continuously updated to adapt to the bodily changes throughout development and other short term changes like changes in environment, ill-health, pregnancy, or tool use.

One of the theories for understanding the construction of body schema is that the generation of body schema can be thought of as an inference problem, to compute the most reliable estimates of body parameters from the inflow of sensory inputs from various sensory modalities, like vision and proprioception (Van Beers et al., 1999). The Bayesian model of multi-sensory integration proposes that the cognitive system computes the most probable output, here, the parameters specifying the body schema, from the sensory cues available, on the basis of the reliability of the sensory cues (noise), and the prior beliefs encoded in the system regarding the relevance of the cues. In a novel bodily situation, online sensory cues will be favoured, versus the beliefs regarding the past body representations. Furthermore, the success of action also informs the construction of the body schema. If the action is successful, the process underlying the generation of body schema is confirmed, whereas, if the action is unsuccessful, the body schema needs to be updated for future actions. Thus, Bayesian view provides a description of how body schema is dynamically constructed.

Furthermore, the role of integration of inputs from multiple modalities in demonstrating the plasticity of the body representation has also been empirically demonstrated through the Rubber Hand Illusion. In this paradigm, when a rubber hand is stroked synchronously with a participant's unseen real hand, they tend to i) perceive the location of their hand as displaced towards the rubber hand (Tsakiris & Haggard, 2005) , ii) perceive tactile sensation through the rubber hand (Pavani, Spence, & Driver, 2000), and iii) experience a feeling of ownership over the rubber hand (Longo, Schüür, Kammers, Tsakiris, & Haggard, 2008).

There are two major challenges regarding the computations of the multi-sensory integration processes underlying the emergence of body schema. The first is how to identify and select which are the relevant sensory signals to bind together- a problem termed as the "parsing problem". The second problem is, once the relevant signals are selected, how are the signals integrated into a coherent body representation? Every sensory signal is encoded in a different modality-specific format, and in a particular frame of reference, which needs to be resolved into a common format and reference frame. Furthermore, different sensory cues are assigned different weights based on their reliability and relevance. How are these weights assigned, is an important question to consider.

Summarizing this section, body schema specifies bodily parameters which are relevant for action planning and control, and the construction of body schema is a dynamic process
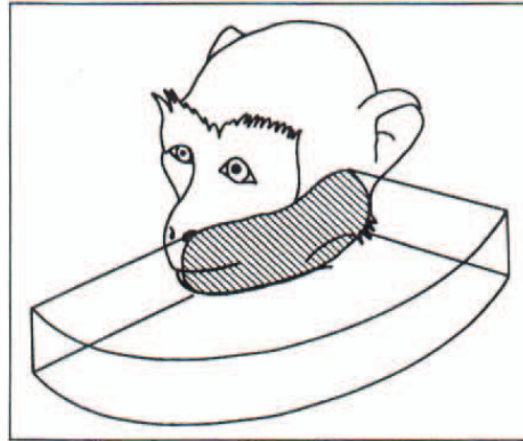
Figure 2.1: Spatially corresponding tactile receptive field (striped) and visual receptive field (boxed) of a bimodal neuron in putamen graziano1994mapping

relying on a process of multi-sensory integration of sensory inputs from vision and proprioception. We will now turn towards a discussion regarding perception of the action target and the representation of its spatial location.

2.2 Peri-personal space

In the cognitive system, the representation of space provides a framework regarding how different objects in the external environment are located with respect to each other, and with respect to the self. Other than abstract and geographical spatial representations, the prominent characteristic of space representation is that body, or a body-part is the origin of spatial representations. Spatial information is processed in several reference frames, depending on the sensory source of the information and the agent environment interaction. For example, the spatial location of an approaching bee will be processed not only in eye-centered and head-centered reference frames due to visual and auditory inputs, but also in reference to the hand to which it is approaching to avoid getting strung (Serino, 2019). However, exponentially, the agent does not have explicit access to these distinct representations, suggesting integration of different reference frames into a unified global reference frame (as reviewed by Serino (2019)).

Peri-personal space is the region of space in close proximity to the body. Its representation is distinct from the representations of extra-personal space (regions of space far from the body), based on findings from lesion studies in monkeys that show that different neuronal populations encode peri-personal and extra-personal space (Rizzolatti, Matelli, & Pavesi, 1983).

2.2.1 Objects in peri-personal space are represented with respect to the body-part in proximity

The spatial location of objects in peri-personal space is represented by an underlying multi-sensory mechanism. This notion is supported by both neurophysiological studies in macaque monkeys and behavioural studies in humans. Results from primate studies show evidence of bimodal neurons in several regions of the brain - premotor cortex, putamen, and parietal area 7B (Graziano & Gross, 1994). These neurons have visual and tactile receptive fields which spatially correspond with each other (refer Figure 2.1), and respond to both visual and tactile stimuli, when they are present in their respective receptive fields. Furthermore, evidence from Avillac, Hamed, and Duhamel (2007) is also indicative of neural computation of multi-sensory integration, where evoked multi-sensory responses are either sub-additive or super-additive than than sum of uni-modal responses. This dual receptive field property of these bimodal neurons is not just indicative of the multi-sensory mechanisms underlying peri-personal space representation, but also that these bimodal neurons integrate the visual signal with the location of the body near which the visual stimulus appears, suggesting that the representation of the spatial location of the objects is anchored to the surface of the body in the object's proximity.

The body-part centered representation of peri-personal space is also supported by multi-sensory interaction effects demonstrated in human behavioural studies. Results from cross-modal congruency effect experiments show that participants localize tactile stimulation faster when a task-irrelevant visual cue is presented congruent to that body part, as compared to the condition where the task irrelevant cue is presented incongruent to the tactile stimulated body part (Spence, Pavani, Maravita, & Holmes, 2004). Similarly, participants are quicker in responding to tactile stimulation where task-irrelevant visual cue is present, compared to the condition where visual cue is absent (Patané et al., 2019). Furthermore, these multi-sensory effects are dependent on the distance of the visual stimuli to the body, as evidenced by the finding that the magnitude of multi-sensory interactions is greater in regions of space near the body as compared to far away from the body (Spence, Pavani, & Driver, 2004), and processing times of stimuli in peri-personal space is faster as compared to extra-personal space (Iachini et al., 2014). Consequently, these effects are used as measures for peri-personal space and as paradigms to understand the effect of several factors on the peri-personal space representation (Bufacchi & Iannetti, 2021).

2.2.2 Action and Peri-personal Space

The nature of neural and behavioural multi-sensory responses as discussed above indicate the body-part specific encoding of objects in peri-personal space, with the differences in their magnitude being indicative of the distance between the object and the body part. However, Bufacchi and Iannetti (2021) proposes that these multi-sensory measures not just encode the spatial information about the object but also behavioural relevance of a stimuli for a set of possible actions. The action field theory of peri-personal space thus proposes that peri-personal space measures are reflective of values of actions that aim to make or avoid contact between the body and the object, and since distance is a necessary parameter for contact related actions, the distance dependent nature of peri-personal space may be a mere complementary characteristic of these measures.

This notion that the representation of peri-personal space serves as a perceptual-motor interface for guiding goal directed actions is supported by several empirical studies which aimed to establish a functional link between actions and multi-sensory interactions occurring in peri-personal space (Ladavas & Serino, 2008). The unfolding of goal-directed actions seem to trigger dynamic and online changes in visuo-tactile interactions in peri-personal space as the action unfolds. Specifically, behavioural measures of peri-personal space are distinct in different stages of action, suggesting an interaction between the representation of objects in space and the action performed on those objects. Even before the action is initiated, Patané et al. (2019) found evidence that that visuo-tactile interactions were most strongly affected during the action planning phase, and further enhanced during the action execution phase. Moreover, different kinds of actions, like grasping and pointing, showed distinct modulations in peri-personal space measures (Brozzoli et al., 2010).

Theories of anticipatory behaviour control propose that actions are initiated by predicting sensory consequences. In line with the claim of the Action Field Theory (Bufacchi & Iannetti, 2021), several studies suggest that peri-personal space is modulated by action possibilities, that is, the extent of peri-personal space, as operationalized by visuo-tactile interaction measures, was extended to action-relevant spatial locations (Lohmann et al., 2019; Iriki et al., 1996). Furthermore, plausibility of goal-oriented actions further seem to increase visuo-tactile interactions during the action execution stage (Senna et al., 2019). Lohmann et al. (2019) and Noel et al. (2018) have reasoned that the a generative mechanism underlies modulation of peri-personal space, which predicts the likelihood of future sensory outcomes. In other words, the cognitive system represents peri-personal space as

a representation of space surrounding the body in terms of a set of probabilities that are constantly updated, which carry information about the likelihood of stimuli in the space around the body making a contact with a particular area of the body.

2.3   Gap in the literature

One of the gaps in the literature is that the involvement of peri-personal space representation in action planning and execution is investigated by how different stages of action modulate the multi-sensory interaction effects. However, the visual stimuli, whose spatial properties interact with the tactile stimuli, are not relevant to the action. We reason that the implication of peri-personal space representation in action is also due to the fact that reach action targets lie in peri-personal space, and therefore, characteristics of peri-personal space representation, like body-part centric representation of objects, should also be applicable to the reach action targets spatial encoding and therefore reflective in the way target's location is estimated by the agent. However, this has not been investigated before, and thus serves as a gap which we seek to investigate in the following study. We reasoned that if multi-sensory integration mechanisms underlie the construction of body-schema, and if reach targets are represented with respect to a proximal body-part, manipulating visual and proprioceptive information should affect the spatial encoding of the target, and would be reflective in its location estimation. The effect of such multi-sensory discrepancies on reach target location estimation has been investigated before (eg. Noel et al., 2018), but the visuo-proprioceptive discrepancy has been induced in the body-part which serves as the effector, that is, the action hand. However, the effect of visuo-proprioceptive discrepancies in body-part proximal to body part has not been studied before, which is a gap that we seek to explore in this thesis.

Chapter 3

Uni-sensory and multi-sensory information of proximal body-part

3.1 Introduction

Performing a goal-directed reaching action requires the specification of the spatial location of the action target. The targets of reaching action are situated in peri-personal space - the region of space proximally surrounding the body. Evidence from several empirical studies suggest that the spatial information of the objects in peri-personal space is represented with respect to the region of the body-surface which is proximal to the object (For review, see Serino, 2019). Furthermore, the information regarding the body-parameters, such as its spatial information, is encoded as a multi-modal representation, constructed by integrating signals from the visual and proprioceptive modalities. Thus, if the spatial information of the reach action target is encoded with respect to the proximal body surface, then manipulations in the proximal body-part's visual and proprioceptive information and its integrative processes should affect the estimation of the spatial location of the target.

In that light, the aim of this experiment was to establish that sensory information regarding the proximal body part, even if it is irrelevant to the reaching task, affects the estimation of spatial location of the target. Specifically, we sought to understand how uni-sensory and multi-sensory inputs regarding the proximal (and task irrelevant) body part affects the accuracy of the target location estimation. If the spatial information of the target is indeed encoded with respect to the target-proximal body part, the conjecture is that the presence of multi-sensory visuo-proprioceptive inputs regarding the target-proximal body

part would lead to a relative increase in accuracy in target location estimation compared to the condition in which no sensory inputs of target-proximal body part is present. Moreover, when only uni-sensory (visual or proprioceptive) information was present, several outcomes were possible. The first possibility was that uni-sensory input increases relative accuracy in target location estimation, but not comparable to the multi-sensory input. If more weight in the encoding mechanism is given to vision, the relative increase in accuracy would be more in the visual uni-sensory input condition as compared to proprioceptive uni-sensory input condition, and vice-versa. Another possibility was that uni-sensory input would lead to no significant relative increase in target location estimation accuracy, which would suggest that multi-sensory integration mechanism underlies the encoding of spatial information of the target with respect to the proximal body-part.

To investigate these hypotheses, a reaching task paradigm was implemented in an immersive virtual reality set-up, in which subjects had to perform goal-directed reaching movements towards a target, while the non-action hand was placed proximal to the target. The virtual reality environment was capitalized to dissociate the visual and proprioceptive information provided to the subject regarding the non-action hand. The visual and proprioceptive information of the non-action hand was therefore manipulated, along with the target's distance from the non-action hand (and consequently, action hand). The estimated target location was measured as the response of the end-point of the reaching action.

3.2   Method

3.2.1   Participants

Twenty-three right handed participants (11 females, mean = 26.39, range = 21-34 years old) were recruited for participation in the experiment. All participants reported normal or corrected to normal vision, and provided formal written consent before participating in the experiment, in accordance to the protocol approved by the Institutional Ethics Committee (IEC). Two participants were excluded from the analysis. One was excluded due to a technical error that arose during the run of the experiment, while another participant was excluded on basis of the outlier criteria, described in the Experimental Design section below.
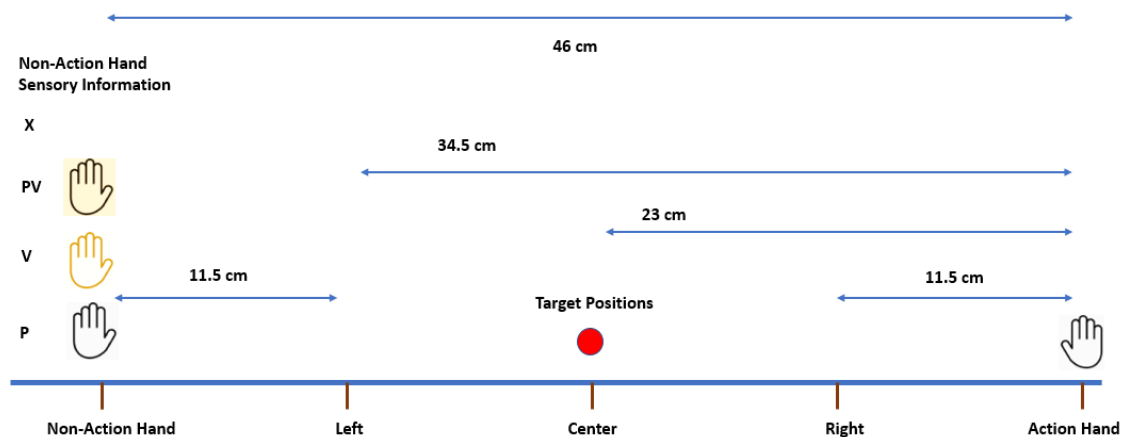
Figure 3.1: Experimental Setup

3.2.2 Materials and Apparatus

The participant sat at a table with surface dimensions = 120cm x 50cm. Two coin-shaped docks (2.5 cm diameter) were attached to the table with a distance of 46 cm between them at the center of the table. A cylindrical barrier of 2.5cm height was attached around the right dock. The participant sat with their arms resting comfortably on the table, with the right and left docks serving as resting positions for the right and left index fingers respectively. Hand motion and position was tracked with Ultraleap Leap Motion Controller.

The virtual reality environment was displayed on Oculus Rift S Head-mounted Display, and developed using Unity Game Engine. The virtual scene consisted of a virtual table situated in a black room. The virtual table was spatially aligned to the physical table at which the participant sat. Reach targets appeared at three locations on the surface of the virtual table, along the axis joining the two docks- at a distance of 11.5cm, 23cm, and 34.5cm from the left dock, acting at Left, Center, and Right Target Positions respectively (see Figure 3.1). The reach action target was a red dot of 0.5cm in diameter. Depending on the experimental condition, virtual hands of the participant was rendered in the virtual environment, using the hand models from the Unity module developed by Ultraleap.

3.2.3 Calibration

At the beginning of the experiment, the participants were given oral instructions to calibrate for spatially aligning the virtual table with the physical table. Five block dots, aligned on x axis, appeared on the table with a distance of 11.5cm between them. The

Block	Action hand Visibility	Non-Action hand Position	Non-Action hand Visibility
Training	Visible	Lap	-
Feedback	Invisible	Lap	-
X	Invisible	Lap	-
P	Invisible	Left dock	Invisible
V	Invisible	Lap	Visible (Model hand rendered)
PV	Invisible	Left dock	Visible

Table 3.1: Experiment Block structure

participant rested their right index finger on the right dock. Only the virtual hands were visible during the calibration process, whereas the physical docks were not visible in the virtual environment. Using the arrow keys on the keyboard, the participant was asked to move the rightmost black dot, such that the dot lies under their right index finger. Then, the table height was adjusted by the participant using the keypad enter and plus keys, such that the hand resting on the physical table also appeared to rest on the virtual table. The participant was then asked to rest their left index finger on the left dock, sequentially make contact with the three center dots with their right hand, and then rest the right hand on the right dock, during which the experimenter calibrated the positions of the hand at these five locations respectively.

3.2.4 Procedure

The experiment was divided into 6 blocks. Depending upon the block, the participant was instructed regarding the placement of the left hand - either with the left index finger resting on the left dock, or the left hand resting on the participants lap. Furthermore, depending upon the block of the experiment, the left and the right hand were rendered visible or invisible.

The participants initiated each trial by bringing their right index finger on the right dock. After an audio cue, the target is presented in one of the three target positions. On presentation of the visual target, the participant attempted to make contact with the visual target by making unspeeded reaching action with their right hand towards the target. On contact with the table, the target disappears, indicating the completion of the trial. After completion of the trial, the participant returns the right hand back to the right dock to initiate the next trial.

The first block of the experiment was the training phase, intended to acquaint the participant with the task and the virtual environment. In the training block, the right hand was rendered visible, while the left hand was placed on the participants lap. The next

	Effect	DFn	DFd	F	p	p<.05	pes
1	block	3.00	60.00	3.482	0.021	*	0.148
2	target_pos	1.14	22.89	10.013	0.003	*	0.334
3	block:target_pos	3.62	72.38	1.435	0.235		0.067

Table 3.2: Repeated measures ANOVA results with Greenhouse-Geisser Sphericity Corrections: Reach error ~ Target Position * Visuo-Proprioceptive Information + (1 | subject)

block was the Feedback block, in which the right hand was rendered invisible. After the participant indicated the position of the target by making contact with the table, a green dot appeared to indicate the location where the participant had made the contact with the table. This served as feedback for the participant, to compare their responses to the actual location of the target.

In the next four blocks, the right hand was rendered invisible as a methodological constraint to introduce uncertainty in the reaching action. The presence and absence of proprioceptive and visual information of the left hand served as experimental manipulations. Proprioceptive information of the the non-action hand was manipulated by asking the participant to rest their left hand on the left dock (Proprioceptive Information present condition) or to rest their left hand on their lap (Proprioceptive Information absent Condition). The Visual information of the non-action hand was manipulated by rendering it visible (Visual Information Present Condition) or invisible (Visual Information Absent Condition). See Table 3.1 for a comprehensive overview of the block design.

For all participants, the first and the second block were Training and Feedback respectively. The order of the remaining four blocks was counterbalanced across participants using the Latin Square Balancing Design. The three target positions were randomized within each block. All six blocks consisted of 60 trials each, with 20 trials for each of the 3 target positions within each block. The participant thus completed a total of 360 trials across six blocks.

3.2.5   Experimental Design

The experiment was set up in a within-subject 4 x 3 factorial design, with two independent variables - Non-Action Hand Information and Target Position. Accuracy in target location estimation was operationalized as *reach error*, which was computed by subtracting the actual target location from the estimated target location indicated by the participants response for each trial. Positive reach error indicates that the estimated target location
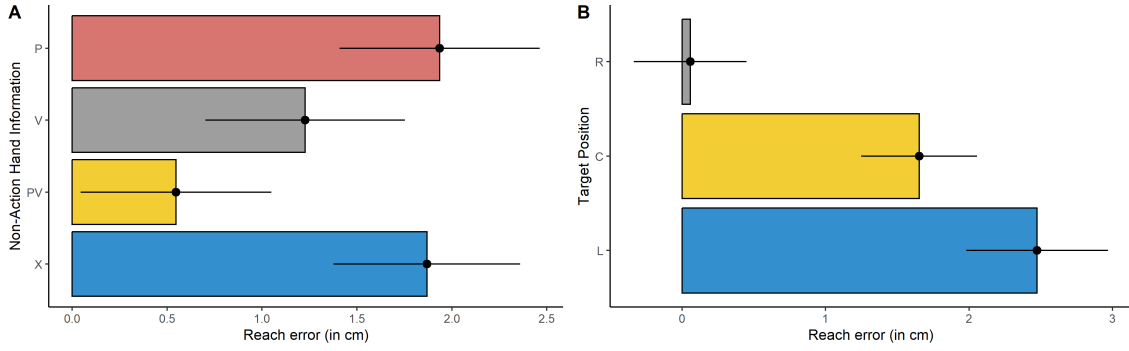
Figure 3.2: Effect of Non-Action hand information and Target Position on Reach error. The positive values of reach error indicate that the end point of the reach was in the direction of the action hand to the right of the target. The error bars indicate standard error of the mean.

was an underestimation towards the direction of the action hand, while negative reach error indicates that the estimated target location was an overestimation towards the direction of the non-action hand. Data was analyzed for the four blocks where Non-Action hand information was manipulated (X, P, V, PV). The experiment thus consisted of 12 conditions- 4 (Non-Action Hand Information conditions) x 3 (Target Position conditions). Subject-wise mean reach error was calculated for each of the 12 conditions by computing the average of mean error across trials for each condition, which served as the dependent variable.

The Outlier criteria was determined as any participant whose any of the mean reach error values fall above Q3 + 3 x IQR or below Q1 - 3 x IQR (Q3 = Third quartile, IQR = interquartile range). On the basis of this criteria, data from one participant was excluded from the analysis.

3.3 Results

A 4 (Non-Action Hand Information: X, V, P, PV) X 3 (Target Position: Left, Center, Right) repeated measures ANOVA was performed to evaluate the effect of the two factors on magnitude of reach error. The analysis demonstrated a significant main effect of both Non-Action Hand Visibility ($F(3, 60) = 3.482, p = 0.021, \eta_p^2 = 0.148$) and Target Position ($F(1.14, 22.89) = 10.013, p = 0.003, \eta_p^2 = 0.334$) (See Table 3.2). No significant interaction effect between the two factors was present($F(3.62, 72.38) = 1.435, p = 0.235, \eta_p^2 = 0.067$).

Post-hoc pair-wise comparisons for analyzing the differences in the reach error between each of the four Non-Action Hand Information condition, using Holm-Bonferroni corrected paired t-test, showed that the magnitude of reach error for X and PV conditions differed significantly ($t(20) = 2.9427216, p = 0.048, d = 0.642$). Pair-wise comparisons for analyzing differences in reach error between the three target position conditions showed that reach error magnitude differed significantly between Left and Right positions ($t(20) = 3.223430, p = 0.009, d = 0.703$), and also between Center and Right positions ($t(20) = 3.728681, p = 0.004, d = 0.813$).

The results of the experiment show that across all conditions, on average, there is an underestimation in the responses to the estimation of spatial location of the target, that is, the estimated target location is in the direction of the action hand relative to the actual target location. Furthermore, the subjects are more accurate (lower reach error magnitude) when both proprioceptive and visual information of the non-action hand is present (PV condition), as compared to condition in which both proprioceptive and visual information of the non-action hand is absent (X Condition). Moreover, the subjects are most accurate in the Right Target Position, lesser accurate in the Center Target Position, and least accurate in the Left Target Position.

3.4 Discussion

The results of the experiment show that the response of the reach action is affected by the position of the target, and the sensory information of the non-action body part which is proximal to the target. The position of the target can be interpreted in two ways- as the distance to be covered by the action hand during the unfolding of action, or as the proximity to the non-action hand. Since there is no interaction effect present between the non-action hand information and the target position, the target position may be interpreted as the distance between the action hand and the target. The decrease in accuracy as the distance between the target position and initial position of action hand suggests that as the distance covered in the reach increases, the greater underestimation is observed in situations of uncertainty, which in this case is the invisibility of action hand.

Regarding the effect of non-action hand information, subjects are more accurate in target location estimation when both visual and proprioceptive information of the non-effector hand is provided, even when the body-part parameters of this hand are not seemingly relevant. This result provides support to our conjecture that even task irrelevant body

part in proximity to the target affects the representation of the spatial information of the target, suggesting that the target of reach action is represented in an ego-centric frame of reference with respect to the part of body near it. Furthermore, the accuracy of estimating target location is not significantly improved when just uni-sensory information of the non-action hand is present. This suggests that atleast some level of integration between visual and proprioceptive sensory signals of the hand may be necessary for the ego-centric representation of target to occur efficiently.

However, if the spatial encoding of the the target is indeed anchored to the sensory information which specifies the position of the body-part proximal to the target, the expectation is that the proximity to the target should increase the accuracy of the spatial representation of the target. However, the results indicate the opposite. Thus overall, this suggests that there are two processes which additively influence the spatial encoding of the target-i) underestimation when the reaches are over longer distances, and ii) biasing the target position towards the non-action hand (thus increasing the accuracy) when the target's spatial information is anchored to it when its multi-sensory information is present.

Chapter 4

Visuo-proprioceptive spatial discrepancies in proximal body-part

4.1 Introduction

In the previous chapter, we discussed evidence supporting the hypothesis that multi-sensory visuo-proprioceptive integration mechanisms of proximal body-part plays a role in reach target's spatial encoding. But how exactly is the multi-sensory integration involved in the target's spatial encoding? Let us consider two hypothesis which may plausibly explain the implication of the integration mechanism in reach target's spatial encoding.

4.1.1 Spatial encoding with respect to the inferred position of proximal body-part

One of the mechanisms postulated regarding the specification of body's spatial information in the representation of body (that is, the body schema) is that of visuo-proprioceptive multi-sensory integration. Here, the functional role of multi-sensory integration is to compute reliable estimates of the location of the body part, by integrating the estimates provided by uni-sensory cues on the basis of its reliability and the relevance attributed to the particular modality (Limanowski & Blankenburg, 2016; Noel et al., 2018; Körding et al., 2007; Van Beers et al., 1999). In this light, the way multi-sensory integration could
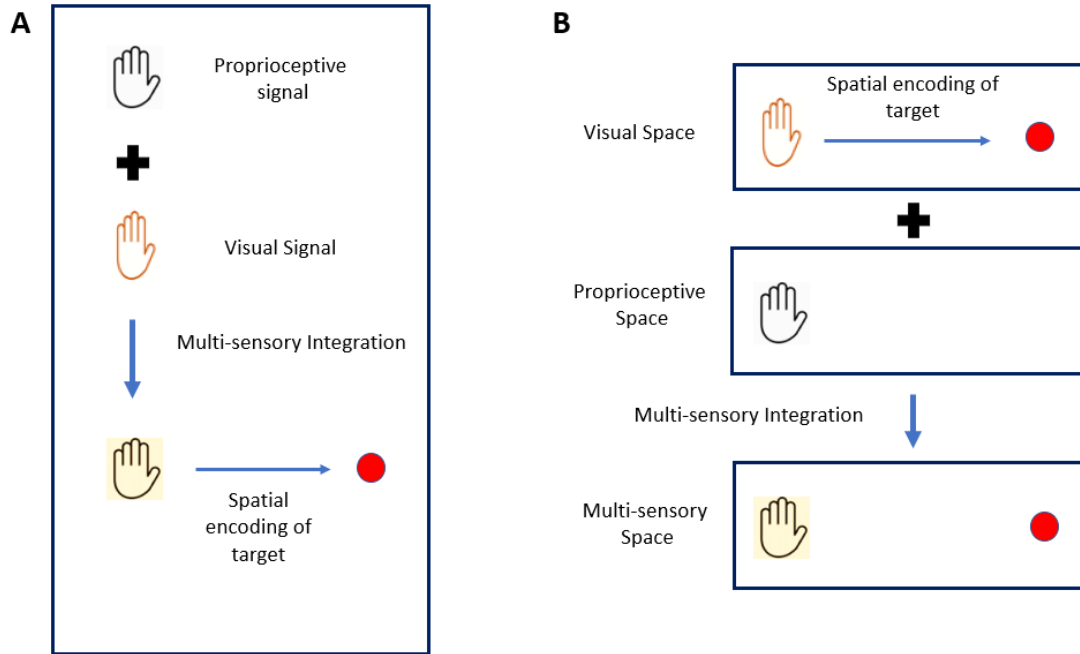
Figure 4.1: Two hypothesis regarding spatial encoding of the reach target. In A), the filled hand denotes the hand position inferred from the multi-sensory integration of visual and proprioceptive uni-sensory signals. The target's spatial location is encoded relative to this inferred hand position. In B) The target's spatial location is encoded relative to the visual signal in the visual space. The visual space is then integrated with the proprioceptive space to construct a multi-sensory representation of body-schema and surrounding peri-personal space.

be involved in spatial encoding of the reach target is as follows- multi-sensory integration results in the inference of a reliable estimate of proximal body-part's spatial location, and thereafter, the spatial location of the reach target is represented with respect to this inferred spatial location. A schematic of this conjecture is shown in Figure 4.1 (A).

4.1.2 Spatial encoding with respect to the visual signal of proximal body-part

An alternative hypothesis is that the spatial encoding of the target could occur with respect to the estimate of the body-part position provided by its visual signal. This visual space, containing a map of visual objects- body part and the target, is then integrated with the proprioceptive space containing a map of proprioceptive objects (see Figure 4.1 (B)). The plausibility of this hypothesis is supported by several empirical studies. For instance, the reach target is a visual object, and several studies have shown increases in accuracy of reaching, when visual "landmarks" are in proximity to the reach target (eg. Conti

& Beaubaton, 1980), suggesting the occurrence of allocentric encoding of reach target. Furthermore, several studies have also suggested that spatial encoding in allocentric reference frames is used in combination with egocentric representation (Byrne, Cappadocia, & Crawford, 2010; Schütz, Henriques, & Fiehler, 2013), although the literature here refers to eye-centered egocentric representation. Nevertheless, it is plausible that the spatial encoding of visual target is with respect to the visual information, and this visual space is then integrated with proprioceptive space, to construct a multi-sensory representation of space which is characteristic of peri-personal space representation.

4.2 Experiment 2

To test for these competing hypothesis, we conducted an experiment in which we manipulated the spatial discrepancy between the actual position of the non-action target proximal hand and its visual input, by spatially displacing the visually rendered hand in an Immersive Virtual Reality environment. The objective of this experiment was to understand how reach target location estimation was affected by the spatial discrepancy in the estimates provided by visual and proprioceptive inputs of the body-part (non-action hand) proximal to the target. We also aimed to qualitatively compare which of the mechanistic hypotheses described above explains the the reach accuracy results. If the inferred position of the proximal hand systematically affects the target location estimation, it would suggest that Hypothesis 1 may be the multi-sensory mechanism that could underlie body-proximal reach target spatial encoding. On the other hand, if the proximity to the visual hand systematically affects the target location estimation, Hypothesis 2 may be the mechanism underlying the reach target spatial encoding.

4.2.1 Method

4.2.1.1 Participants

Nineteen right-handed subjects(2 Females, Mean age = 22.68, Range = 18-33) were recruited to participate in the experiment. All participants reported normal or corrected to normal vision. They were naive to the purpose of the experiment and the experimental procedure, and had provided written consent to participate in the study according to the norms approved by the Institute Ethics Committee (IEC) of Indian Institute of Technology, Kanpur.

4.2.1.2   Materials and Apparatus

The participants sat across a table with surface area dimensions of 120cm x 50cm. Two coin-shaped docks (2.5 cm diameter) were attached to the table with a distance of 46 cm between them at the center of the table. A cylindrical barrier of 2.5cm height was attached around the right dock. The participant sat with their arms resting comfortably on the table, with the right and left docks serving as resting positions for the right and left index fingers respectively. Hand motion and position was tracked with Ultraleap Leap Motion Controller. The virtual reality environment was displayed on Oculus Rift S Head-mounted Display, and developed using Unity Game Engine. The virtual scene consisted of a virtual table situated in a black room. The virtual table was spatially aligned to the physical table at which the participant sat. The target of the reach was a red dot of 0.5cm in diameter.

4.2.1.3   Procedure

The experiment began with a calibration session, similar to the session performed in Experiment 1 (Refer Chapter 3). The experiment consisted of 5 Experimental blocks which were interleaved with filler blocks. For the experimental blocks, the participants were instructed to place their left hand on the dock, whereas for the filler blocks, the left hand was placed on the lap. The experimental blocks differed in the position at which the left hand was rendered visually in the virtual scene. The visual rendering of the left hand, henceforth termed as the visual hand, was positioned at five locations - congruent to the left hand placed at the left dock (henceforth termed as the proprioceptive hand), 5.75cm, and 11.50cm to the right of proprioceptive hand, and 5.75cm and 11.50cm to the left of the proprioceptive hand. The order of the experimental blocks was counterbalanced across participants using Latin Square counterbalancing design.

Within each block, the target appeared at one of the following 6 locations along the axis joining the two docks, at a distance of 28.75, 25.30, 21.85, 18.40, 14.95, and 11.50 cm from the action hand (see Figure 4.2). The experimental blocks each consisted of 10 x 6 (target positions) = 60 trials, while the filler blocks consisted of 5 trials each.

The task performed by the participant was same as that of experiment 1. Each trial was initiated by bringing the right index finger on the right dock. On an audio cue, the target appeared at one of the above mentioned target locations, after which the participant attempted to make contact with the target with their right index position. The trial ended with the disappearance of the target after the contact was detected between the table and
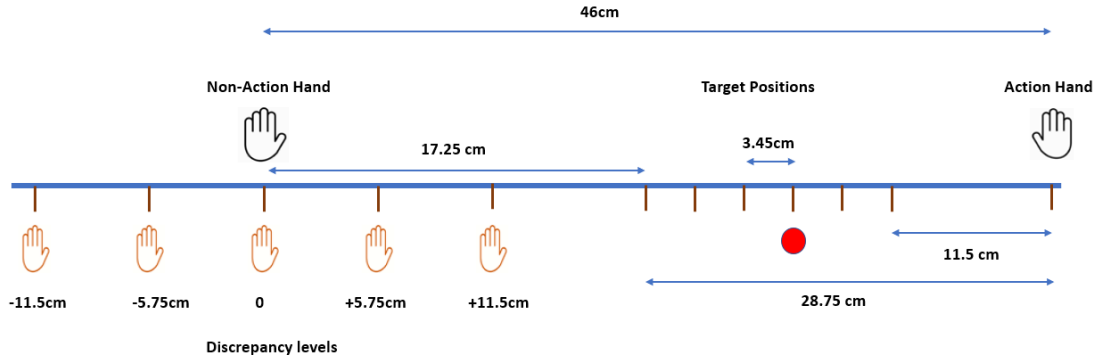
Figure 4.2: Experimental Setup

the right index finger. The participant brought the right hand back to the right dock to initiate the next trial. Throughout the course of the experiment, the right (action hand) was rendered invisible.

4.2.1.4 Experimental Design

The accuracy of the spatial encoding of the target in reach action was operationalized as *reach error*, which was computed as the difference between the actual location of the target and the estimated location of the target, indicated by the participant by the end-point of the reach action. Positive reach error indicates that the relative to the actual position of the target, the estimate of the target location is in the direction towards the action hand. On the other hand, negative reach error indicates that target location estimate is to the left of the actual target location in the direction of the the non-action hand.

The experiment was set-up in a within-subjects experimental design, with reach error as the dependent variable. The two independent variables that were experimentally manipulated were - i) The non-action hand visuo-proprioceptive spatial discrepancy, that is, the distance between the actual position of the hand (proprioceptive hand) and the visual rendering of the hand (visual hand), and ii) The distance between the initial position of the action hand and target (AHT).

Model	Fixed effect	AIC	BIC
Null Model	None	28746	28779
Discrepancy Model	Discrepancy	28731	28770
AHT Model	AHT	28698	28737
Additive Model	Discrepancy + AHT	28683	28729
Interaction Model	**Discrepancy + AHT + (Discrepancy * AHT)**	**28666**	**28719**

Table 4.1: AIC and BIC Values for fitted Linear Mixed Effect models. Random effects of all models consisted of intercepts for subjects and by-subject random slopes for effect of discrepancy ($1 + discrepancy|subject$). The interaction model has the lowest AIC and BIC value.

4.2.2 Data Analysis

To understand the effect of visuo-proprioceptive spatial discrepancy on the reach error, we performed a linear mixed effects analysis using the *lme4* (Bates, Mächler, Bolker, & Walker, 2014) and *lmerTest* packages in R. We considered visuo- proprioceptive spatial discrepancy and the distance from action hand (AHT) as fixed effects in our models. Random effects of all models that we fit consisted of intercepts for the subjects and by-subject random slopes for the effect of discrepancy. Table 4.1 states the fixed effects of the various fitted models and their Akaike information criterion (AIC) and Bayesian Information Criteria (BIC) values. Likelihood Ratio Test was performed to compare the models and as a criteria for model selection. Refer Appendix B for the results of the Likelihood Ratio test.

The AIC, BIC, and the results of the Likelihood Ratio Test reveal that the best fit model consisted of discrepancy and AHT as the fixed effects, along with an interaction between the two predictors (see Table **??**). There were no obvious deviations from homoscedasticity and normality on the basis of visual inspection of residual plots (See Figure B.2 and Figure B.1). Furthermore, we checked for multi-collinearity between the predictors by using Variance Inflation Factor (VIF) as a measure of collinearity using the package *performance* in R. The results show low correlation (VIF ¡ 2) between the fixed effect terms of the best-fit model (see Table B.7).

Moreover, we performed bootstrapping to get credible confidence intervals at 95% for the estimates of the fixed effects (see Table B.8). None of the confidence intervals for the fixed effects contain the value 0, which further supports the significance of the effects of the three predictors (discrepancy, AHT, and interaction) on the reach error.
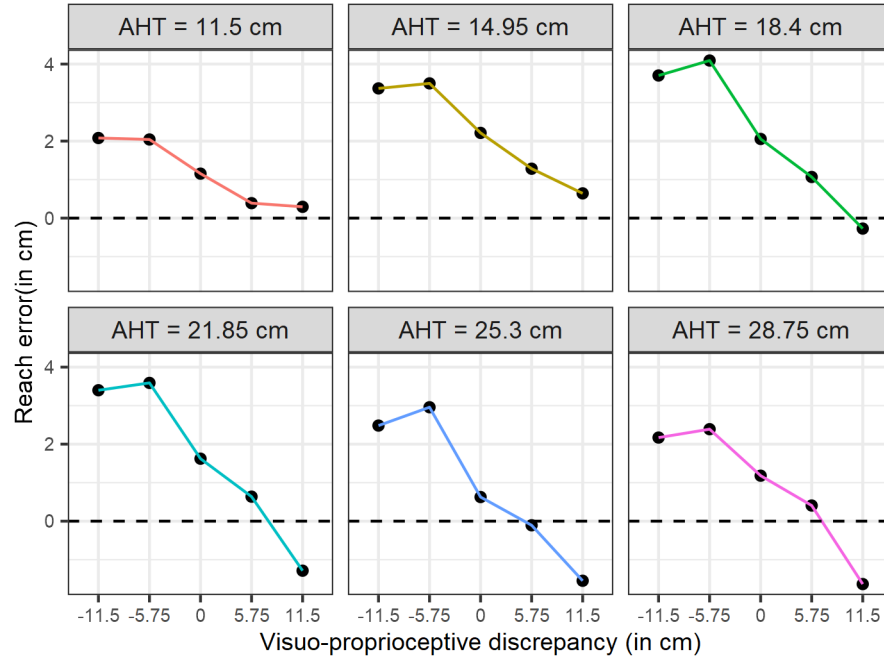
Figure 4.3: Reach error as a function of spatial discrepancy between the position of visual hand and the position of actual (proprioceptive) non-action hand. Each panel shows the distance between the target and the initial position of the action hand (AHT). Negative reach error indicates that the estimation of the target location was in the direction of the non-action hand towards the left, while positive reach error indicates that the location of the target was estimated in the direction of action hand towards the right of the actual location of the target.

Thus, the results of the Linear Mixed Effects analysis shows that the interaction between the spatial discrepancy between the visual and proprioceptive inputs of the proximal non-action hand and the distance between the target and the action hand has a statistically significant effect on reach error. This suggests that the way visuo-proprioceptive spatial discrepancy of proximal non-action hand affects the target estimation depends upon the distance of the target from the the action hand. Therefore, to interpret the results, we will consider the effect of visuo-proprioceptive discrepancy separately at different distances of target from the action hand.

4.2.3 Support for visual anchoring mechanism

Figure 4.3 shows the effect of visuo-proprioceptive spatial discrepancy of proximal non-action hand on reach error, at different reach distances, that is, distance between the target and the initial position of the action hand.

Since, the location of the proprioceptive signal is fixed according to the design of the experiment, the values of visuo-proprioceptive discrepancy also may be interpreted as different positions of the visual hand. The results in Figure 4.3 show that there is a general underestimation (reach error is positive), when the target is closer to the action hand (Panel AHT = 11.5, 14.95, 18.4). However, as the visual hand is rendered closer to the target, the reach error decreases and the target estimation is more accurate. However, as the distance between the target and the action hand increases (Panels 21.85, 25.3, 28.75), the reach error is negative when the visual hand is rendered very close to the target (Discrepancy = +11.5cm). This suggests that it is not the accuracy that is improving as the visual hand is placed closer to the target, rather the estimate of the target's location is "pulled" towards the proximal non-action hand, as the distance between the visual hand and target decreases.

Overall, the results suggest that it is not the visuo-proprioceptive discrepancy itself that is affecting the target's location estimation, since that would have displayed a convex or a concave curve, with same magnitudes of discrepancy showing similar reach errors. Rather, the proximity of the visual hand to the target biases the estimation of the target in its direction, suggesting that the spatial encoding of the target may be with respect to the visual information of the body part in proximity.

These results seem to contradict the hypothesis that the spatial encoding of the target is with respect to the position of the body-part estimated as a result of multi-sensory integration inference mechanism.

4.2.4   Lack of support for Bayesian Causal Inference Mechanism

The competing hypothesis that we will now consider is that the spatial information of the reach target is encoded with respect to the position of the proximal body-part which is estimated as a result of multi-sensory integration process. One of the models of multi-sensory integration is the *Bayesian Causal Inference Model*. This model specifies how noisy uni-sensory signals are weighted based on the probability that the signals arose from the same source, and then integrated according to Bayesian Inference (Körding et al., 2007; Meijer & Noppeney, 2020). Previous literature suggests that this model can reasonably explain how position of the body part is estimated from noisy visual and proprioceptive signals (Noel et al., 2018; Fossataro et al., 2020). We implemented the model similar to the implementation by Fossataro et al. (2020) to estimate the position of the non-action

hand from the visual and proprioceptive signal positions as specified by the experimental manipulations described in the above experiment.

According to the model, the uni-sensory information is modeled as a probability distribution, specifying the probability at all possible locations that the uni-sensory object is present at that particular location. In this implementation, the visual and the proprioceptive signals are assumed to follow a Gaussian distribution, centered at the veridical location of the signal (specified by the experimental manipulation), with variance indicative of the sensory noise. In our implementation, we set the value of visual precision (inverse of variance) as 0.97, and the value of proprioceptive precision as 2.1, on the basis of results found by Noel et al. (2018). The third parameter of the model is P_π, which represents the "prior" beliefs that the visual and proprioceptive signals belong to the same source. We set the value of $P_\pi = 0.5$ in our implementation.

The complete model is specified as -

$$\begin{aligned}
P(hp|x_v, x_p) &= \Sigma_{c=own,diff} P(hp|x_v, x_p, C = c)P(C = c|x_v, x_p) \\
&= P(hp|x_v, x_p, C = own)P(C = own|x_v, x_p) \\
&+ P(hp|x_v, x_p, C = diff)P(C = diff|x_v, x_p)
\end{aligned}$$

where, $P(hp|x_v, x_p)$ is the probability of inferred location of the non-action hand, given the uni-sensory signals x_v and x_p. $C = own$ refers to the situation where the source of the two signals is same, while $C = diff$ refers to the situation where the source of the two signals is different. When $C = own$, $P(hp|x_v, x_p, C = own)$ is the Gaussian posterior distribution, centered at the average of the uni-sensory signals, with a variance equal to the inverse of the sum of the precision of the uni-sensory signals. When $C = diff$, $P(hp|x_v, x_p, C = diff)$ is the probability distribution relying only on proprioceptive signal - Gaussian centered at mean of proprioceptive signal, with the proprioceptive signal noise as its variance. The weighing factors $P(C = own|x_v, x_p)$ and $P(C = own|x_v, x_p)$ specify the probabilities that the two uni-sensory signals arise from the "own" hand or "different" sources conditional on the uni-sensory signals.

We implemented this model to derive the probability distributions of non-action hand position at different levels of visuo-proprioceptive discrepancy. Figure 4.4 shows these distributions with panels ordered according to the mode of the inferred hand position. Reach target veridical position at 21.85 cm from action hand and its average estimated
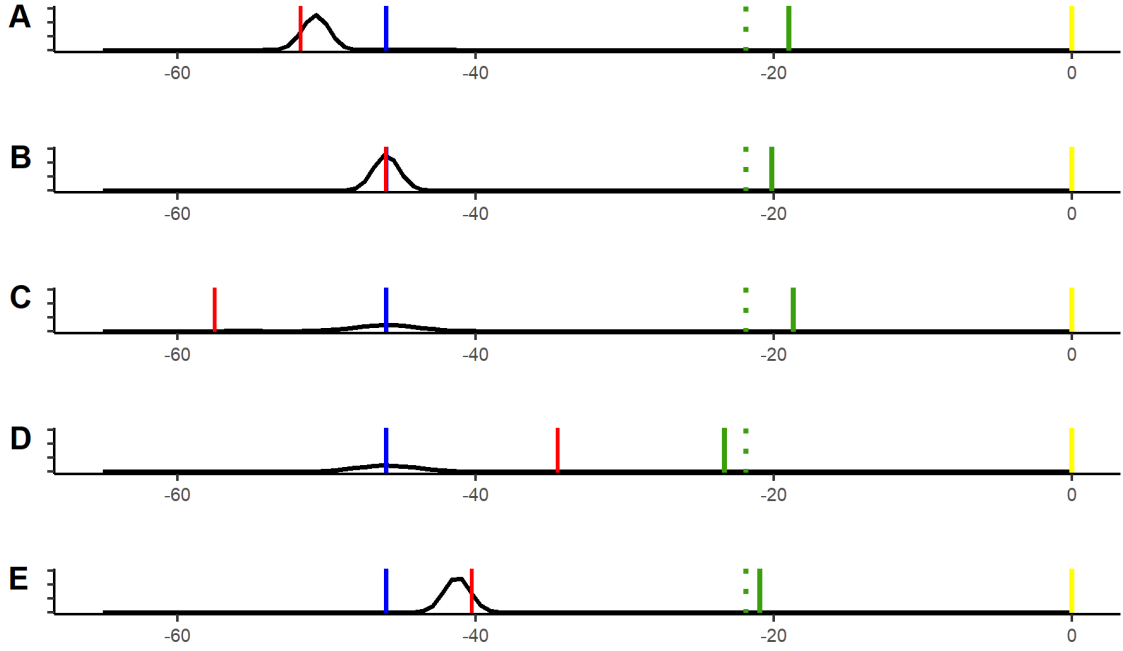
Figure 4.4: Estimated target location for Target Location = 21.85 cm, for different proximal hand's inferred positions according to the Bayesian Causal Inference Model. The red line indicates position of the visual signal of the proximal hand, while the blue line indicates the position of the proprioceptive signal, that is, the location of the placement of the real hand. The black curve indicates the probability density function denoting the probability of the location of the hand, as postulated by the multi-sensory integration process specified in the Bayesian Causal Inference Model. The mode of the distribution indicates the inferred position of the hand. The green dotted line indicates the veridical position of the reach target, while the solid green line indicates the (average) located estimated by the subjects. The yellow line indicates the initial position of the action hand. Panels A to E are ordered by the increasing proximity of the inferred hand position to the reach target, As seen in the figure, there is no systematic effect of proximity between the target and inferred hand position, and the target location estimation.

position is also plotted (for other target positions, refer to Appendix C). We observe that in the case of a high magnitude of visuo-proprioceptive discrepancy (Panel C and D, discrepancy = -11.5cm and +11.5cm), the mode of the inferred hand position is situated at the position of the proprioceptive hand. This is due to high probability attributed to the source of the uni-sensory signals being different (C = diff), due to the higher magnitude of spatial discrepancy between them. On the other hand, when the visuo-proprioceptive signals are relatively close (Panel A and E, discrepancy = -5.75cm and +5.75cm), we observe that the mode of the inferred hand position is shifted towards the visual hand. This is due to the high probability attributed to (C = own) and higher precision of visual signal.

Qualitative inspection of the figure reveals that there is no systematic effect of the inferred position of the non-action hand on the reach target location estimation. In particular, consider Panel C and D. While the inferred position of the non-action hand is same in both cases, along with same uncertainty associated with the probability distribution, we observe differences in estimates for the reach target location. No systematic effect of inferred hand position is observed at other target positions as well (see Appendix C).

Therefore, the results of the experiment do not support the hypothesis that spatial encoding of the reach target occurs with respect to the inferred position of proximal body-part according to the Bayesian Causal Inference Model. On the other hand, the spatial encoding of reach target can be explained by encoding of the target's spatial information with respect to the visual signals of the proximal body-part, based on our results.

Chapter 5

General Discussion

The purpose of this thesis was to investigate the following questions: i) Is the spatial information of the target of reach action represented with respect to body-part situated in its proximity, even if that body-part is not the effector involved in the reach action? ii) If yes, and considering that visuo-proprioceptive integration mechanism may underlie the construction of the representation of the body (*body schema*), how is the integration of visual and proprioceptive signals specifying the spatial parameters of the proximal body-part involved in the spatial encoding of the reach target? To investigate these questions, we experimentally manipulated the presence of visual and proprioceptive information of a body-part proximal to the reach action target, and measured the accuracy of the target location estimated by the subjects in reaching action.

The results of Experiment 1 show that when a body part is present proximal to the target, the target location estimation is more accurate, compared to the situation in which the body-part is not present proximal to the target. This result thus supports the hypothesis that spatial information of the target is indeed represented in an egocentric reference frame, with respect to the body-part proximal to the target. Thus, our finding confirms the results from previous literature that the objects in peri-personal space (here, the reach targets) are spatially encoded in a somatotopic (body-part centric) frame of reference (Serino, 2019). Notably, we have shown that this hypothesis is further supported through the paradigm of accuracy of reaches, and not merely through the paradigm of multi-sensory interaction effects, that has been a norm in the literature (See Chapter 2).

However, the target location estimation is not improved significantly when only uni-sensory information (visual or proprioceptive) about the proximal body-part is present. This is a significant finding, because it suggests that it is more merely the presence of a visual

allocentric landmark that is affecting the target location estimation. Rather, the accuracy in target location estimation is improved only when multi-sensory visuo-proprioceptive signals of proximal body part are present. This suggests that visuo-proprioceptive multi-sensory integration mechanism plays a role in the spatial encoding of the reach target. This contradicts the results from various studies that show that presence of visual landmarks improves reaching performance (e.g. Krigolson, Clark, Heath, & Binsted, 2007); (For review, see Filimon, 2015). However, these studies used memory guided reach paradigm, instead of invisible action-hand reach paradigm to induce variability in the reaches. It is possible that memory-guided reaches have another (allocentric) mechanism underlying the reach target spatial encoding.

Fossataro et al. (2020) showed that when there is high visuo-proprioceptive discrepancy of a body-part, the peri-personal space as operationalized by the magnitude of multi-sensory interactions, enlarges, suggesting that since the estimation of body-part position (according to Bayesian Casual Inference) is less precise due to the high visuo-proprioceptive discrepancy, peri-personal space is enlarged to optimize reactions to external events. Particularly, their findings show that the magnitude of multi-sensory interaction effects are comparable when the visual stimuli occurs near to the proprioceptive or the visual signal of the body part. Our results contradict Fossataro et al. (2020)'s finding that the inference of body-part position affecting the spatial encoding of visual objects in peri-personal space. When visuo-proprioceptive spatial discrepancy is induced in the proximal body-part, the results of the experiment show that the estimation of the target is "pulled" towards the proximal body-part as a function of proximity of the target to the visual signal. This finding suggests that although the results from the previous experiment suggest that a body-part proximal multi-sensory integration mechanism underlies the spatial encoding of the target, the spatial encoding itself seems to occur with respect to the visual signal regarding the proximal body-part. Furthermore, we implemented the Bayesian Causal Inference Model to estimate the inferred location of the proximal body-part based on inputs from the spatially discrepant visual and proprioceptive signals. These inferred proximal body-part positions do not systematically explain the reach errors measured in the experiment. Thus, our results do not support the hypothesis that the reach target is spatially encoded with respect to the position of the proximal body-part as inferred from a multi-sensory integration process by an ideal Bayesian observer.

The discrepancies between the results may be due to the difference in the paradigms implemented in the two studies. In our experiment, we implemented a localization of visual stimuli in peri-personal space paradigm, while Fossataro et al. (2020) implemented a

multi-sensory interaction effects paradigm which is based on reaction times. It is plausible that the spatial properties attributed to the multi-sensory interaction effects observed in peri-personal space and the the spatial encoding of objects in peri-personal space may have different conceptual and mechanistic basis. Indeed, Bufacchi and Iannetti (2021) have discussed the idea that there is not a single peri-personal space representation, rather, a set of responses arising in the space surrounding the body, which may be indicative of distinct cognitive processes.

Overall, our novel findings are i) Multi-sensory visuo-proprioceptive information of the proximal body part may be necessary for the spatial encoding of reach target to occur with respect to it, however ii) the encoding may occur with respect to the visual signal, as opposed to the inferred position of the proximal body-part estimated by an Ideal Bayesian observer. There are however several limitations of our study which must be acknowledged. Firstly, there are some intrinsic limitations in our experimental design. The distance between proprioceptive signal and the target (which keeping the distance between the action hand and target constant) has been been systematically manipulated, which prevents us from making stronger claims regarding the spatial encoding of the target with respect to the proprioceptive signal. Secondly, in our study, the Bayesian Causal inference Model has been implemented to get predictions regarding the inferred body-part position according to an ideal Bayesian observer, and reach responses are compared for different levels of proximity of inferred body-part position to the reach target. While our results and analysis merely point towards the probable multi-sensory mechanism, a stronger case can be made if the reach error itself can be modeled in the Bayesian Causal Inference framework, while experimentally determining the parameters for the model.

In summary, the findings from our work suggest that a proximal body-part's visuo-proprioceptive multi-sensory integration mechanism underlies the spatial encoding of reach target, and that the spatial encoding occurs with respect to the visual information before integration with the proprioceptive information, however, further empirical investigation is required to substantiate this claim.

Appendix A

Experiment 1: Data Analysis

A.1 Assumptions of Repeated Measures ANOVA

	Effect	W	p	p<.05
1	block	0.90	0.87	
2	sti_pos	0.25	0.00	*
3	block:sti_pos	0.14	0.02	*

Table A.1: Mauchly's Test for Sphericity

	Effect	GGe	DF[GG]	p[GG]	p[GG]<.05	HFe	DF[HF]	p[HF]	p[HF]<.05
1	block	0.94	2.81, 56.17	0.02	*	1.10	3.31, 66.29	0.02	*
2	sti_pos	0.57	1.14, 22.89	0.00	*	0.58	1.17, 23.37	0.00	*
3	block:sti_pos	0.60	3.62, 72.38	0.23		0.75	4.52, 90.35	0.22	

Table A.2: Spericity Corrections

A.2 Post-hoc Analysis Results

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.signif
1	avg_mre	L	C	21	21	2.06	20.00	0.05	0.05	ns
2	avg_mre	L	R	21	21	3.22	20.00	0.00	0.01	**
3	avg_mre	C	R	21	21	3.73	20.00	0.00	0.00	**

Table A.3: Pair-wise t test between levels of Target Position. Adjusted P-values are Holm-Bonferroni corrected.

	.y.	group1	group2	n1	n2	statistic	df	p	p.adj	p.adj.signif
1	avg_mre	X	PV	21	21	2.94	20.00	0.01	0.05	*
2	avg_mre	PV	P	21	21	-2.55	20.00	0.02	0.10	ns
3	avg_mre	PV	V	21	21	-1.51	20.00	0.15	0.59	ns
4	avg_mre	V	P	21	21	-1.32	20.00	0.20	0.61	ns
5	avg_mre	X	V	21	21	1.31	20.00	0.20	0.41	ns
6	avg_mre	X	P	21	21	-0.14	20.00	0.89	0.89	ns

Table A.4: Pair-wise t test between levels of Visuo-proprioceptive information of Non-action hand. Adjusted P-values are Holm-Bonferroni corrected.

Appendix B

Experiment 2: Data Analysis

B.1 Linear Mixed effects model comparisons

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m.null	5.00	28745.70	28778.67	-14367.85	28735.70			
m.d	6.00	28730.54	28770.11	-14359.27	28718.54	17.16	1	0.0000

Table B.1: Comparing Null Model and Discrepancy Model

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m.null	5.00	28745.70	28778.67	-14367.85	28735.70			
m.a	6.00	28697.76	28737.33	-14342.88	28685.76	49.94	1	0.0000

Table B.2: Comparing Null Model and AHT model

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m.d	6.00	28730.54	28770.11	-14359.27	28718.54			
m.d.a	7.00	28682.54	28728.70	-14334.27	28668.54	50.00	1	0.0000

Table B.3: Comparing Discrepancy Model and Discrepancy + AHT model

	npar	AIC	BIC	logLik	deviance	Chisq	Df	Pr(>Chisq)
m.d.a	7.00	28682.54	28728.70	-14334.27	28668.54			
m.d.a.int	8.00	28665.72	28718.48	-14324.86	28649.72	18.82	1	0.0000

Table B.4: Comparing Discrepancy + AHT model and Discrepancy * AHT model

B.2 Best fit model summary

Groups	Name	Variance	Std.Dev.	Corr
pid	(Intercept)	7.74261	2.7826	
	discrepancy	0.01535	0.1239	0.05
Residual		11.43551	3.3816	

Table B.5: Random effects for the best-fit model

	Estimate	Std. Error	df	t value	$Pr(> |t|)$	
(Intercept)	0.0391762	0.6712434	22.9858763	0.058	0.954	
discrepancy	-0.2634481	0.0383537	58.0775067	-6.869	4.83595E-09	***
AHT	-0.0556558	0.0077649	5365.490895	-7.168	8.67E-13	***
discrepancy:AHT	-0.0041789	0.0009624	5369.621227	-4.342	1.43692E-05	***

Table B.6: Fixed effects estimates for the best-fit model

B.3 Assumptions for best-fit LME Model

B.3.1 Absence of Collinearity

	Term	VIF	VIF_CI_low	VIF_CI_high	SE_factor	Tolerance	Tolerance_CI_low	Tolerance_CI_high
1	discrepancy	1.75	1.69	1.82	1.32	0.57	0.55	0.59
2	AHT	1.00	1.00	Inf	1.00	1.00	0.00	1.00
3	discrepancy:AHT	1.75	1.69	1.82	1.32	0.57	0.55	0.59

Table B.7: Variance Inflation Factor values as a measure for collinearity between predictors

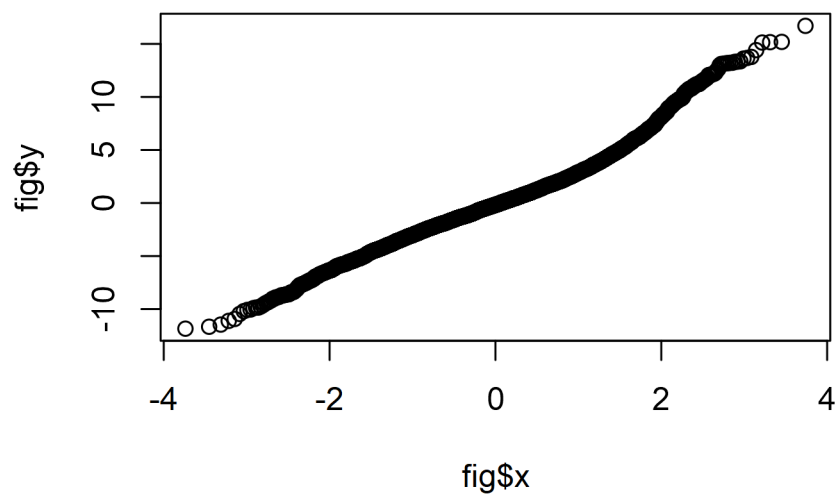## B.3.2   Normality of residuals
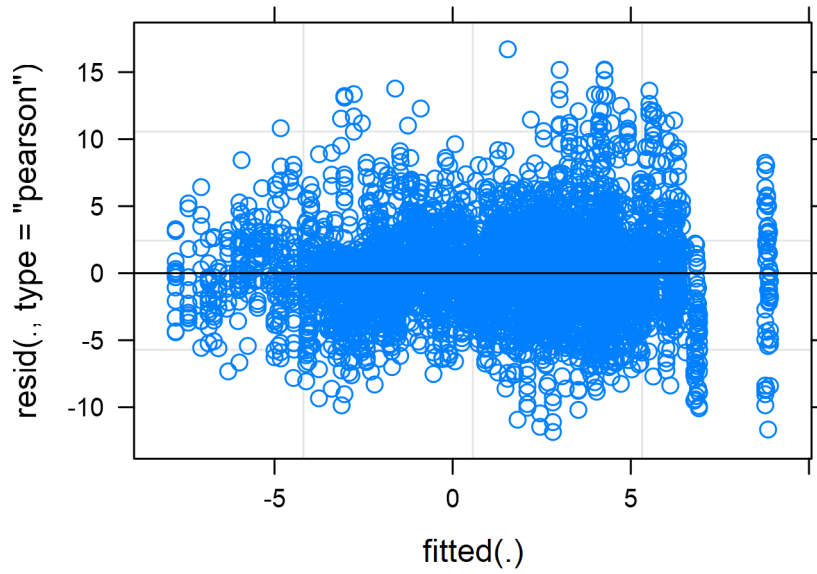

Figure B.1: Normal Q-Q plot

B.3.3 Homoscedasticity

Figure B.2: Residual plot for the best-fit model

B.4 Results of Bootstraps estimates

	term	estimate	lower	upper	type	level
1	(Intercept)	0.04	-1.34	1.38	perc	0.95
2	discrepancy	-0.26	-0.33	-0.19	perc	0.95
3	AHT	-0.06	-0.07	-0.04	perc	0.95
4	discrepancy:AHT	-0.00	-0.01	-0.00	perc	0.95

Table B.8: Bootstrapped Confidence Interval estimates for the best-fit model

|	term	observed	rep.mean	se	bias
1	(Intercept)	0.04	0.02	0.68	-0.02
2	discrepancy	-0.26	-0.27	0.04	-0.00
3	AHT	-0.06	-0.06	0.01	-0.00
4	discrepancy:AHT	-0.00	-0.00	0.00	-0.00

Table B.9: Bootstrapped Estimates for the best-fit model

Appendix C

Bayesian Causal Inference Model Estimates

The plots in this appendix show the predictions of proximal hand's inferred positions according to the Bayesian Causal Inference Model, along with the estimated target locations (measured experimentally) for different target locations.

The red line indicates position of the visual signal of the proximal hand, while the blue line indicates the position of the proprioceptive signal, that is, the location of the placement of the real hand. The black curve indicates the probability density function denoting the probability of the location of the hand, after multi-sensory integration process postulated by the Bayesian Causal Inference Model. The mode of the distribution indicates the inferred position of the hand. The green dotted line indicates the veridical position of the reach target, while the solid green line indicates the (average) located estimated by the subjects. The yellow line indicates the initial position of the action hand. Panels A to E are ordered by the increasing proximity of the inferred hand position to the reach target.
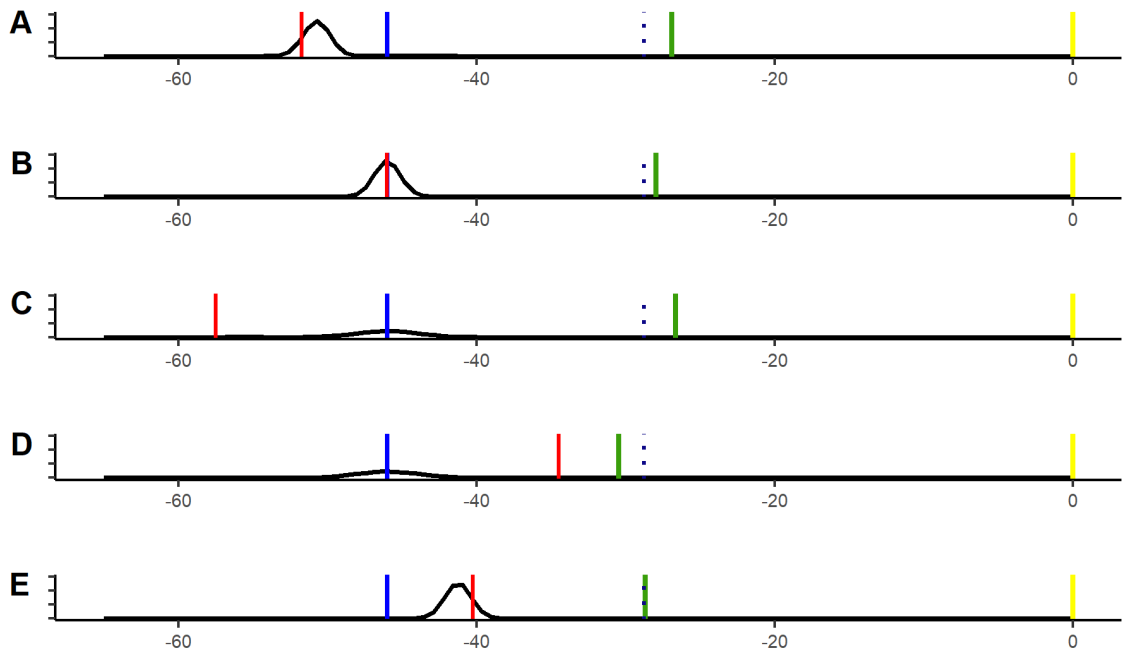
Figure C.1: Target Location = 28.75 cm
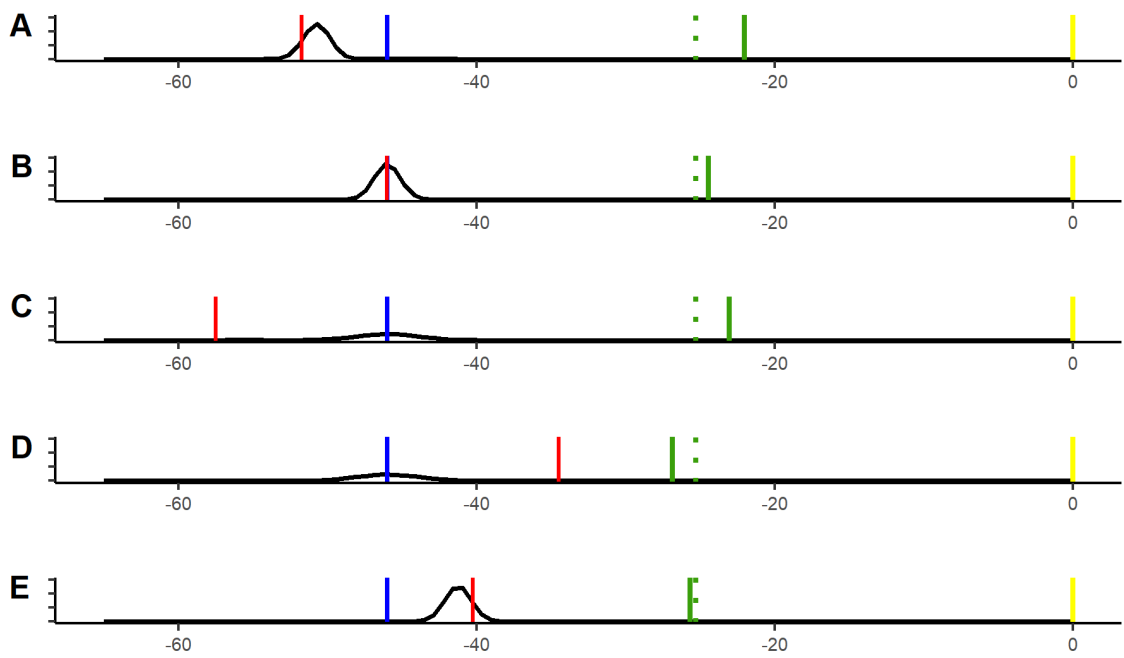
Figure C.2: Target Location = 25.3 cm
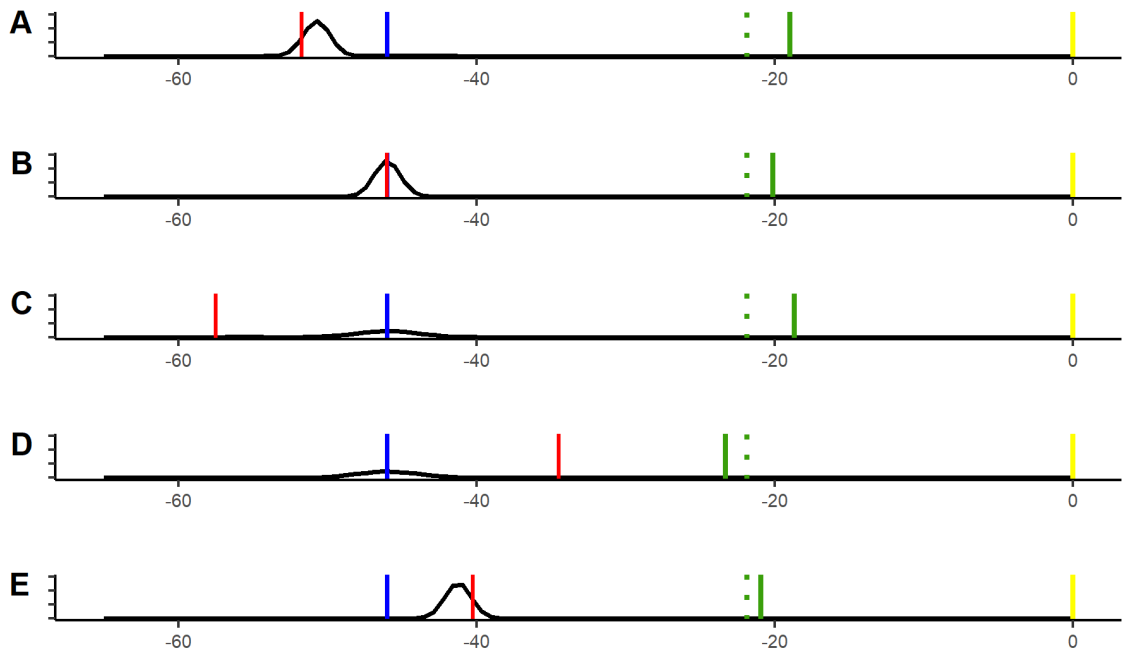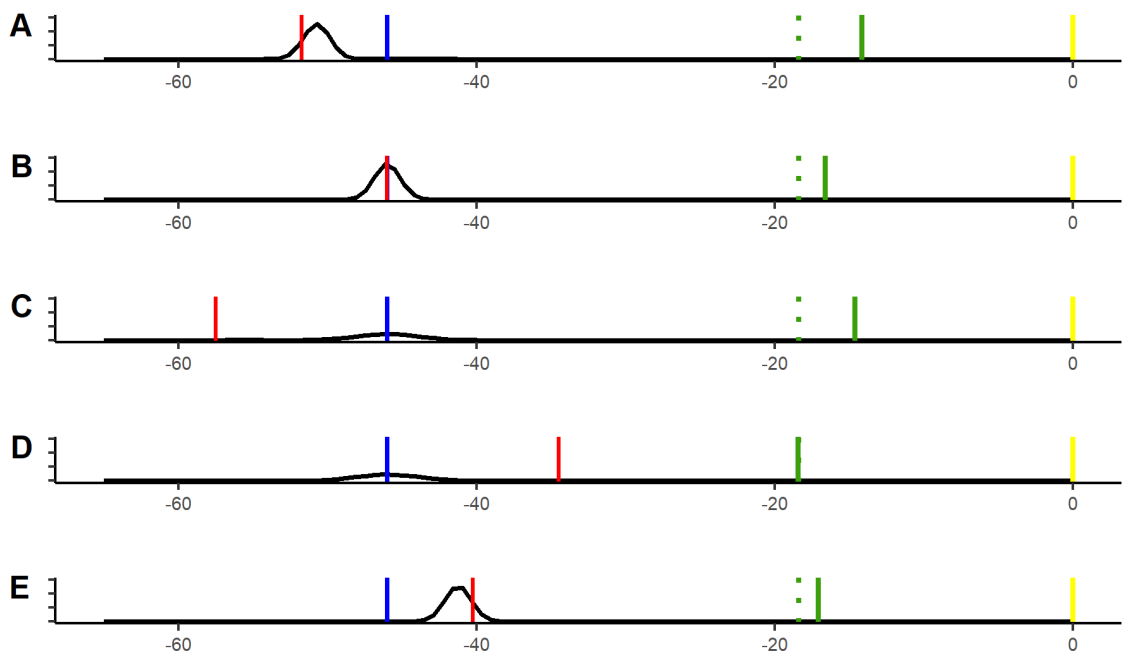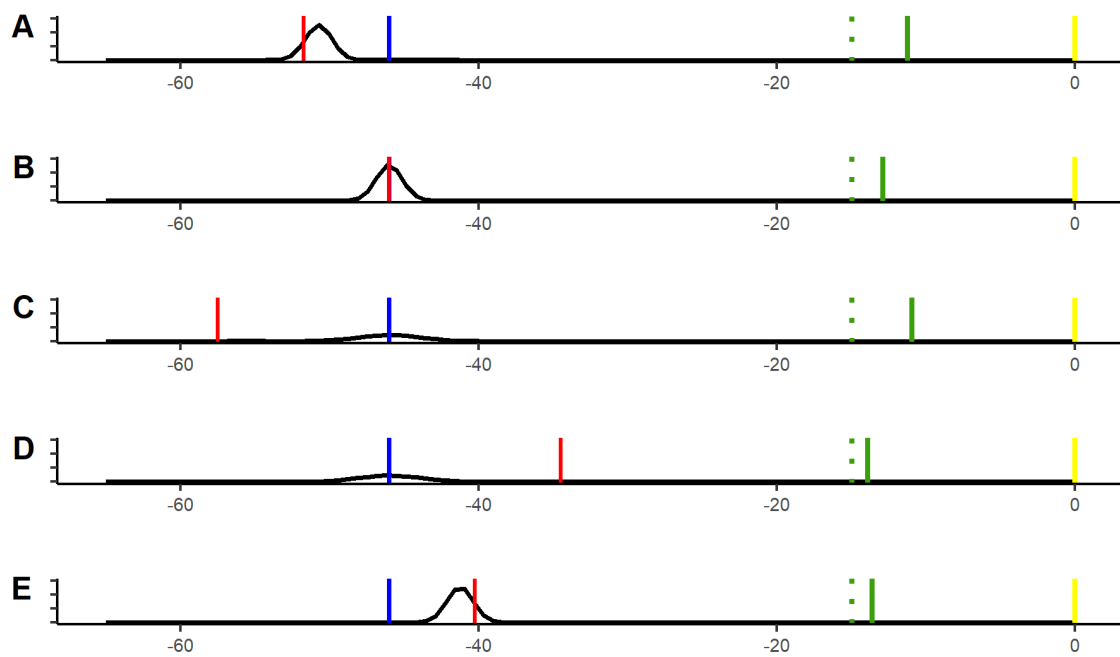
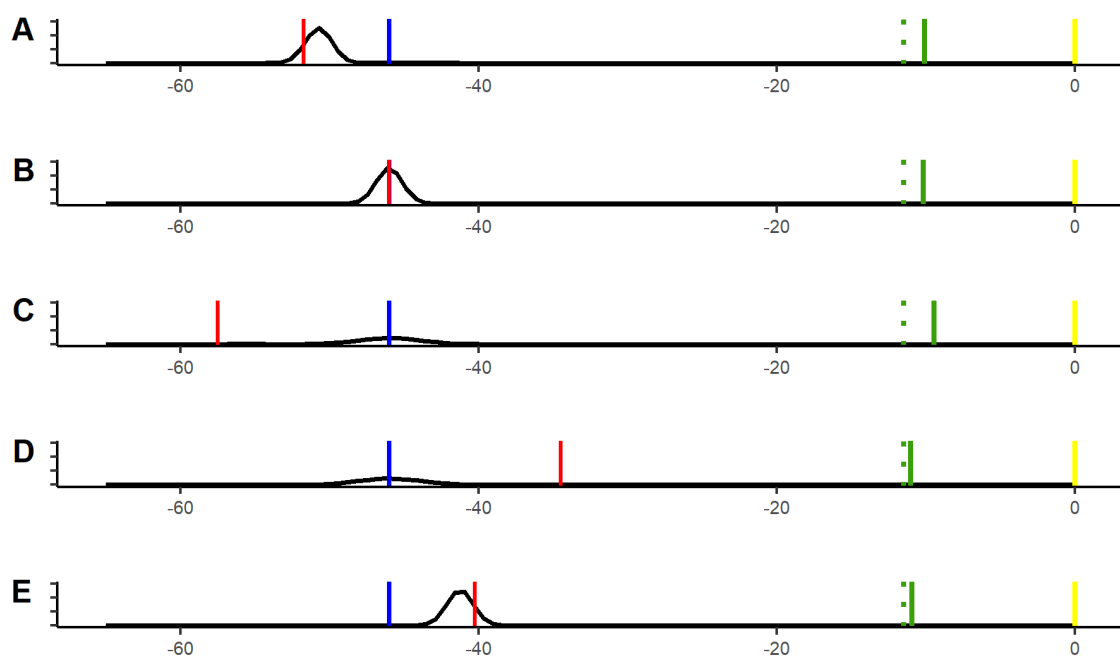**Figure C.3:** Target Location = 21.85 cm

Figure C.4: Target Location = 18.4 cm

Figure C.5: Target Location = 14.95 cm

Figure C.6: Target Location = 11.5 cm

Appendix D

Instructions and Consent Documents

Instructions for the experiment

Please read the following instructions carefully.

1. You will be sitting across the table throughout the course of the experiment.
2. There are two docks attached to the table. You will rest your index fingers on these docks.
3. The starting position for the task is to rest your right arm on the table comfortably, with the right index finger resting in the right dock.
4. You will be asked to follow verbal instructions for calibration at the beginning of the experiment.

The Task:

1. There are 6 phases in this experiment. Depending upon the phase, **you will be instructed regarding the placement of your left hand.** Make sure that when the instructions appear, you pause and read the instructions carefully, before continuing the task.
2. The task begins with your right index finger resting on the right dock.
3. On an audio cue, your task is to touch the red dot appearing on the table with your **right index finger.**
4. After touching the red dot, the red dot will disappear.
5. Whenever you are ready to begin the next trial, bring your finger back to the dock.
6. As soon as you touch the dock, an audio cue will indicate to repeat this task again.

Some points to note:

1. Depending on the phase, you may or may not see your hands. However, in all phases, you have to perform the task with your right hand.
2. If your right hand is not visible, you will not know whether or not you have touched the red dot. Just try to **do your best to be accurate**.
3. In this experiment, it is more important for you to be accurate in your task rather than the speed. Try to make **unspeeded** movements at a comfortable pace.

Figure D.1: Experiment 1: Instructions

<u>**INSTRUCTIONS**</u>

- There are two docks attached to the table. You will rest your index fingers in these docks.
- The starting position for the task is to rest your right arm on the table comfortably, with the right index finger resting in the right dock.

<u>The Experiment:</u>

1. The experiment is divided into multiple phases.
2. There are two positions for your LEFT hand - on the left dock, OR on your lap. Depending upon the phase, **you will be instructed regarding the placement of your left hand.**
3. Make sure that when the instructions appear, you pause and read the instructions carefully, before continuing the task.

The Task:

1. The task begins with your right index finger resting on the right dock.
2. When you hear a PING sound, you have to **touch the red dot** appearing on the table with your **right index finger.**
3. After touching the red dot, the red dot will disappear.
4. Whenever you are ready to begin the next trial, bring your finger back to the dock.
5. As soon as you touch the dock, the PING sound will indicate to repeat this task again.

Some points to note:

- Since your right hand will not be visible, you will not know whether or not you have touched the red dot. Just try to do your best to be accurate. **It is more important to be accurate in your task than speed. Try to make unspeeded movements at a comfortable pace.**

Figure D.2: Experiment 2: Instructions

Virtual Reality Health Disclaimer

The experiment that you are about to participate in involves the use of Virtual Reality(VR) Technology. Virtual Reality is an immersive experience, and there are some risks associated with its use.

Please read carefully the following list of potential issues that you may be affected with while using Virtual Reality Technology.

1. Discomfort: : eye strain; blurred, altered or double vision; motion sickness; nausea; loss of awareness; dizziness; light-headedness; drowsiness; impaired balance or hand-eye coordination; disorientation
2. Seizures: triggered by flickering screens and flashing images or patterns which may occur while experiencing the VR
3. Interference with Medical Devices: the VR equipment may affect the operation of some medical devices such as cardiac pacemakers or other implanted medical devices
4. Pre-existing Health Issues: the VR equipment (including the headsets and rumble chairs) and content may affect those with pre-existing medical (such as a heart ailment), mental health conditions (such as anxiety or post-traumatic stress disorder), or aggravate physical injuries (such as back or neck conditions)
5. Use of the VR experience may adversely affect pregnant women
6. It is possible to feel anxiety, claustrophobia, or other phobias when using the VR equipment.

Disclaimer Statement

By signing this statement, I agree with the following,

1. I recognise and understand that participation in the VR experiment involves certain risks.
2. During the course of the experiment, I am NOT extremely tired, ill, under the influence of alcohol and drugs, hung-over, in pain, or feeling stressed or anxious in any way, as these factors increase my risk of potential issues.
3. I am participating in the Virtual Reality Experiment voluntarily.
4. I understand that I can stop using the Virtual Reality Equipment immediately if I feel any discomfort whatsoever.
5. I will follow the instructions and safety warnings given by the experimenter during the course of the experiment.

Signature of the Participant   Signature of the Experimenter

Participant Name   Experimenter Name

Date   Date

Figure D.3: Virtual Reality Health Disclaimer

Consent Form

Study Title: **Visuo-proprioceptive Hand-Target coupling in reaching action**

Participant's Full Name _____
Date of Birth/Age _____
Address _____

I confirm that I have been explained the nature and procedure of the study by the
Investigator and had the opportunity to ask questions.

1) I understand that my participation in the study is voluntary and that I am free to
   withdraw at any time, without giving any reason and without my medical care or
   legal rights being affected.
2) I understand that the Ethics Committee and the regulatory authorities will not need
   my permission to look at my health records both in respect of the current study and
   any further research that may be conducted in relation to it, even if I withdraw
   from the study.
3) I have been informed that a record of my personal data/ experimental
   data/transcribes resulting from my participation in this will be confidentially
   maintained by the Principal Investigator of this study and under no circumstances
   will this data be released to third parties.
4) I agree not to restrict the use of any data or results that arise from this study
   provided such a use is only for scientific purpose(s). I agree/permit the use of data
   arising from this study to be used for publication provided the confidentiality and
   anonymity is maintained.
5) I agree to undergo the proposed study at Indian Institute of Technology, Kanpur. I
   have been explained about the complications and side effects, if any, and have fully
   understood them.

Signature of the Participant   Signature of the Investigator

Signatory's Name   Signatory's Name

Date   Date

Figure D.4: Consent for participation

References

Ataria, Y., Tanaka, S., & Gallagher, S. (2021). *Body schema and body image: New directions.* Oxford University Press.

Avillac, M., Hamed, S. B., & Duhamel, J.-R. (2007). Multisensory integration in the ventral intraparietal area of the macaque monkey. *Journal of Neuroscience, 27*(8), 1922–1932.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Brozzoli, C., Cardinali, L., Pavani, F., & Farnè, A. (2010). Action-specific remapping of peripersonal space. *Neuropsychologia, 48*(3), 796–802.

Bufacchi, R., & Iannetti, G. (2021). What do 'peripersonal space measures' really reflect? the action field perspective.

Byrne, P. A., Cappadocia, D. C., & Crawford, J. D. (2010). Interactions between gaze-centered and allocentric representations of reach target location in the presence of spatial updating. *Vision research, 50*(24), 2661–2670.

Conti, P., & Beaubaton, D. (1980). Role of structured visual field and visual reafference in accuracy of pointing movements. *Perceptual and Motor Skills, 50*(1), 239–244.

Filimon, F. (2015). Are all spatial reference frames egocentric? reinterpreting evidence for allocentric, object-centered, or world-centered reference frames. *Frontiers in human neuroscience, 9*, 648.

Fossataro, C., Sebastiano, A. R., Tieri, G., Poles, K., Galigani, M., Pyasik, M., ... Garbarini, F. (2020). Immersive virtual reality reveals that visuo-proprioceptive discrepancy enlarges the hand-centred peripersonal space. *Neuropsychologia, 146*, 107540.

Graziano, M. S., & Gross, C. G. (1994). Mapping space with neurons. *Current Directions in Psychological Science, 3*(5), 164–167.

Iachini, T., Ruggiero, G., Ruotolo, F., & Vinciguerra, M. (2014). Motor resources in peripersonal space are intrinsic to spatial encoding: Evidence from motor interference. *Acta psychologica, 153*, 20–27.

Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport*, *7*(14), 2325–2330.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS one*, *2*(9), e943.

Krigolson, O., Clark, N., Heath, M., & Binsted, G. (2007). The proximity of visual landmarks impacts reaching performance. *Spatial vision*, *20*(4), 317–336.

Ladavas, E., & Serino, A. (2008). Action-dependent plasticity in peripersonal space representations. *Cognitive neuropsychology*, *25*(7-8), 1099–1113.

Limanowski, J., & Blankenburg, F. (2016). Integration of visual and proprioceptive limb position information in human posterior parietal, premotor, and extrastriate cortex. *Journal of Neuroscience*, *36*(9), 2582–2589.

Lohmann, J., Belardinelli, A., & Butz, M. V. (2019). Hands ahead in mind and motion: Active inference in peripersonal hand space. *Vision*, *3*(2), 15.

Longo, M. R., Schüür, F., Kammers, M. P., Tsakiris, M., & Haggard, P. (2008). What is embodiment? a psychometric approach. *Cognition*, *107*(3), 978–998.

Meijer, D., & Noppeney, U. (2020). Computational models of multisensory integration. *Multisensory Perception*, 113–133.

Noel, J.-P., Samad, M., Doxon, A., Clark, J., Keller, S., & Di Luca, M. (2018). Peripersonal space as a prior in coupling visual and proprioceptive signals. *Scientific reports*, *8*(1), 1–15.

Patané, I., Cardinali, L., Salemme, R., Pavani, F., Farnè, A., & Brozzoli, C. (2019). Action planning modulates peripersonal space. *Journal of Cognitive Neuroscience*, *31*(8), 1141–1154.

Pavani, F., Spence, C., & Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological science*, *11*(5), 353–359.

Rizzolatti, G., Matelli, M., & Pavesi, G. (1983). Deficits in attention and movement following the removal of postarcuate (area 6) and prearcuate (area 8) cortex in macaque monkeys. *Brain*, *106*(3), 655–673.

Schütz, I., Henriques, D., & Fiehler, K. (2013). Gaze-centered spatial updating in delayed reaching even in the presence of landmarks. *Vision Research*, *87*, 46–52.

Schwoebel, J., & Coslett, H. B. (2005). Evidence for multiple, distinct representations of the human body. *Journal of cognitive neuroscience*, *17*(4), 543–553.

Senna, I., Cardinali, L., Farnè, A., & Brozzoli, C. (2019). Aim and plausibility of action chains remap peripersonal space. *Frontiers in Psychology*, *10*, 1681.

Serino, A. (2019). Peripersonal space (pps) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience &*

Biobehavioral Reviews, *99*, 138–159.

Spence, C., Pavani, F., & Driver, J. (2004). Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 148–169.

Spence, C., Pavani, F., Maravita, A., & Holmes, N. (2004). Multisensory contributions to the 3-d representation of visuotactile peripersonal space in humans: evidence from the crossmodal congruency task. *Journal of Physiology-Paris*, *98*(1-3), 171–189.

Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology: Human perception and performance*, *31*(1), 80.

Van Beers, R. J., Sittig, A. C., & Gon, J. J. D. v. d. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of neurophysiology*, *81*(3), 1355–1364.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current biology*, *11*(18), R729–R732.