Faculty of Medicine and Psychology

PhD Program in Behavioral Neuroscience

XXXII cycle

PhD thesis in

Do your actions help me remember better? Effects of pointing movements performed by self and others on visuo-spatial working memory: Evidence from behavioral, eye-tracking and event-related potential data

Submitted by

Divya Bhatia

Supervised by

Prof. Clelia Rossi-Arnaud

Table of Contents

Abstract

The current dissertation investigates how pointing movements performed towards to-be-remembered target locations together with a task-partner influence visuo-spatial working memory (VSWM). Previous studies have shown that, under specific conditions, spatial arrays that have been pointed to during encoding are better recognized than passively viewed ones. The first experiment described here was aimed at replicating these results. Participants were presented with two spatial arrays, one of which was encoded through mere visual observation (the no-move array), while the other was encoded through observation accompanied by pointing movements towards target locations (the move array). The encoding was followed by a recognition test in which participants had to decide whether a probe array was equal or different from one of the arrays presented during the study phase. Results replicated the memory advantage of pointed-to arrays as compared to merely observed ones. Since a number of studies suggest that perceiving an action and executing it are functionally equivalent and activate common representational structures, we proceeded to investigate whether pointing movements performed by a task-partner could have similar effects. No memory advantage was found when participants passively observed the pointing movements performed by the task-partner. A joint condition of the task was thus tested where the participant and the task-partner took turns in performing the pointing movements towards the target locations in the move array (recognition was always performed by the participants). Results showed that in this joint-action condition the pointing movements performed by the task-partner benefited individual's visuo-spatial working memory to the same extent as self-performed pointing movements. Further experiments were conducted to investigate the processes underlying the encoding of no-move and move arrays, either jointly pointed by the participant and the task-partner or pointed only by the task-partner.

In two eye-tracking experiments, the eye movement patterns for the two types of arrays were recorded. The analysis of eye movements yielded results that are in line with the behavioral findings and showed that fixations to move arrays were higher in number and longer in duration than the fixations to no-move arrays but only when the participant and the task-partner alternate in making the pointing movements. No difference in eye movement patterns was found for the move and no-move arrays when the participant merely observed pointing movements performed by the task-partner. We propose that, in the joint-action condition, self and others' actions are coded at the same representational level and their functional equivalency is reflected in a similar pattern of eye fixations. Behavioral experiments investigating the effects of a number of other factors on visuo-spatial working memory are also reported, including the partner's involvement in the task or the animacy of the task-partner. Findings show that: a) in the joint-action condition, the infrequent execution of pointing movements by the participants (about 16%) is sufficient to successfully produce a social context which leads to the co-representation of partner's actions; and b) a memory advantage occurs when the turn-taking task-partner is replaced by a moving cartoon hand, but only in the joint-action condition (there is no advantage when the movements were only carried out by the cartoon hand). The latter finding indicates that the mere presence of the cartoon hand is not sufficient to trigger the memory advantage. Altogether, the results suggest that the social context between the participant and the task-partner modulates the task co-representation process. Finally, event related potential (ERP) data were recorded in order to investigate the encoding processes involved in processing self and other's pointing movements in the joint-action condition. In summary, the current dissertation investigates, from behavioral and physiological points of view, the question of how performing an action together with a partner

influences individual’s encoding processes and suggests future directions for research investigating the influence of jointly performed tasks on individual’s memory.

Acknowledgements

This dissertation would not have been possible without the constant support and guidance from my PhD advisor Prof. Clelia Rossi-Arnaud. I thank her for encouraging focus, clarity of ideas and freedom to think and test my ideas which subsequently resulted into this dissertation. I also thank her for guiding me to navigate the challenges of PhD throughout these three years. It would not have been an easy process if it would not be a supportive, kind and generous mentor as she is. I thank her for providing me with a model of dedication to both on one hand, mentorship and research and on the other for keeping personal commitments in both professional and personal life that will keep inspiring me throughout my career. Dr. Pietro Spataro is also due thanks for his constant feedback on my dissertation research.

I also thank Prof. Guenther Knoblich who generously welcomed me in his lab during my participation in the Joint Mobility Program for PhD and providing me with the facilities and constant support and guidance to perform my EEG experiment. I also thank Dr. Dimitrios Kourtis for his guidance during the designing of the EEG study. Dr. Anna Zamm is also due thanks for sharing her EEG knowledge with me. Thanks are also due to Dr. Stefano Lasaponara for helping me with analyzing EEG data at the point when I had such a limited time for submission. Thanks are also due to Prof. Ramesh Mishra who happily collaborated on the eye-tracking project.

I would like to thank my partner for his constant love, care and support during the entire PhD process; from the days I was dull to the days I was overjoyed.

I would also like to thank my brother who read most of my dissertation and highlighted errors which were left unnoticed by me during the process of writing. Additional thanks to my close friends from India who have provided me with inspiration and friendship over the years. Finally, I would like to thank my parents who have been incredibly supportive and loving throughout this entire process and who are always proud of each little achievement I make.

Organization of the thesis

The current dissertation comprises three chapters, one based on a literature review, one illustrating the 10 different experiments undertaken and the last one reporting a general discussion of the results obtained in all the different studies performed. The first chapter (Chapter 1) gives an overview of the three different lines of research which make up the current research question: a) how performing an action together with a partner influences individual’s working memory? The first line of research comes from the studies investigating the effect of self- performed pointing movements towards a visuo-spatial array. The studies discussing detrimental and beneficial effect of pointing movements on spatial memory are reported; b) the second line of research comprises the studies investigating the effect of jointly performed tasks on individual’s performance. An overview of studies discussing the shared representational structures for perceiving and performing an action is provided; c) the third line of research comprises the studies investigating the memory of action events. An overview of the studies suggesting how and when enacting action phrases are remembered better than other modalities such as verbal encoding or observation encoding is provided.

The second chapter (Chapter 2) reports the experimental work and is divided into four sections. The first (section 2.1) reports four behavioral experiments investigating how and when partner's actions influence individual's memory in a visuo-spatial working memory task. Section 2.2 complements the major findings obtained in section 2.1 by providing eye-movements data in two experiments. Section 2.3 reports three behavioral experiments investigating the boundary conditions for the findings obtained in section 2.1. Finally, section 2.4 complements the behavioral findings obtained in section 2.1 with the electrophysiological data and provides further insights on the mechanisms of how partner's actions influence one's motor system and subsequent memory.

Chapter 3 is a general discussion of the behavioral, eye movements and electrophysiological findings of the current study. The findings are discussed with regard to different accounts such as the motor simulation account or the task co-representation account. A discussion based on attentional processes is also provided. The limitations of the present study and the future research questions that remain open to investigation are also discussed.

Finally, a bibliography of the studies cited in the current dissertation is provided.

Chapter 1: General Introduction

1.1 Visuo-spatial Working Memory

1.1.1 Introduction:

In the working memory model originally proposed by Baddeley and Hitch (1974), and later modified by Logie (1995), the visuo-spatial sketch pad or scratch pad (VSSP), later known as visuo-spatial working memory (VSWM), served as a “slave” system alongside the articulatory loop, under the control of the central executive. The central executive is thought to be responsible for reasoning, decision making, and for coordinating the functions of its specialized slave systems. The articulatory loop (also known as the Phonological Loop, Baddeley 1990; 1992; Baddeley and Logie, 1992) is considered to be responsible for the storage and manipulation of verbal information whereas VSSP serves a similar function for visual and spatial information. ‘Visual’ information refers to the static representation of item properties such as shape, color and brightness whereas ‘spatial’ information is a reference to the location of items in space and geometric relationships between those items or moving from one item to another (Logie, 1991). More specifically, spatial refers to a representation that involves movements in its broader sense, such as movement in the form of scanning a visual array, movement to a target in an array or movement of objects in an array. The cognitive structure of VSSP had been considered unitary until the studies showed disruptive effects of movement while maintaining Brooks matrix material. In the Brooks Task (1967), the subjects were asked to imagine digital numbers in consecutive squares of a 4 x 4 matrix. Subjects either read and listened, or only listened to the task instructions such as in the first square put a 1, then place a 4 in the upper left corner etc. Upon completion of the presented sequences, subjects responded with recall of the series of sentences vocally. Brooks (1967) found that the visual presentation of the sentences interfered with their recall, while verbal presentation did not. He concluded that internal representation of spatial material uses mechanisms specialized for visual perception. Byrne (1974) showed that visually guided arm movements disrupted the recall of a sequence of pictures when they were spatially arranged. Baddeley, Grant, Wight and Thomson (1975) using the Brooks task, explored the effect of concurrent perceptual-motor tracking on retention of the Brooks matrix. While listening to the sequence of sentences, subjects were asked to concurrently follow a moving target with their arm. It was found that tracking had a significantly disruptive effect on the retention of the number positions in the Brooks matrix. Conclusions were made that VSWM not only involves visual-perceptual input but also is needed for the production and control of movement. The disruption of VSWM performance has been shown by a number of studies since then using the Brooks matrix (Brooks, 1967) and Corsi Block Task (De Renzi and Nichelli, 1975). These set of studies involved the work by Mary Smyth and her colleagues (Smyth, Pearson and Pendleton, 1988; Smyth and Pendleton, 1989), Gerry Quinn and colleagues (Quinn and Ralston, 1986; Quinn, 1988, 1991), the experiments conducted by Alan Baddeley (Baddeley and Lieberman,1980) and Robert Logie (Logie, Zucco and Baddeley, 1990). These experimental data led to the view of the division of VSWM into two separate systems: proposed by Logie (1995), that VSWM be subdivided into two systems: the visual cache and the inner scribe. The visual cache is a passive store system handling static information such as colors and shapes, whereas the inner scribe is responsible for information related to movement sequences and works as an active rehearsal mechanism for the visual cache.

1.1.2 Interaction of visuo-spatial working memory with movement

1.1.2.1 Detrimental Interaction: Movements performed during the encoding and/or retention of spatial information occupies the inner- scribe, therefore leaving less available resources for the rehearsal of information, resulting into a lower memory performance. Most of the evidence in support of this hypothesis comes from dual task studies showing that the temporary retention of spatial information is interfered with by concurrent movements generated in response to unrelated targets (Baddeley & Lieberman, 1980; Logie & Marchetti, 1991; Smyth & Scholey, 1994; see Quinn, 2008, for a review). Using the Brooks task, Baddeley and Lieberman (1980) demonstrated a substantial fall in performance when participants were concurrently asked to maintain a beam of light on a swinging pendulum. Similarly, Quinn and Ralston (1986) combined the Brooks paradigm with a secondary task in which participants moved their hand on an actual matrix taped to the table. The matrix on the table and the subject's arm were covered, and movements had to be completed without the subjects being able to see their arm. This ensured that the movement task does not involve any visual processing. The executed arm movements were either compatible or incompatible to the verbally presented sentences in the Brooks task. They found that incompatible movement disrupted recall of the Brooks matrix material but compatible movement did not. They further added a "passive movement" condition to this experiment, where participant's hand was being moved by the experimenter. The disruption of recall of Brooks matrix continued to be seen with incompatible passive arm movement. Other studies have reached similar conclusions using the continuous sequential tapping task (also known as the matrix-tapping task), in which participants repeatedly tap four keys arranged in a square (Morris, 1987; Smyth & Pelky, 1992). Smyth and colleagues asked their participants to recall the sequence of blocks indicated by experimenter in the Corsi blocks task. They found that the performance on the Corsi blocks task was not disrupted by arm movements but was disrupted by tapping a square pattern of four switches. These data showed a distinction between body movements and movements to specified target locations (also see, Smyth and Pendleton, 1989). Vandierendonck, Kemps, Fastame, and Szmalec (2004), further, investigated the effects of this type of interference on the computerized version of the Corsi blocks task and found that recall performance was impaired for both short and long block sequences (see also Rossi-Arnaud, Pieroni, Spataro, & Baddeley, 2012). These results can be easily interpreted within the Logie's model of VSWM by supposing that the execution of the movements in the interference tasks occupies the inner scribe, leaving less resources available to rehearse the spatial locations presented in the primary task.

1.1.2.2 Beneficial interaction: Despite the overwhelming evidence showing detrimental effects of unrelated movements towards to be remembered stimuli, recent studies have shown a beneficial effect of pointing movements performed towards target stimuli. In a more recent study by Chum, Bekkering, Dodd and Pratt (2007) pointing movements were found to be beneficial when performed towards to-be-remembered spatial locations. In this study, participants were presented with two visuo-spatial arrays one after the other, made up of white-filled circles and squares. Both arrays contained the same number of items (varying between three, four and five items) but differed in shape. Each item from two spatial arrays was presented for 1 sec and occupied a different location on the screen. Participants were asked to pay attention and try to remember these item locations. On each trial one array was encoded via visual observation only (no-move array) whereas the other array was encoded via visual observation accompanied by pointing movements (move array). Their memory for one of the two array locations was tested in a recognition task at the end of each trial. An illustration of the paradigm is shown in Figure 1.1.

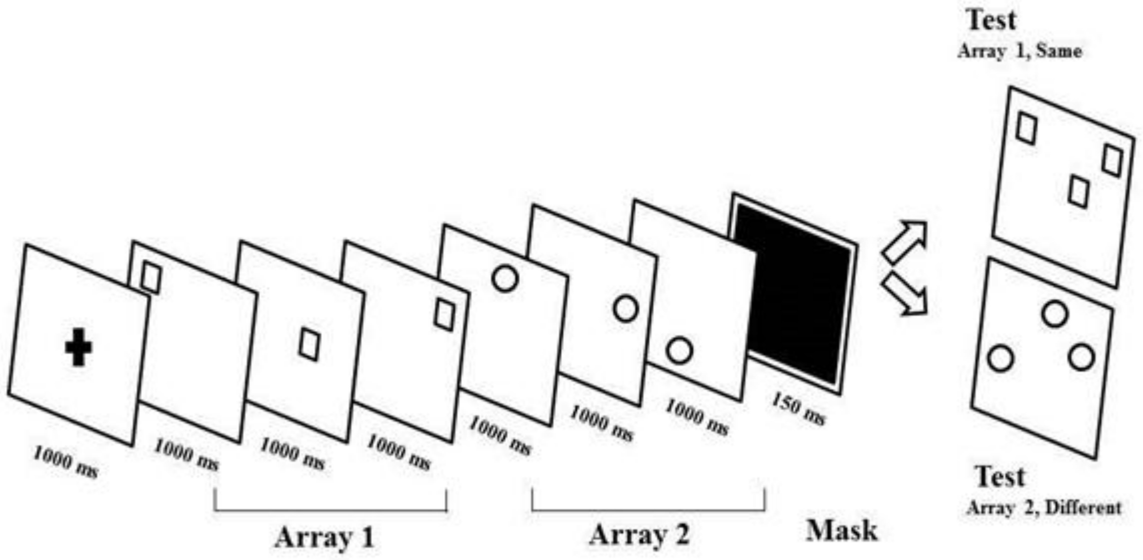


Figure 1.1: Illustration of a trial from the visuo-spatial working memory task used by Chum et. al. 2007. Two visuo-spatial arrays (with 3, 4 and 5 items each) differing in shape were presented. One array was encoded visually while the other was encoded by making pointing movements towards the item locations during its presentation. Following a mask, memory for one of the two arrays was tested.

Results from an immediate recognition task showed that the move arrays were recognized significantly better than the no-move arrays. Chum and colleagues proposed that action-based encoding increased the spatial-based perceptual saliency of target locations in the case of move arrays, subsequently resulted in their higher recall. They also suggested that action-based encoding may have resulted into stronger form of egocentric (body-based) encoding whereas simply viewing the items may have resulted into the allocentric (scene-based) encoding. These two different encoding processes led to the memory advantage obtained for the move arrays over no-move arrays. They also related the memory advantage obtained for move arrays to the enactment effect which refers to the improvement in the recall of action phrases when their encoding is accompanied by action-performance (Mohr, Engelkamp and Zimmer, 1989; Engelkamp, 1998). Finally, Chum and colleagues concluded that the memory advantage in the case of move arrays was caused by either increased spatial-based perceptual bias caused by the act of pointing or because of the formation of a stronger egocentric spatial map in parallel with allocentric map or some integration of these two mechanisms. The results obtained by Chum et al. (2007) are shown in Figure 1.2.

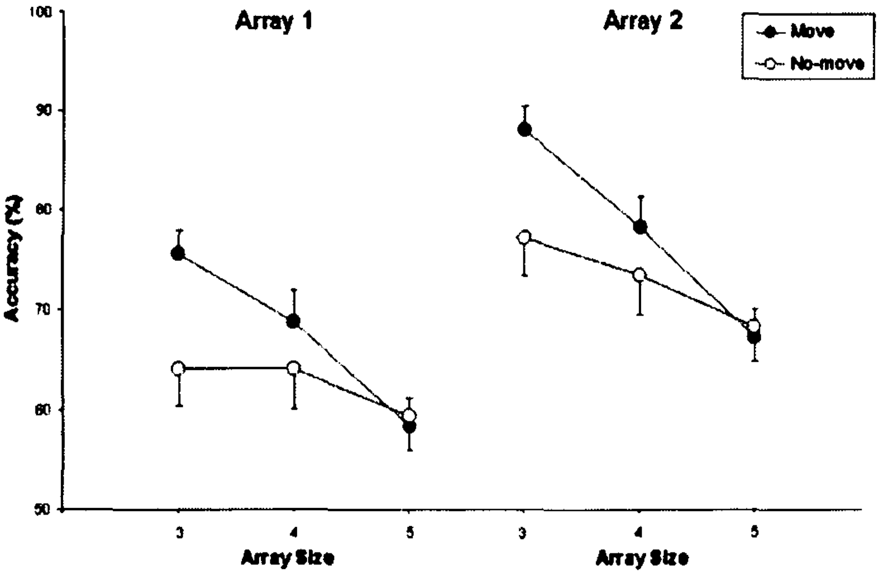


Figure 1.2: The results shown by Chum et al. (2007, p. 1191) for Experiment 1. Recognition accuracy (%) is shown as a function of Array order (Array 1 and Array 2), Array size (3, 4 and 5) and Condition (Move and No-move). Move arrays were remembered better than no-move arrays at array sizes 3 and 4. The memory for move and no-move arrays for array size 5 was found to be similar. The general recency effect was observed such that the arrays presented lastly were remembered better than the arrays presented first. Error bars depict 95% confidence intervals. Reprinted from “Motor and visual codes interact to facilitate visuospatial memory performance,” by Chum, M., Bekkering, H., Dodd, M., & Pratt, J. 2007, *Psychonomic Bulletin and Review,* 14 (6), p. 1191. Copyright 2007 by the Springer Nature. Reprinted with permission.

Furthermore, following the pointing manipulation made by Chum et al. (2007), Dodd and Shumborski (2009) first manipulated the pointing instruction within trial (e.g. in a given trial, half of the items were encoded through visual observation only while the other half were encoded through visual observation accompanied by pointing movements) and replicated the pointing advantage for the move arrays over the no-move arrays (Experiment 1). They further suggested that while manipulating the pointing instruction within a trial, it was unclear as to whether the memory advantage for move arrays reflected enhanced processing of these arrays due to the act of pointing, or inhibited processing of no-move arrays or some combination of these two factors. To investigate this, they manipulated the pointing instruction between blocks (Experiment 2) such that in one block, all the items in a trial were encoded through either by visually observing them (no-move array) or by performing pointing movements towards item locations (move array). In Experiment 2, a very different pattern of results emerged showing superior memory for the no-move arrays relative to the move arrays. Moreover, by comparing the memory performances from Experiment 1 and 2, Dodd and Shumborski (2009) showed that the enhanced memory for move arrays in Experiment 1 was also accompanied by poorer memory for no-move arrays in this condition as compared to when the pointing instruction was manipulated between blocks in Experiment 2. Dodd and Shumborski (2009) emphasized the requirement of selection (Allport, 1987) that is when a subset of items is selected for additional processing, the memory for these items increases, on the other hand, the memory for unselected items gets hurt. Experiment 2 showed that when pointing instruction is manipulated between blocks, pointing action actually decreased the VSWM performance. They suggested that the act of pointing to half of the items enhances memory for these items relative to the other half of the items which were passively viewed. They further added that the assumption of having superior memory when encoding items with both a perceptual and motor code relative to encoding items with only perceptual code, seems not to be true. In fact, memory for items which were encoded via both perceptual and motor code, seems to be inferior than the items encoded only via perceptual code when all the items in a trial were encoded in the same way. They therefore suggested that the critical variable is that of selection; when a subset of items requires selection for additional processing, not only the processing of selected items is enhanced but also the processing for unselected items is reduced. On the other hand, when no selection is required, all the items are treated equally. The results obtained by Dodd and Shumborski (2009) are shown in Figure 1.3. The finding that the memory for the selected items is enhanced and memory is reduced for the items not selected for action has also been demonstrated in other selection based memory effects such as retrieval-induced forgetting (Anderson, Bjork and Bjork, 2000, 1994). In retreivel-induced forgetting paradigm, participants study a series of category-exemplar pairs and are instructed to learn them for a later recall test. Participants are then cued to recall half of the items from half of the studied categories. It has been found that participants recall the items better that they have studied and practised repeatedly than the items which were unpractised.

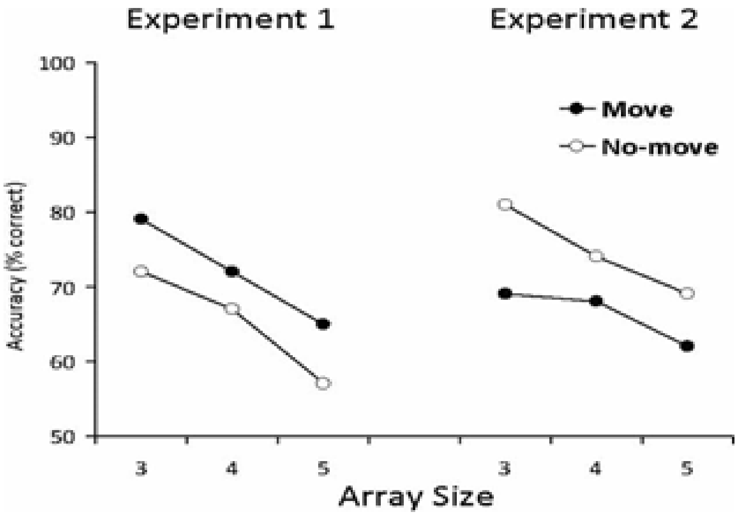


Figure 1.3: Results shown by Dodd and Shumborski (2009, p. 1240) in Experiment 1 and 2. In Experiment 1 pointing instruction is manipulated within trial that is in each trial, one array is encoded by passive observation whereas the other is encoded by passive observation accompanied by pointing movements. In Experiment 2 pointing instruction is manipulated between blocks that is in one block all the items in a trial are either passively

observed or pointed. Reprinted from “Examining the influence of action on spatial working memory: The importance of selection. By Dodd, M., & Shumborski, S. 2009, *Quarterly Journal of Experimental Psychology* , *62* (6), p. 1240. Copyright 2009 by the Taylor & Francis. Reprinted with permission.

Dodd and Shumborski (2009) concluded that the pointing manipulation probably taps into the inner scribe while passive observation manipulation of encoding engages only visual cache. This creates the possibility that different representations are being formed at encoding and effectiveness of representation will be a function of task type. These findings altogether suggest that the interaction between action and working memory is quite complex.

1.2 Shared Action Representations

1.2.1 Common representations of perceived and planned actions:

An interesting part of human nature is that we not only execute our own actions but also imagine them and observe others doing them. The idea that perceiving others' actions creates a tendency to perform these actions, was first postulated by the ideomotor theory (Lotze, 1852; James, 1890). Originally, it was postulated that imagining an action would create a tendency to perform the same action; a later modified version of the ideomotor theory extended the idea to social interactions by suggesting that not only when we imagine an action, but also when we observe someone performing an action, creates a tendency to perform the same action (Greenwald, 1970; Prinz, 1997). It was then suggested that imagining and observing an action and performing the same action should be coded in a common representational domain, known as a *common-coding approach* (Prinz, 1990, 1997; Hommel et al., 2001). The core point to the common-coding approach is that the perceived actions and planned actions share a common representational domain (Prinz, 1997). The evidence supporting the common-coding approach have been provided by both behavioural and neuroimaging studies. Studies have shown that performance of an action is facilitated when the observed action matches with the action to be performed and is impaired when the observed action and the action to be performed do not match (Brass, Bekkering and Prinz, 2001; Kilner, Paulignan and Blakemore, 2003; Bertenthal, Longo and Kosobud, 2006; Brass, Bekkering, Wohlschläger and Prinz, 2000). Furthermore, studies on nonconscious mimicry have shown that people unintentionally adopt the postures (Bernieri and Rosenthal, 1991; Chartrand and Bargh, 1999; La France, 1979, 1982; La France and Broadbent, 1976), mannerisms (Chartrand and Bargh, 1999) and facial expressions (Anisfield, 1979; Bavelas, Black, Lemery and Mullett, 1986) while interacting with others in many social situations. These sets of findings can be explained by the assumption that observing an action creates a tendency in the observer to perform the same action, because observed and performed actions share common representational structures (Chartrand and Bargh, 1999). Moreover, in neuroimaging and neurophysiological studies, it has been evidenced that observed actions are represented similar to the executed actions. An increasing body of research has been demonstrating the existence of mirror- like neuronal population in humans which fires both when we perform an action and when we observe someone performing the same action (Grèzes and Decety, 2001). A complex network of neuronal circuits including superior temporal sulcus, the inferior parietal lobe and the inferior frontal gyrus is thought to play a key role in action imitation and observation (Grèzes and Decety, 2001) and found to be activated during both execution and observation of an action (Grèzes and Decety, 2001; Blakemore and Decety, 2001; Buccino, Binkofski and Riggio, 2004; Grèzes, Passingham and Frith, 2004). Mirror neurons were first discovered in area F5 of the monkey premotor cortex which discharge both when the monkey performs an action and when it observes a similar action performed by another monkey or experimenter (Rizzolatti, Fadiga, Gallese and Fogassi, 1999; Rizzolatti et al., 1996; Gallese et al., 1996; Hari et al. 1998; Cochin et al., 1999; Buccino et al., 2001). In a similar vein, the *motor simulation theory* (MST; Jeannerod, 1994, 2001, 2006a; Decety & Grèzes, 2006; Grèzes & Decety, 2001) also predicts a neural similarity of covert action states such as action observation and imagination and overt execution of an action. The motor simulation theory postulates that the motor system is part of a simulation network that is activated under a variety of conditions such as action planning and execution or observing actions from other individuals. The theory is supported by several evidences providing similar temporal characteristics for imagining an action and execution of the same action (Decety et al., 1989). Studies have shown that the speedaccuracy trade-off (Fitts's Law) can be observed in motor imagery (Toni, Gentilucci, Jeannerod and Decety, 1996; Sirigu et al., 1995). Several evidences have also been provided in support of the motor simulation theory by physiological, neurophysiological and neuroimaging studies throughout the years. For instance, physiological measures such as respiratory rate, skin resistance etc. have been found to be similar in both imagined and executed movements (Guillot and Collet, 2005a; Guillot et al., 2007). At the neurophysiological level, the mirror neuron system (a complex network including the superior temporal sulcus, the inferior parietal lobe and the inferior frontal gyrus) have reportedly been found to be activated during covert action states and overt execution of same action (for a review see, Jeannerod, 2001;Grèzes & Decety, 2001). It has been suggested that the mirror system enables us to understand the intentions of a conspecific through an internal embodied simulation of observed actions (Enticott, Johnston, Herring, Hoy and Fitzgerald, 2008; Jeannerod, 2001). Thus, they provide a neural mechanism for the direct mapping of perceived actions onto motor representations of the same actions.

1.2.2 Task co-representation in jointly-performed actions:

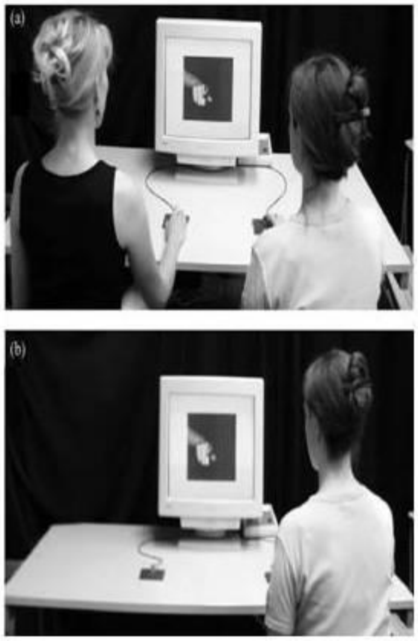
People engage in social interactions all the time such as playing games together, cooking together, dancing, playing a duet or moving towards a common goal together. Such joint actions make people achieve complex goals together which would be difficult to achieve alone. The term *joint action* is typically defined as the ability to coordinate one's own actions with those of others 'in space and time to bring about a change in the environment' (Knoblich and Sebanz, 2006, p. 100). Ideomotor theory provides an understanding of how perceiving others induces specific tendencies in us to engage in the same actions (Greenwald, 1970; James, 1890; Jeannerod, 1999; Prinz, 1997). Ideomotor theory suggests that observing someone's actions might activate event representations that are functionally equivalent to the event representations used in one's own control of these actions. To test these assumptions, Sebanz, Knoblich and Prinz (2003), using a joint version of the classical Simon task (Craft and Simon, 1970; Simon, 1990), investigated how performing a task together with a partner influences individual's task performance. They devised a joint version of the Simon task in which responses were distributed between two participants. Participants executed left and right button press, in response to green and red colored rings placed on the index finger of a pointing hand stimulus (irrelevant stimulus spatial dimension). The index finger of the hand either pointed to the left, or the right, or straight. In the compatible trials, the index finger pointed in the same direction as response required by the ring color (for example, the finger pointed to the left, and required response was left button press). On the other hand, in the incompatible trials, the finger pointed in the opposite direction of the response required by the ring color (for example, the finger pointed to the left, and required response was right button press). A compatibility effect was found both in the two-choice RT condition and joint condition (see 1.2.2.1, for details). Authors suggested that participants experienced response activation and conflict whether the alternative response was their own (as in two-choice RT condition) or their partner's (joint condition). They further suggested that a joint compatibility effect emerged as a result of the co-representation of partner's responses in participant's mind, that is participants took their partner's actions into account when performing a task together and represented partner's actions as their own. A tremendous number of studies thereafter provided behavioural and physiological evidence of shared representations of action and perception in joint task settings. Atmaca, Sebanz, Prinz and Knoblich (2008) distributed Eriksen flanker task among pairs of participants where both partners took care of complementary part of the task. They showed that participants were slower in responding to stimuli containing flankers which were potential targets for their partner (incompatible trials) as compared to the compatible or neutral stimuli. This joint flanker effect was also observed when participants merely believed to be performing the task with a partner. A similar interference effect was observed when participants named pictures alongside a task-partner (Kuhlen and Rahman, 2017). Moreover, it has been shown that knowing a [1]co-actor's task affects one's own action planning and performance even when the task does not require to take partner's actions into account (Atmaca et al., 2008; Sebanz, Knoblich and Prinz, 2003, 2005). With regard to memory, in a further study by Eskenazi and colleagues (Eskenazi, Doerrfeld, Logan, Knoblich and Sebanz, 2013), a pair of participants performed a categorization of stimuli task such as participant A responded to animals, participant B responded to fruit/vegetable items and household items were not assigned to anyone. A surprise recall test revealed that co-actor's relevant items were remembered better than no-one's relevant items. Authors suggested that shared task representations can establish shared memories.

1.2.2.1 The joint Simon task paradigm: Engaging in joint actions requires participants to perform a task together with a partner, where each partner is responsible for a complementary part of the task. To successfully perform a task together with a partner, participants require sharing representations with others, anticipating their behavior, coordinating individual's actions with others and integrating partner's action plans into their own action plans. But the answer of how people do that, comes from the common coding approach which suggests that observing an action performed by another person activates corresponding motor representations in the observer, and these motor representations can then be used to successfully coordinate with partner's actions (Brass, Bekkering and Prinz, 2001; Brass, Bekkering, Wohlschläger and Prinz, 2000; Tsai and Brass, 2007). Such motor corresponding representations

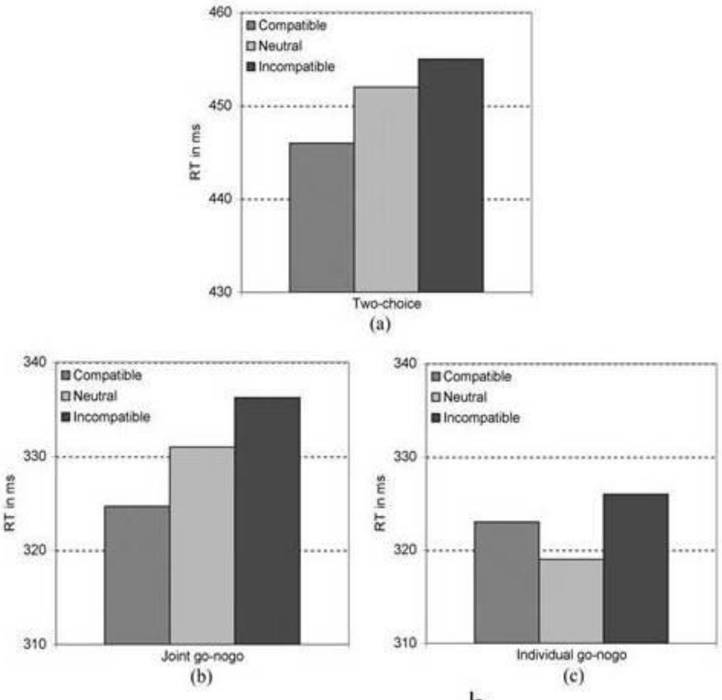
---

[1] The term task-partner is mostly used in the current dissertation to refer to the partner in the pointing

termed as co-representations have been studied using a joint version of the classical Simon task in joint action literature. Sebanz, Knoblich and Prinz (2003) developed a go-nogo version of the classical Simon task. Participants responded to digital photographs of a right human hand pointing either to the left, or to the right, or to the middle. The index finger of the hand wore a red or green coloured ring. Two groups of participants were assigned to two different task conditions. One group of participants performed the two-choice RT task, in which one participant seated in front of the screen (centrally) performed the whole task alone. Participants were instructed to respond to one ring color with the left button press and the other color with the right button press. Another group of participants performed both an individual go-nogo and joint go-nogo task. In the joint go-nogo condition, pair of participants sat side-by-side in front of a screen. In individual go-nogo task, an empty chair remained beside each participant. In the joint go-nogo task, one participant responded to one ring color (e.g. red) with left button press while the other responded to the other ring color (e.g. green) with the right button press. In the individual go-nogo task, the same task was performed alone such that each participant responded to only one ring color (go trials) while no-one responded to the other ring color (nogo trials). Responses were coded as compatible (when the finger pointed towards the person requiring to give a response), incompatible (when the finger pointed towards the other person than the one requiring to press the response button) and neutral (when the finger pointed to no-one i.e. to the middle). A compatibility effect was found both in the two-choice RT condition and in the joint go-nogo condition. Participants were faster to respond when the finger pointed towards the person requiring to give a response and were slower when the finger pointed towards the person not requiring to give a response. No such effect was found in the individual go-nogo condition. The results showed that the same go-nogo task is performed differently depending on whether a person performs the task alone or alongside another agent performing a complementary part of the task. They suggested that participants experienced an action selection conflict when they had to respond to a stimulus associated with partner's action acting in the same social context. Given that the two action alternatives and their spatial dimensions are represented, thus it overlaps with the spatial dimension provided by the stimulus. Therefore, just as in a normal Simon task, the response corresponding to the spatial dimension provided by the stimulus is automatically activated. For example, when the finger points left, the left is activated, irrespective of the person in charge of the left response. When activated and required responses match, there is no selection conflict but when they do not match, response conflict arises. For this reason, participants take longer to respond when the stimulus points to the other person and are faster when it points to themselves. This suggests that they represented an action alternative at a co-actor's disposal in a functionally similar way as to their own.



a.



b.

Figure 1.4: Sitting arrangements in joint and individual go-nogo task conditions and the results reported in the study by Sebanz, Knoblich and Prinz (2003, p. 14, p. 16) (a) Top panel shows the sitting arrangements for the joint go-nogo condition, bottom panel shows the sitting arrangements for the individual go-nogo condition. (b) Top graph shows the compatibility effect in the two choice RT condition, bottom left graph shows the compatibility effect in the joint go-nogo condition, bottom right graph shows the RTs in different trials in the individual go-nogo condition. Reprinted from "Representing others' actions: just like one's own?," by Sebanz, N., Knoblich, G., & Prinz, W. , 2003, *Cognition, 88*(3), p. 14, 16. Copyright 2003 by the Elsevier. Reprinted with permission.

1.2.2.2 Electrophysiological evidence of task co-representation: Sebanz, Knoblich, Prinz and Wascher (2006) in an event-related potentials study investigated how planning and control of actions are shaped by acting together with another individual. They further investigated whether partners' actions become part of one's own action plans and also whether acting in turns places special demands on the action control. EEG was recorded when participants performed an individual go-nogo task alone (individual condition) and together with a partner (joint go-no condition). By comparing the ERP components in the two task settings, they showed that a co-actor's action alternative is taken into account during action planning in the joint go-nogo condition and that holding one's response when it is partner's turn to respond requires increased inhibitory control. Results showed that in the go trials in both individual and joint go-nogo conditions, the P3 amplitude was larger in the compatible trials where the pointing direction of the finger corresponded to the required response as compared to the incompatible trials and it was more pronounced at central and parietal sites. They suggested that this effect might reflect perceptual interference. Moreover, they showed that at frontal sites, incompatible stimuli evoked a larger P3 in the joint condition than in the individual condition. This suggested that the stimuli referring to the partner's action seemed to be processed in a way similar to the stimuli referring to one's own actions. They explained this by the assumption that the two action alternatives were represented in a functionally equivalent way in the joint condition. In contrast, in the individual condition, stimulus in the incompatible trials did not refer to an action, therefore no such representations were processed. Furthermore, nogo P3 reflecting response inhibition was stronger in the joint as compared to the individual condition. This finding clearly suggested that perceiving a stimulus relevant to the co-actor's actions leads to the activation of representational structures involved in one's own execution of this action. Therefore, increased response inhibition was required to prevent oneself from responding on nogo trials in the joint condition. Another study by Tsai, Kuo, Jing, Hung and Tzeng (2006) further showed that nogo-N2 component reflected action inhibition in all conditions (Joint go-nogo, Individual go-nogo and 'being observed' condition, where participants were being observed by another person sitting next to them) whereas nogo-P3 was more pronounced in the joint condition reflecting higher control processing for action monitoring. They also showed LRPs in three social contexts to reflect motor preparation. LRPs from compatible trials of nogo responses and incompatible trials of go responses were larger in the joint than in the individual condition suggesting a significant modulation by the social context. Furthermore, Tsai and colleagues (Tsai, Kuo, Hung and Tzeng, 2008) showed the electrophysiological responses when the participants performed the Simon task under the believed intentionality of performing the task either with a human partner or with a computer. However, all the responses in both conditions were generated by a computer program. They analysed go-stimulus locked LRPs which had a positive deflection (100-200 ms) followed by a negative deflection (200-400 ms). The amplitude for the positive dip was larger for incompatible trials when participants were performing the task under the belief of interacting with a biological agent as compared to the incompatible trials when they performed under the belief of interacting with a computer. It reflected the conflict between the equivalent-purpose coding attributed to self-response and the response of a biological agent. Furthermore, when analysing the negative deflection in the LRPs they found that the amplitude for compatible trials was significantly higher than the amplitude for incompatible trials reflecting the conflict between the spatial and response dimensions. For the no-go stimulus locked LRPs, agency-dependent effects were found; amplitude for the positive deflection (100-200 ms) was significantly higher in the case of performing the task under the belief of interacting with a biological agent than the computer. For the negative deflection, it was found that the compatibility effect was only significant when they were performing the task under the belief of interacting with a biological agent. (see Figure 1.5 for the LRPs shown by Tsai et al., 2008).

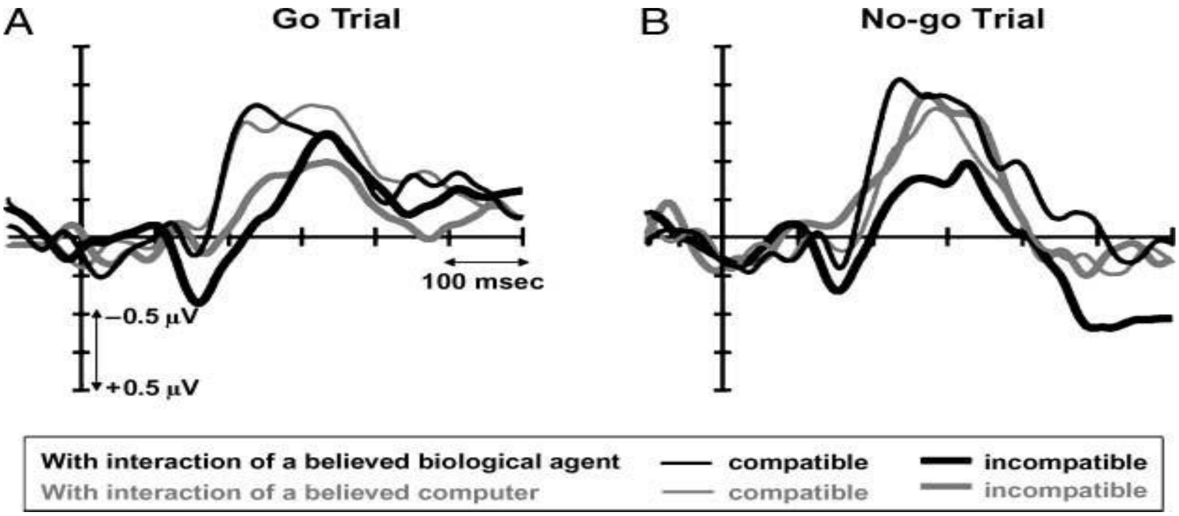


Figure 1.5: LRP results shown by Tsai et al. (2008, p. 2021). Compatibility difference in the two interacting contexts: performing the task under the belief of co-acting with a biological agent and computer. (A) Go trials (B) NoGo trials. Reprinted from “Action co-representation is tuned to other humans,” by Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J., 2008, *Journal of Cognitive Neuroscience*, *20*(11), p. 2021. Copyright 2008 by the Massachusetts Institute of Technology. Reprinted with permission.

1.2.2.3 Task co-representation process is modulated by social-context:

Sebanz and colleagues suggested that co-representations of partner’s actions are formed as a result of ‘socialness’ formed between the participants while performing a task jointly (Sebanz and Knoblich, 2009; Kourtis, Sebanz and Knoblich, 2010, 2013). It was further suggested that processing information related to the actions performed by other people requires special social mechanisms (Sebanz et al., 2006; Sebanz and Knoblich, 2009). An EEG study by Kourtis, Sebanz and Knoblich (2010, 2013) showed brain activations while modulating the social-context between the task partners. In Experiment 1, the participant, interaction partner and a loner (non interaction partner) were seated around a table forming an equilateral triangle. They performed a choice-reaction task where a cue indicated the type of action to be performed, followed by an imperative stimulus indicating them to act. Their task was to lift an object placed in the middle of the table, either returning it to its original position (individual action) or giving/receiving it to/from the partner (joint action condition). In joint action condition participants always interacted with the partner and the loner always acted alone. The late CNV (contingent negative variation) was more pronounced when participants anticipated the partner to act individually than when they anticipated that the same action would be performed by the loner. In Experiment 2, the participant and interaction partner faced each other on the table while the loner sat on the side. This more salient contrast between the partner and the loner showed a larger CNV for the partner as compared to the loner. Thus, stronger motor activations were found when participants anticipated an action to be performed by a partner they interacted with as compared to the same action performed by a person they did not interact with. They suggested that the simulation of another person's action gets stronger the more the other person is perceived as an interaction partner. Finally, it is clear that shared representations are formed as a consequence of interactive-context established between the task partners while performing a task together.

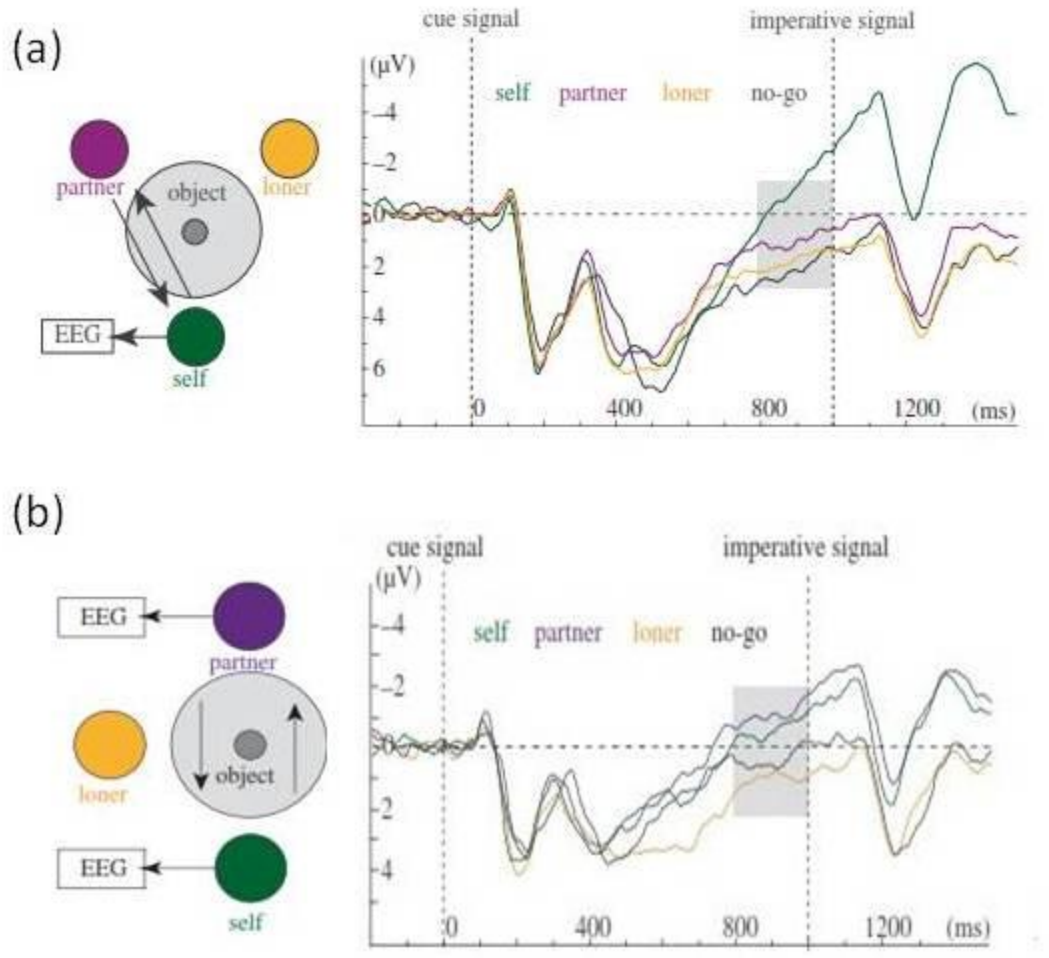


Figure 1.6: The effect of social context between the task-partners illustrated by CNV reported by Kourtis, Sebanz and Knoblich (2010, p. 759). (a) The left panel is the color-coded drawing of the experimental set up in Experiment 1. Right panel is the EEG results from Experiment 1: anticipating partner's actions showed more pronounced CNV than a loner. (b) The left panel is the color-coded drawing of the experimental set up in Experiment 2. Right panel is the EEG results from Experiment 2: higher CNV for anticipating actions for facing partner than the CNV for the loner sitting on the side. Reprinted from "Favouritism in the motor system: Social interaction modulates action simulation," by Kourtis, D., Sebanz, N., & Knoblich, G., 2010, *Biology Letters* , *6* (6), p. 759. Copyright 2010 by The Royal Society. Reprinted with permission.

Similarly, in the study by Tsai et al. (2008), participants performed the Simon task under the belief of co-acting with either an unseen person (biological agent) or a computer (non-biological agent), while they were actually co-acting with a computer program in both conditions. A joint Simon effect occurred only in the “believed” biological agent condition, indicating that the belief of co-acting with an intentional agent can influence the representations of alternative action events. However, another study by Welsh, Higgins, Ray and Weeks (2007) found a joint spatial compatibility effect only when the confederate sat beside the participant and performed his part of the task. Such an effect was absent when the confederate left the room and participants believed of co-acting with him on a networked computer. These results revealed that co-actor’s actions may only be represented when the participant is able to witness a portion of the action. Furthermore, Sellaro and colleagues (Sellaro, Treccani, Rubichi and Cubelli, 2013) showed that the belief of interacting with an intentional agent is not sufficient to elicit the required interference for a joint Simon effect to occur but this belief has to be attached with a spatial event occurring next to the participant such as imagining a response. Thus, once such an alternative spatial event is established, ongoing sensory information is not needed for cognitive representations (Sebanz et al., 2003; Vlainic et al., 2010). Furthermore, studies investigated the effect of interpersonal relationships between the task partners on joint Simon effect. Hommel, Colzato and van den Wildenberg (2009) showed that joint Simon effect occurs only when the co-actors are involved in a positive relationship (friendly and cooperative confederate) but not if they are involved in a negative relationship (intimidating, competitive confederate) (also see, Kuhbandner, Pekrun and Maier, 2010 and Iani et al., 2011). These results indicate that the interpersonal relationship between the task-partners modulates the task co-representation process.

1.2.3 Task co-representations are not so social: The task co-representation account suggests that the co-actor must be an intentional agent and the joint Simon effect is modulated by the social context in which two people are sharing the task. However, a finding that does not seem to be in line with the idea of co-representation being social and automatic is that the joint Simon effect is observed in autistic participants who have difficulties processing social information (Sebanz, Knoblich, Stumpf and Prinz, 2005). It was assumed that the co-actor's role in the joint Simon task may be to provide a spatial reference that allows coding of one's own action as left or right with respect to the other person exactly as one's own action alternatives provide a reference for relative response coding (Hommel, 1996; Guagnano, Rusconi and Umiltà, 2010). Dolk and colleagues (Dolk et al., 2011, 2013) in a series of experiments demonstrated that the Simon like effect emerged in the absence of any social/biological agent. First, they showed that when participants performed the Simon task alongside an inactive co-actor who was instructed to watch the participant's behavior and the co-actor's response button was associated with an attention-attracting event, a Simon effect emerged. Further, they showed that the Simon effect emerged even in the absence of co-actor leaving only the attention-attracting event present. Furthermore, in the next series of experiments (Dolk et al., 2013), they replaced the attention-attracting event with a Japanese waving cat, a clock and a ticking metronome; and found a Simon effect in all the cases. This led them to suggest the emergence of Simon effect in joint settings on the basis of referential coding - the spatial coding of one's actions relative to the other events. They finally concluded that as long as events attract attention in a bottom-up fashion and provide a spatial reference for one's own actions, Simon like effect emerges.

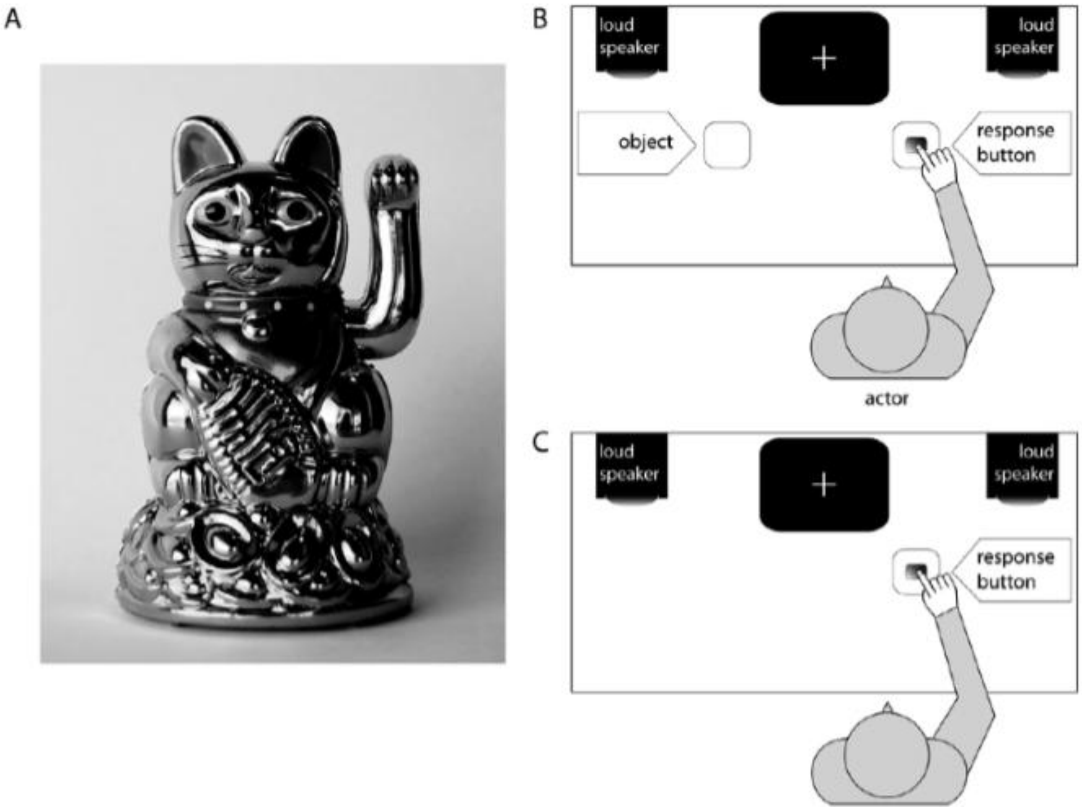


Figure 1.7: Attention-attracting object used by Dolk et al. (2013, p. 1251). A. The Japanese cat used in the experiment. B. Cat present condition. C. Cat absent condition. Reprinted from "The (not so) social Simon effect: a referential coding account," by Dolk, T., Hommel, B., Prinz, W., & Liepelt, R., 2013, *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), p. 1251. Copyright 2013 by the American Psychological Association. Reprinted with permission.

1.3 Memory for Actions: Enactment encoding and Observation encoding

1.3.1 Enactment effect on memory: 'Learning by doing' often appears superior than 'learning by viewing'. This assumption of superior learning by performing actions (i.e. hands-on experience) over other conditions has been addressed in action memory research for years and is commonly referred to as '*enactment effect*'. This assumption is applied to a wide variety of contexts from instructional design to navigation (von Stülpnagel and Steffens, 2012, 2013). In memory research, over the years it has been consistently observed that the action phrases like 'lift the pen', 'bend the wire' etc. are recalled much better when the actions are performed (i.e enacted) than when they are only read or listened to (Cohen, 1989; Engelkamp and Zimmer, 1994; Saltz and Donnenwerth-Nolan, 1981; see Engelkamp, 1998, for a review). It has been shown that the motor programs activated during enactment of action phrases become part of the memory representations of these action phrases that are not usually activated during verbal processing. The general paradigm requires subjects to remember action phrases (such as open the book, roll the ball etc.) either by enacting them in response to verbal instructions or without enacting them. Following instructions, subjects have to freely recall the verbal instructions. It was found that the memory of enacted actions was superior to that of action phrases encoded without enactment (Backman, Nilsson and Chalom, 1986; Cohen, 1981; Engelkamp and Zimmer, 1985). In a slightly different procedure, subjects were presented with verbal instructions which were encoded with or without enactment. Instead of a free recall, subjects were given a cued recall task. A cue-word from the verbal instructions was presented for each sentence and subjects had to complete the instruction by recalling the rest of the sentence. Again, it was found that enacted instructions were recalled better than those encoded without enactment (Saltz, 1988; Saltz and Donnenwerth-Nolan, 1981). Following Cohen (1981), enacted actions were referred to as subject-performed tasks (commonly known as SPTs) and non enacted actions, when participants only listened to the action phrases and tried to retain them, were referred to as verbally learned tasks (commonly known as VTs). As a general result, memory performance for action phrases was found to be better under SPTs as compared to verbal-learning condition. On the one hand, SPT advantage was said to result from the richer encoding of action events due to the activation of sensory and motor programs during enactment which lead to multimodal representation of actions (Backman and Nilsson, 1985; Backman, Nilsson and Kormi-Nouri, 1993). On the other hand, Engelkamp and colleagues focused more on the distinct motor representations produced during enactment suggesting that performing each action during verbal instruction learning might lead to the formation of a motoric (i.e. kinesthetic) code, which might help recall of enacted phrases by using additional component of motor traces (Engelkamp, 1998; Engelkamp and Zimmer, 1985, 1989, 1994).

Further experiments included an intermediate condition of action observation referred to as experimenter-performed tasks (EPTs) and reported mixed results about having a similar recall performance for subject-performed (SPTs) and Experimenter performed tasks (EPTs). A question that is more relevant to the present context is whether the observed actions (EPTs) are recalled in a similar way to the subject-performed actions (SPTs) or not. On one hand, Cohen and colleagues have consistently reported similar recall performances for SPTs and EPTs using concrete objects (Cohen, 1981, 1983; Cohen and Bean, 1983; Cohen, Peterson and Mantini-Atkinson, 1987). On the other hand Engelkamp and colleagues who used imaginary objects in the learning of verbal instructions have consistently found that SPTs are recalled better than EPTs (Engelkamp and Krumnacker, 1980; Engelkamp and Zimmer, 1983, 1984, 1989, for a review, see, Engelkamp and Cohen, 1991). These contradictory findings suggest that the studies involving actions requiring manipulation of a real object did not yield SPT advantage. To explain these differences in the findings, Steffens (1999) and Koriat et al. (1991) suggested that enactment using a real object focuses attention on the object but away from relational and contextual information. In other words, increasing attention to the unique attributes of an object (size, color, shape etc.) reduces relational processing (i.e. relations between the object and the action to be performed) in the SPT condition, which in turn, reduces enactment effect in recall tasks, producing equal recall performances for SPTs and EPTs. Furthermore, in a recent review by Steffens, Stülpnagel and Schult (2015), after comparing a number of experiments varying in stimuli and study conceptualizations suggested that a memory advantage for enactment encoding over observation encoding does not generally exist (i.e. independent of study and test conditions). They showed that an enactment effect over observation may emerge, for instance, in recognition tasks rather than recall tasks or by using a within-list design (i.e. where observation and enactment encoding is manipulated within the same list, i.e. intermixed) rather than a pure-list design (where all the actions were either performed or observed).

1.3.1.1 Recognition vs. free recall: During recognition the participants are given the verbs, objects or phrases they have learnt among distractors and they are then asked which ones were presented during study. Given that the recognition tasks are sensitive to item-specific information (specific properties of an item such as its size, shape), it was shown that the recognition of action phrases improved, if they were enacted rather than observed during study. Enactment enhances item-specific information, leading to enacted actions being better recognized than recalled (Mohr, Engelkamp and Zimmer, 1989; Engelkamp, 1986; Hornstein and Mulligan, 2004; Mulligan and Hornstein, 2003). During free recall, participants are asked to recall as many action phrases as they can out of those they have learnt during study. Free recall has been considered to be a function of relational processing in addition to item-specific processing (Hunt and Einstein, 1981; Hunt and McDaniel, 1993; McDaniel and Bugg, 2008). Although enactment improves item-specific processing, it does not generally enhance relational processing, particularly, if the relations among action phrases do not become salient during enactment (Steffens, 1999). This probably explains, at least in part, why the effect of enactment is clear in recognition tasks but less clear for free recall. For example, an SPT advantage over EPT was found in free recall by Engelkamp and Zimmner, (1983) and Dick, Kean and Sands (1989) but a similar memory performance was found for SPT and EPT by Cohen (1981, 1983; Cohen and Bean, 1983). It was speculated that both types of experimental design (i.e. recognition or free recall) and list length (use of a short or a long list) determine the pattern of results in free recall tasks (Engelkamp, 1990; Engelkamp and Zimmer, 1997). For short study lists, up to 18 items, an SPT advantage was found if the encoding condition was manipulated within subjects and there were no differences between SPT and EPT. However, for long study lists, containing more than 30 items, an advantage of SPT over EPT is generally observed regardless of the experimental design (Engelkamp and Dehn, 2000).

1.3.1.2 Pure-list vs. mixed-list design: In a pure-list design, participants either enact or observe all action phrases whereas in a mixed-list design they switch between these encoding conditions. A robust enactment effect in free recall has been found in mixed-list designs (Zimmer and Engelkamp, 1984; Engelkamp and Zimmer, 1997; Engelkamp and Dahn, 2000). One explanation is that the relational processing taking place during observation is impaired by the intervening enactment (as the observation and enactment are intermixed) which impairs free recall during observed actions. Another reason for poor recall for observed actions in mixed-list designs could be selectively displaced rehearsal that is actions carried out appear to be more important than observed actions, therefore relatively higher attention is devoted to enacted actions (Slamecka and Katsaiti, 1987; Schult and Steffens, 2011, 2013). On the other hand, in pure-list designs, enactment and observation have mostly yielded similar memory performance (for a review, see, Steffens, Stülpnagel and Schult, 2015).

1.3.2 Observation inflation effect:

Observing simple actions performed by someone else can produce a false memory phenomenon in which people falsely remember having performed an action on their own, when they merely observed someone else performing it: this is called the observation inflation effect (Lindner, Echterhoff, Davidson and Brand, 2010; also see Lindner, Schain and Echterhoff, 2016). Lindner and colleagues had their participants perform some of the action sentences they read in Phase 1. In Phase 2, participants watched a video of someone else performing the actions, some of which had not been performed by participants in Phase 1. A source-memory test was taken after a delay of 2-weeks in which participants had to indicate whether they performed the actions in Phase 1. False memories of self-performance were observed for the actions that were not performed by participants or were not presented at all in phase 1. According to the motor simulation account (Grèzes and Decety, 2001; Jeannerod, 2001; Decety and Grèzes, 2006), observing someone performing an action activates similar motor representations as performed actions; therefore motor representations of observed action would contain both sensory and motor information giving rise to the false memory of performing an action even if merely observing it (Lindner et al., 2010; Lindner, Schain and Echterhoff, 2016). It was hypothesized that, if the activation of the motor system through a simulation mechanism during action observation is fundamentally responsible for the false memories regarding self-performance, such false memories should be disrupted when the observer's motor system is occupied with a motor task during the observation. With regard to this hypothesis, Lindner, Schain and Echterhoff (2016) demonstrated that performing incongruent (vs. congruent) movements during observation reduced the observation-inflation effect. However, a more recent study by Lange, Hollins and Bach (2017) showed that preventing participants from encoding observed actions motorically by engaging their motor system in a concurrent motor task did not affect the observation-inflation effect. Lange et al. (2017) explained their findings with regard to the source-monitoring framework (Johnson, Hashtroudi and Lindsay, 1993) which suggests that the source of any information is not encoded but later inferred from the features encoded alongside the item information. Lange and colleagues (2017) suggested that both performed and observed actions create memories of events that share similarities such as body parts used, their trajectories and manner in which the task is performed. Such similarities predict the general confusion about the source of encoded actions. Linder et al. (2010, 2016) however acknowledge that even if motor simulation occurs during observation, additional processes such as consolidation of associations between observed actions and motor codes containing self-performance cues are necessary to produce self-other confusions.

1.3.3 Effect of jointly performed actions on memory (Joint Memory Effect):

While the joint action studies have typically explored the effects of acting together with others on non-memory tasks (i.e., usually reaction times measures in go-nogo tasks), there is now emerging evidence suggesting that the mechanisms of action monitoring and inhibition at play during joint task conditions could have broad consequences for higher-order cognitive systems, including memory. More specifically, a study by Eskenazi, Doerrfeld, Logan, Knoblich and Sebanz (2013) examined how acting together with another person affected the incidental encoding of information. Pairs of participants were presented with a series of words belonging to three different categories (e.g., animals, fruit/vegetables, household items). In the joint condition, each participant in a pair was instructed to press a button when the presented word belonged to one specific category (e.g., participant A responded to animals, whereas participant B responded to household items). Words of a third category (e.g., fruit and vegetables) did not require a response and served as control stimuli. Following this categorization task, participants were administered an individual surprise test in which they were asked to recall as many words as possible – regardless of whether they, their co-actors, or nobody had responded to them. The key result was that, relative to control words, recall was better not only for the words that participants had responded to themselves, but also for the words that their co-actors had responded to (an effect referred to as 'joint memory effect'). Eskenazi et al. (2013) proposed that this result was the consequence of motor simulation: that is, upon seeing the words relevant for the other, participants activated a related action plan and simulated the co-actor's motor responses, resulting in the formation of more recallable memory traces. However, a later study by Elekes et al. (2016) demonstrated that the joint memory effect (Eskenazi et al., 2013) did not depend on simulating motor responses, since the enhanced recall of the words relevant to the co-actor continued to be observed even when participants performed a non-motor task, such as counting the number of target words. Elekes et al. (2016) proposed that the presentation of the co-actor's relevant words evoked stronger response monitoring processes, as participants had to continuously prevent themselves from switching to the co-actor's task. Put in a different way, the fact that the words relevant to the co-actor were also associated to a counting task created a selection conflict: the consequence was that "it took longer for participants to decide whether they should respond to words from the other-relevant category than to words from the non-task-relevant category" (Elekes et al., 2016, p.2387).

1.4 Aim of the dissertation

The series of experiments reported in the current dissertation is aimed at investigating how performing an action (i.e. pointing movements) or observing the same action being performed by a task-partner towards to-be-remembered target locations during encoding influence one's memory for these target locations. Previous studies have shown that when encoding differs between two spatial arrays presented in a trial, where one array is encoded by mere viusal observation but the other is encoded by pointing movements performed towards item locations, their memory performance differs too. Self-performed pointing movements towards to-be-remembered target locations result into better memory performance for these locations as compared to visually-observed target locations (Chum et al., 2007). A different line of research suggests that perceiving an action can have the same functional and neural effects as performing the same action (for a review, Galantucci & Sebanz, 2009). At the neuro-physiological level it has been shown that a complex network of mirror neuron system in primates is active both while performing an action or observing it being performed by another individual (Grèzes & Decety, 2001). The first series of four experiments thus investigates whether observing pointing movements being performed by a task-partner will benefit individual's spatial memory in a similar way as self-performed movements.

Further, previous studies have shown that similar eye-movement programs are activated when performing a task and observing the same task performed by a partner (Flanagan and

Johansson, 2003). Therefore, two eye-tracking experiments in section 2.2 investigate whether similar patterns of eye movements are obtained when participants are performing the pointing movements and when they are observing them being performed by the partner. Eye-movement patterns were recorded during the observation of pointing movements performed in different encoding conditions: passive observation condition when the participant is instructed to passively observe the pointing movements performed by the task-partner and joint-action condition when both participant and the task-partner take turns to perform pointing movements.

The next three experiments investigated the boundary conditions of the findings obtained from jointly performing pointing movements in section 2.1. Experiment 1 in this section investigates whether limiting the extent to which participants are actively involved in the task could replicate the findings previously obtained. Experiment 2 and 3 investigate whether the memory effect obtained from jointly performing pointing movements is limited to a real life task-partner.

In the last experimental section, Section 2.4, in order to better understand the electrophysiological basis of the encoding processes underlying the memory advantage shown in previous experiments, with an EEG experiment we recorded the Event Related Potential components during the encoding phase, when pointing movements towards to-be-remembered target locations are performed by the participant and the task-partner by taking turns.

Overall, the aim of the current dissertation is to understand how actions performed by self and others are processed and subsequently influence individual’s memory, and the underlying mechanisms through with such actions come to influence individual’s memory.

Chapter 2: Experimental Work

2.1 I co-represent you: Effect of joint pointing movements on visuo-spatial working memory

2.1.1 Abstract: Previous studies have shown that, under specific conditions, arrays that have been pointed to during encoding are recognized better than passively-viewed ones. According to one interpretation, the superior recognition of pointed-to arrays can be explained by the motor inhibition of passively-viewed arrays. The present study sought to determine whether a similar motor inhibition can be induced also when the participants observed a task-partner perform the pointing movements. Participants were presented with two spatial arrays, one of which was encoded via observation only (the no-move array), while the other was encoded with pointing movements (the move array); movements were performed either by the participant or by the experimenter. Experiment 1 replicated the advantage of self-pointed arrays over passively-viewed arrays. Experiment 2 showed that, when participants passively observed the pointing movements performed by the experimenter (the task-partner), move arrays were recognized no better than no-move arrays. Finally, Experiment 3 and Experiment 4 demonstrated that, in a joint-action condition in which participants alternated with the experimenter task-partner in making pointing movements, the advantage of experimenter-pointed arrays over passively-viewed arrays was significant and similar in size to the advantage produced by self-performed movements. Importantly, a series of cross-experiment comparisons indicated that the higher

recognition of both self- and experimenter-pointed arrays in Experiment 3 and 4 could be explained by the motor inhibition of no-move arrays. We propose that, in a joint condition, the pointing movements performed by the experimenter were represented in the same functional way as self-performed movements and this produced the motor inhibition of passively-viewed arrays.

2.1.2 Introduction: When encoding or recalling a spatial array, being asked to simultaneously perform movements that are unrelated to the to-be-remembered items is generally detrimental for performance (see Quinn, 2008, for a review). These data are well accounted for in the Visuo-Spatial Working Memory model proposed by Logie (1995), a later version of the working memory model originally formulated by Baddeley and Hitch (1974). In Logie's account (1995), VSWM can be divided into a passive visual store (called the *visual cache*) and a movement-based spatial store (called the *inner scribe*) associated with rehearsal processes. Any movement unrelated to the to-be remembered spatial information occupies the inner scribe, leaving less resources available to rehearse spatial information, and produces a decrement in spatial memory. Most of the evidence for this hypothesis comes from dual-task studies. For instance, Vandierendonck, Kemps, Fastame, and Szmalec (2004) found that a matrix-tapping task, in which participants repeatedly tapped four keys arranged in a square, interfered with participants' performance in a computerized version of the Corsi block task (see also Rossi-Arnaud, Pieroni, Spataro, & Baddeley, 2012).

However, researchers have proposed that making pointing movements towards to-be-remembered locations might result in a more spatially-oriented perception of external objects (Fisher & Hoellen, 2004; Sapkota, Pardhan, & van der Linde, 2013). To test this hypothesis, Chum, Bekkering, Dodd, and Pratt (2007) devised a novel paradigm in which two consecutive spatial arrays were presented in each trial: one was encoded through passive visual observation (the no-move array), while the other was encoded through visual observation accompanied by pointing movements (the move array)[1]. The task required participants to maintain both arrays since they did not know in advance which array would be tested in the immediate recognition test that followed encoding at the end of each trial. Results of the recognition test showed that accuracy was significantly higher for arrays coded by observation accompanied by movements than for no-move arrays. Dodd and Shumborski (2009, Exp.1) replicated the superior recognition of pointed-to arrays when pointing instructions were manipulated within trials (i.e., when each trial included a pointed-to array and a passively-viewed array); however, they also showed that, when pointing instructions were manipulated between blocks (i.e., when the move trials required participants to touch all the items in both the arrays presented within a trial, while the no-move trials required them to passively view all the items), then pointing movements actually decreased VSWM performance (Dodd & Shumborski, 2009, Exp.2) – a result confirmed by later studies (Rossi-Arnaud, Spataro, & Longobardi, 2012; Spataro, Marques, Longobardi, & Rossi-Arnaud, 2015). Importantly for the present purposes, when Dodd and Shumborski (2009) compared the results of their first two experiments, they found that the memory advantage of pointed-to arrays in Experiment 1 could not be fully explained by the hypothesis that pointing movements increased memory performance by adding a motor code to the passive, visual code (Dodd and Shumborski, 2009). In fact, while recognition accuracy for pointed-to arrays increased from 66% when pointing instructions were manipulated between blocks (Exp.2) to 72% when pointing instructions were manipulated within trials (Exp.1), it was also apparent that the recognition accuracy for passively-viewed arrays decreased from 75% when pointing instructions were blocked (Exp.2) to 65% when pointing instructions were manipulated within trials (Exp.1). The authors concluded that at least part of the difference between move and no-move arrays in their first experiment was due to the fact that the requirement to point to one of the two arrays led to the *inhibition of motor action* for passively-viewed items, which in turn hurt memory for this type of arrays. The aim of the present study was to determine whether a similar motor inhibition occurred even when the pointing movements were observed being performed by a co-actor.

Regarding the role of observed actions, it worth noting that in the above-mentioned studies only the effects of self-performed actions on the encoding of spatial positions have been examined. However, an increasing body of research has begun to illustrate that just viewing another person's movements can have the same functional and neural effects as performing the movements by oneself (see Galantucci & Sebanz, 2009, for a review). At the neurophysiological level, the mirror neuron system (a complex network including the superior temporal sulcus, the inferior parietal lobe and the inferior frontal gyrus) is thought to play a key role in action imitation and observation, in both humans and primates (Grèzes & Decety, 2001). Moreover, it has been suggested that this system enables us to understand the intentions of a conspecific through an internal "embodied" simulation of their behavior (Enticott, Johnston, Herring, Hoy, & Fitzgerald, 2008; Jeannerod, 2001), possibly providing a powerful mechanism which might underlie effective social interactions in a large variety of domains, including empathy, theory of mind and facial emotion recognition (Iacoboni et al., 2005; Rizzolatti & Craighero, 2004). The mirror neuron system has also been involved in memory for observed actions. For example, Lindner, Echterhoff, Davidson, and Brand (2010; see also Lindner, Schain, & Echterhoff, 2016) and, more recently, Lange, Hollins and Bach (2017) investigated the so-called observation inflation effect, in which participants erroneously judge that they performed an action in a previous encoding phase while they had merely observed a confederate performing that action.

According to the Motor Simulation Account (Decety & Grèzes, 2006; Grèzes & Decety, 2001; Jeannerod, 2001), such a confusion arises because observing an action generates an internal replica that approximates the motor experience of action performance; as a consequence, the mental representations of the observed actions will contain both visual and motor/proprioceptive information, making them virtually indistinguishable from the representations of self-performed actions (Lange et al., 2017; Lindner et al., 2010).

From a broader theoretical perspective, ideomotor theories suggest that observing other's actions can induce specific tendencies to engage in the same actions, because observation leads to the activation of the same response processes that control the observer's performance (Brass, Bekkering, & Prinz, 2001). In support to this idea, there is now ample evidence indicating that, in the context of joint-action tasks, one's own actions and others' actions are represented in a functionally equivalent way in the cognitive system. Sebanz, Knoblich and Prinz (2003), for example, compared the classical individual version of the Simon task with a joint go/no-go version in which the responses were distributed between two people and concluded that the co-actor's actions were represented in the same way as one's own actions and were therefore subject to similar interference effects. This conclusion has been substantiated by later studies showing that observing co-actor's actions activated corresponding motor representations in the participants' brain (Kourtis, Sebanz, & Knoblich, 2013). In particular, Sebanz, Rebecchi, Knoblich, Prinz, and Frith (2007) and Ramnani and Miall (2004) found activations in the supplementary and ventral motor cortices during the no-go trials where participants anticipated the co-actor's actions (although note that more recent evidence has questioned the involvement of motor representations by demonstrating that the 'joint Simon effect' can be induced by any salient event that provides a spatial reference for coding one's own actions: Dolk, Hommel, Prinz, & Liepelt, 2013).

The present study was aimed at bringing together these different lines of research by investigating whether the recognition difference between move and no-move arrays occurred even when the pointing movements were performed by a task-partner (in this case, the experimenter); in addition, we were interested in determining whether observed pointing movements produced a motor inhibition of passively-viewed arrays similar to that obtained with self-performed movements (Dodd & Shumborski, 2009). To this purpose, we adopted the same methodology illustrated by Chum et al. (2007), in which each trial involved the presentation of two consecutive spatial arrays, one of which was encoded by visual observation alone (the no move array), and the other was encoded by visual observation followed by pointing movements (the move array). Importantly, in the present study the pointing movements could be performed only by the participant (in Experiment 1), only by the experimenter, the task-partner (in Experiment 2), or by the participant and the experimenter taking turns either in random (in Experiment 3) or in regular alternation (in Experiment 4). In the subsequent recognition task, memory for the move arrays pointed by participants or by the experimenter was compared with memory for the arrays encoded in the no-move condition.

To recap, previous studies suggest that: (a) when pointing instructions are manipulated within trial (e.g., when participants point to one array but passively view the other array), move arrays are recognized better than no-move arrays because self-pointing movements induce a motor inhibition of passively-viewed arrays (Chum et al., 2007; Dodd & Shumborski, 2009); (b) observing other people’s actions activates the same neuronal populations that are recruited during the execution of these actions (Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Gallese,

& Fogassi, 1996); and (c) in joint task settings, the task-partner's actions are represented just as self-performed actions and can trigger similar inhibitory effects (e.g., Galantucci & Sebanz, 2009; Sebanz et al., 2003). Thus, we hypothesized that sharing the VSWM task with the experimenter could produce a joint condition in which both the arrays pointed to by the participant and those pointed to by the experimenter should be recognized significantly better than passively-viewed arrays. In addition, if the experimenter's movements are represented just as the participant's movements in a joint condition (Sebanz et al., 2003), then we expected to observe the phenomenon of motor inhibition not only in the trials in which the pointing movements were performed by the participant, but also in the trials in which the pointing movements were performed by the experimenter.

2.1.3 Experiment 1: Self- Performed Pointing

2.1.3.1 Aim: The aim of Experiment 1 was to replicate the findings previously reported by Chum et al. (2007) and Dodd et al. (2009). In agreement with these studies, we expected to find better recognition memory for self-pointed arrays than for passively-viewed arrays.

2.1.3.2 Method:

2.1.3.2.1 Participants. Twenty-five students (19 females) of the Faculty of Medicine and Psychology of the University Sapienza of Rome volunteered to participate (age: *M* = 27.1 years).

2.1.3.2.2 Design and materials. Experiment 1 followed a 2 (array order: first vs. second) × 2 (array size: 3 vs. 4 items) × 2 (condition: no-move vs. move) design, in which all variables were manipulated within participants.

Stimuli were 192 visuospatial arrays, containing 3 or 4 items arranged in a 5 × 5 matrix which could not be seen by participants. Of these, 96 arrays contained only circles (48 for each length), whereas the other 96 arrays contained only squares. Both circles and squares were 2 cm × 2 cm in size, with two adjacent items separated by 1 cm (see Rossi-Arnaud et al., 2012). For each pattern, we constructed a test lure having all items in common with the original array except one, which was instead shifted by one or two positions. All the stimuli were presented as black outlines against a white background.

Array size was varied because, in their original study, Chum et al. (2007) reported a significant interaction between array size and condition, such that the advantage of pointed-to arrays tended to decrease as set size increased (it was non-significant for 5-item arrays).

2.1.3.2.3 Procedure. The general procedure, which mimicked that illustrated by Chum et al. (2007) and Dodd and Shumborski (2009), is outlined in Figure 1.1. Each trial comprised an encoding phase, in which two arrays of items were presented, and a test phase in which memory for the items' locations of one of the two arrays was tested. During the encoding phase, participants viewed two arrays of items, with one array including three or four white-filled squares, and the other array including three or four white-filled circles (the number of items in each array was equivalent). More specifically, each trial started with a fixation point (a centered cross) for 1000 ms, after which two spatial arrays were presented, one after the other. Both arrays contained the same number of items (either 3 or 4) but differed in shape (either squares or circles). Each item in the two arrays was presented sequentially (i.e., one by one) for 1000 ms, such that each subsequent item appeared when the previous item disappeared. The sampling of locations for each item was completely random (with the constraint that, within a given trial, no item occurred in a location previously occupied by another item), so that there was no apparent regularity in the presentation. In addition, the assignment of a square or a circle to a location was also random. There were 48 trials for each array size, for a total of 96 trials. Array size was blocked and the order of the blocks was counterbalanced across participants: half of the participants performed the 3-item trials before the 4-item trials, whereas the other half performed the task in the reverse order.

Each trial in the study phase contained both a move and a no-move array. At the beginning of the experiment, participants were shown an illustration of the procedure and were instructed to memorize the location of the items in both arrays, because their memory for one of the two arrays would be tested at the end of each trial. In addition, participants were instructed to passively view the locations of the presented items for one array (the no-move array) and to move their hands towards the locations of each item until touching the screen for the other array (the move array). Thus, the no-move array was encoded only through a passive perceptual code, whereas the move array was encoded through both a perceptual and a motor code. In half of the trials the no-move array was presented first, followed by the move array; for the other half of the trials the order was reversed. Half of the participants were instructed to tap only on the arrays containing the squares, while the other half tapped only on the arrays containing the circles (this information was provided together with the instructions, at the beginning of each experimental session).

The test phase began immediately after the presentation of the second array, starting with a black screen for 150 ms. The purpose was to examine participant's memory for the locations of the items in either the move or the no-move array. Thus, each test screen contained all the items from one of the two presented arrays (either squares or circles). Participants did not know in advance which array would be tested in each trial, meaning that they needed to memorize both arrays at study. They were instructed to judge whether the locations of the items presented at test matched those shown at encoding, by pressing the keys ‘S’ and ‘D’ for “same” and “different” responses, respectively. The time limit was set to 5000 ms. On half of the trials, the locations of the items in the test screen matched those at which they were presented during the encoding phase, while on the other half of the trials one item was shifted to a previously unoccupied location. Participants were only required to indicate that a change had occurred, without specifying which of the three (or four) locations had changed position.

2.1.3.3 Results:

For all the experiments illustrated in this study, performance was measured as the proportions of correct hits. For Experiment 1, the proportions of correct hits in each condition are illustrated in Figure 2.1. A 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) repeated measures ANOVA found:

a) a significant main effect of condition [$F(1, 24) = 11.22$, $p = 0.003$, $\eta_p^2 = 0.31$], indicating that move arrays ($M = 86.63$ %) were recognized better than no-move arrays ($M = 77.03$ %);

b) a significant main effect of array order [*F*(1, 24) = 19.92, $p < 0.001$, η_p^2 = 0.45], indicating that recognition performance was higher for arrays presented as second (*M* = 85.89 %) than for arrays presented as first (*M* = 77.76 %);

c) a significant interaction between array order and array size [*F*(1, 24) = 6.30, p = 0.019, η_p^2 = 0.21], showing that the advantage of the arrays presented as second was significant at size 4 [*M* = 87.66 % vs. *M* = 75.26 %: *F*(1, 24) = 23.38, $p < 0.001$, η_p^2 = 0.49], but not at size 3 [*M* = 84.13 % vs. *M* = 80.26 %: *F*(1, 24) = 2.56, p = 0.12, η_p^2 = 0.09].

No other effects or interactions were significant [highest *F*(1, 24) = 0.59, p = 0.44].

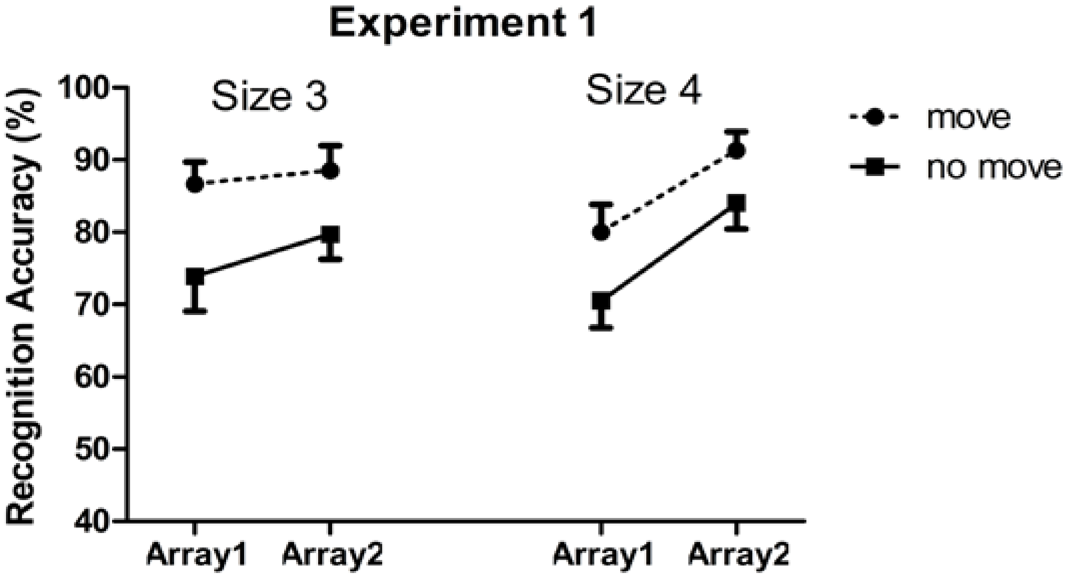


Figure 2.1: Experiment 1: mean percentages of correct hits as a function of condition (move vs. no-move), array order (array 1 vs. array 2) and array size (size 3 vs. size 4). Bars represent standard errors.

As expected, Experiment 1 replicated the results reported by Chum et al. (2007) and Dodd et al. (2009), confirming that self-pointed arrays were recognized better than no-move arrays when a subset of items was actively selected for additional processing – i.e., when each trial included a pointed-to array and a passively-viewed array

2.1.4 Experiment 2: Experimenter- Performed Pointing

2.2.4.1 Aim: Experiment 2 investigated whether the recognition difference between move and no-move arrays occurred even when the pointing movements were performed only by the experimenter, the task-partner (i.e., when participants were passive observers). Previous studies have demonstrated that observing other people's actions leads to the activation of matching motor representations in the observers' mind (Decety & Grèzes, 2006; Grèzes & Decety, 2001; Jeannerod, 2001) and recruits the same neuronal populations that are involved in the execution of those actions (Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996; Rizzolatti & Matelli, 2003). This framework predicts that a motor inhibition of passively-viewed arrays should be observed in Experiment 2, resulting in a significant memory advantage of pointed-to arrays.

On the other hand, a different set of studies indicate that the ability to simulate other's actions may depend on the social interaction between the actor and the observer, such that motor activation is less strong when participants anticipate that a particular action will be performed by a third person they do not interact with (Kourtis, Sebanz and Knoblich, 2010; Kourtis, Knoblich, & Sebanz, 2013). In Experiment 2, participants were required to be completely passive: this condition was unlikely to be represented as a meaningful social interaction. Moreover, the minimal architecture required for joint action suggests that an individual has to represent his own task for motor simulation of others' actions to occur (Vesper, Butterfill, Knoblich and Sebanz, 2010). Since participants did not perform any task in Experiment 2, it was unnecessary for them to represent the experimenter's task as well. Collectively, these results suggest that the passive observation of the task-partner's pointing movements may not produce the motor inhibition of passively-viewed arrays because participants could not represent the task-partner's actions in the same way as self-performed actions when they themselves were not actively engaged in the task.

2.1.4.2 Method:

2.1.4.2.1 Participants. Twenty naive students (14 females) of the Faculty of Medicine and Psychology of the University Sapienza of Rome volunteered to participate (age: *M* = 24.8 years).

2.1.4.2.2 Design and materials. The design and the stimuli were the same as those illustrated in Experiment 1.

2.1.4.2.3 Procedure. The general procedure mirrored that adopted in Experiment 1, with the critical exception that the items in the move arrays were pointed by the experimenter, the task-partner, sitting next to the participants (who were passive and explicitly instructed not to move their hands). The position of the experimenter (right or left) was counterbalanced across participants. Note that, while pointing movements were performed by the experimenter, the recognition task was always undertaken by the participant.

2.1.4.3 Results: The proportions of correct hits for each condition are reported in Figure 2.2. A 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) repeated measures ANOVA revealed only a significant main effect of array order [$F(1, 19) = 9.33$, $p = 0.007$, $\eta_p^2 = 0.33$], indicating that recognition performance was higher for arrays presented as second ($M = 89.58$ %) than for arrays presented as first ($M = 82.76$ %). The main effect of condition was not significant [$F(1, 19) = 1.12$, $p = 0.30$, $\eta_p^2 = 0.05$], as they were all other effects and interactions [highest $F(1, 24) = 1.31$, $p = 0.26$]. To estimate the null effect of condition, we ran the Bayesian paired-sample *t*-test which confirmed that the null hypothesis was 3.02 times more likely than the alternate hypothesis (Rouder, Speckman, Sun, Morey and Iverson, 2009).

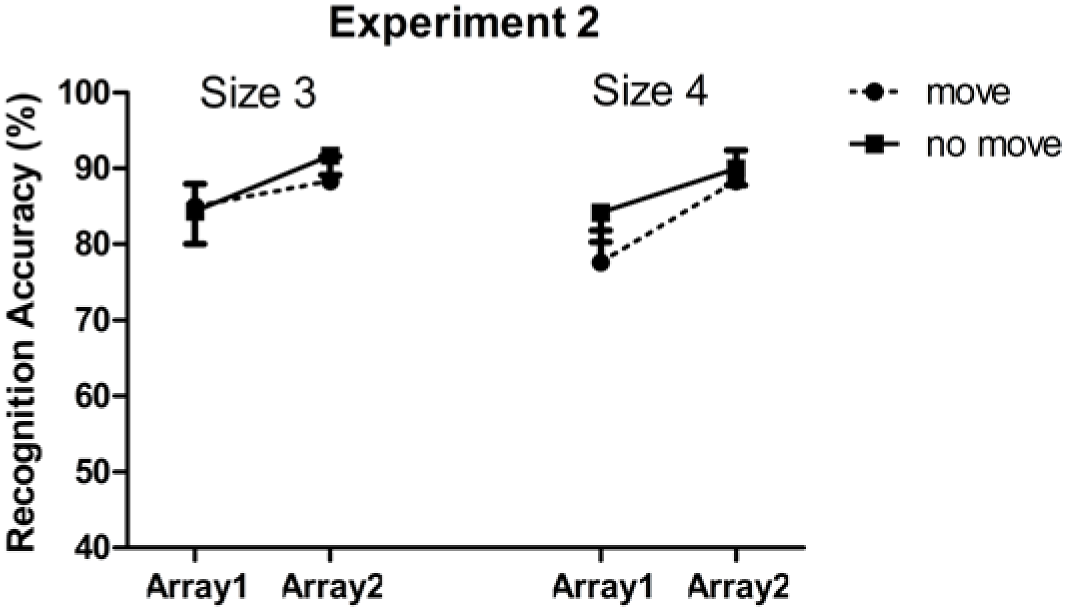


Figure 2.2. Experiment 2: Mean percentages of correct hits as a function of condition (move vs. no-move), array order (array 1 vs. array 2) and array size (size 3 vs. size 4). Bars represent standard errors.

2.1.4.4 Discussion: The results from Experiment 2 suggest that the mere observation of others' pointing movements did not result in a significant difference in the recognition of move and no-move arrays. However, given the relatively limited sample size of Experiment 2, we conducted a post-hoc analysis to determine whether the lack of a significant result was due to insufficient statistical power. The to-be-reached effect size was computed from the means and standard deviations of Experiment 1 ($d = 0.49$). Using the software G-Power 3 (Faul, Erdfelder, Lang & Buchner, 2007), we estimated that, with $N = 20$, $\alpha = 0.05$ and $r = 0.74$ (the correlation between the move and no-move conditions in Experiment 1), the power to detect an effect of condition comparable to that observed in Experiment 1 was 0.90 (*t*-test for dependent means). Thus, it appears that the present results cannot be accounted for by a reduced sample size.

2.1.4.5 Cross-Experimental Analysis: As mentioned above, one of the aims of the present study was to determine whether the memory advantage produced by self-pointing movements in Experiment 1 was due to an increase in the recognition of self-pointed arrays (Chum et al., 2007) or to a decrease in the recognition of passively-viewed arrays due to the motor inhibition resulting from the instruction to point to half of the items in a trial (Dodd & Shumborski, 2009). To clarify this issue, we performed a direct comparison between the results of Experiment 1 and Experiment 2. The assumption was that, given the absence of a significant memory difference between move and no-move arrays, the performance in Experiment 2 could be considered as an appropriate baseline against which the improvements produced by self-pointing movements in Experiment 1 could be fruitfully verified. The collapsed data were submitted to a repeated-measures ANOVA, in which array order (first vs. second array), array size (3- vs. 4-item array) and condition (move vs. no-move array) were the within-subject factors and Experiment (Exp. 1 vs. Exp. 2) was the between-subject factor. As expected, this analysis revealed a significant interaction between condition and experiment [$F(1, 43) = 9.75$, $p = 0.003$, $\eta_p^2 = 0.18$]. A follow-up analysis of simple effects showed no difference in the recognition of move arrays between the two experiments [Experiment 1: $M = 86.63\%$ vs. Experiment 2: $M = 84.81\%$; $F(1, 43) = 0.25$, $p = 0.61$, $\eta_p^2 = 0.006$], therefore suggesting that self-pointed movements did not increase spatial memory, relative to an appropriate baseline. On the other hand, the recognition of no-move arrays was significantly lower in Experiment 1 ($M = 77.03\%$) than in Experiment 2 ($M = 87.54\%$) [$F(1, 43) = 10.30$, $p = 0.003$, $\eta_p^2 = 0.19$], suggesting that the advantage due to self-pointing movements in Experiment 1 could be entirely ascribed to the negative effect of motor inhibition on the recognition of passively-viewed arrays.

2.1.5 Experiment 3: Self- Performed and Experimenter- Performed Pointing in random alternation.

2.1.5.1 Aim: Experiment 2 failed to show a better recognition of pointed-to arrays when participants were not engaged in the pointing task and were instead instructed to passively observe the pointing movements performed by the task-partner, the experimenter. Many previous joint action studies have adopted paradigms in which the task is divided between two individuals such that each participant performs a subtask which is independent of his/her task-partner, but collectively the dyad performs the full task in a social, joint-action environment (Atmaca et al., 2011; Constable, Pratt, & Welsh, 2018; Sebanz et al., 2003; Welsh et al., 2005). The logic behind these experiments is that, if individuals working independently in a joint-action environment co-represent the actions and goals of the task-partner, then the behavioral effects that occur when the participant completes the whole task alone should re-emerge in the collective performance of the task-partners. Based on this literature, in Experiment 3, we devised a joint condition in which the participant and the task-partner (the experimenter) randomly alternated in making pointing movements. Their turn was signaled by a cue presented at the beginning of each trial. We expected to replicate the finding that self-pointed arrays were recognized better than passively-viewed arrays; more importantly, we expected this advantage to extend to the arrays pointed by the experimenter. Specifically, if self- and experimenter-pointing movements were represented in a functionally equivalent way in the joint condition, then we expected the memory advantage associated to self-pointed arrays to be driven by the same motor inhibition of passively-viewed items induced by self-performed movements in Experiment 1.

2.1.5.2 Method:

2.1.5.2.1 Participants. Twenty naive students (9 females) of the Faculty of Medicine and Psychology of the University Sapienza of Rome volunteered to participate (age: *M* = 26.93 years).

2.1.5.2.2 Design and materials. Experiment 3 followed a 2 (array order: first vs. second) × 2 (array size: 3 vs. 4 items) × 2 (condition: no-move vs. move) × 2 (agent cue: participant vs. experimenter) within-subjects design. Stimuli were the same 192 arrays used in Experiment 1.

2.1.5.2.3 Procedure. The general procedure resembled that adopted in Experiment 1 and 2, with the exception that, during the encoding phase, each trial was preceded by a cue (a single uppercase letter: P for the participant, E for the experimenter) for 2000 ms, which signaled who had to point to the items in the move array for the following trial. Note that both the P- and E-cued trials included a no-move array that was never pointed; thus, the function of the cue was to inform the participant and the experimenter about who had to make a pointing movement on a given trial for one of the two arrays. The participant and the experimenter pointed to the same type of items (either squares or circles) and the assignment was decided and communicated at the beginning of each experimental session. Overall, the participants and the experimenter pointed for an equal number of trials (48, with P- and E-cued trials being randomly alternated). However, the memory test in each trial was always responded by the participant, irrespective of the agent cue.

2.1.5.3 Results:

The mean proportions of correct hits were analyzed with a 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (agent cue: P-cued vs. E-cued) repeated measures ANOVA, which found:

a) a significant main effect of condition [$F(1, 19) = 13.35$, $p = 0.002$, $\eta_p^2 = 0.41$], indicating that move arrays ($M = 79.78$ %) were recognized better than no-move arrays ($M = 66.97$ %);

b) a significant main effect of array order [$F(1, 19) = 5.55$, $p = 0.029$, $\eta_p^2 = 0.22$], indicating that recognition performance was higher for arrays presented as second ($M = 76.66$ %) than for arrays presented as first ($M = 70.10$ %);

c) a significant main effect of agent cue [$F(1, 19) = 9.36$, $p = 0.006$, $\eta_p^2 = 0.33$], indicating that the move and the no-move arrays preceded by the P cue ($M = 76.65$ %) were recognized better than the move and the no-move arrays preceded by the E cue ($M = 70.11$ %);

d) a marginal interaction between condition and array size [$F(1, 19) = 3.69$, $p = 0.070$, $\eta_p^2 = 0.16$], which is illustrated in Figure 2.3: a follow-up analysis of simple effects demonstrated that the advantage of move arrays (over no-move arrays) was significant at size 3 [$M = 81.66$ % vs. $M = 63.12$ %: $F(1, 19) = 18.86$, $p < 0.001$, $\eta_p^2 = 0.49$], but not at size 4 [although the means were in the same direction: $M = 77.91$ % vs. $M = 70.83$ %: $F(1, 19) = 2.07$, $p = 0.16$, $\eta_p^2 = 0.09$].

No other effects or interactions were significant [highest $F(1, 19) = 2.62$, $p = 0.12$]. In particular, the interaction between condition and agent cue was negligible [$F(1, 19) = 0.00$, $p = 0.96$], indicating that a significant difference between move and no-move arrays was observed for the arrays pointed to by the participants [$M = 83.12$ % vs. $M = 70.20$ %: $F(1, 19) = 14.27$, $p \leq 0.001$, $\eta_p^2 = 0.43$], as well as for the arrays pointed to by the experimenter [$M = 76.45$ % vs. $M = 63.74$ %: $F(1, 19) = 6.48$, $p = 0.020$, $\eta_p^2 = 0.25$] (see Figure 2.4).

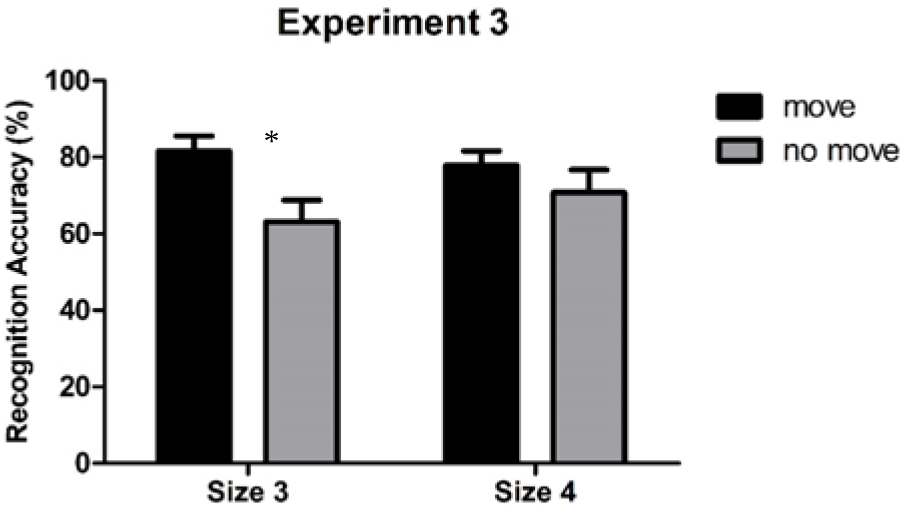


Figure 2.3. Experiment 3: Mean percentages of correct hits as a function of condition (move vs. no-move) and array size (size 3 vs. size 4). Bars represent standard errors. * indicates the significant different values, independent of the significance level.

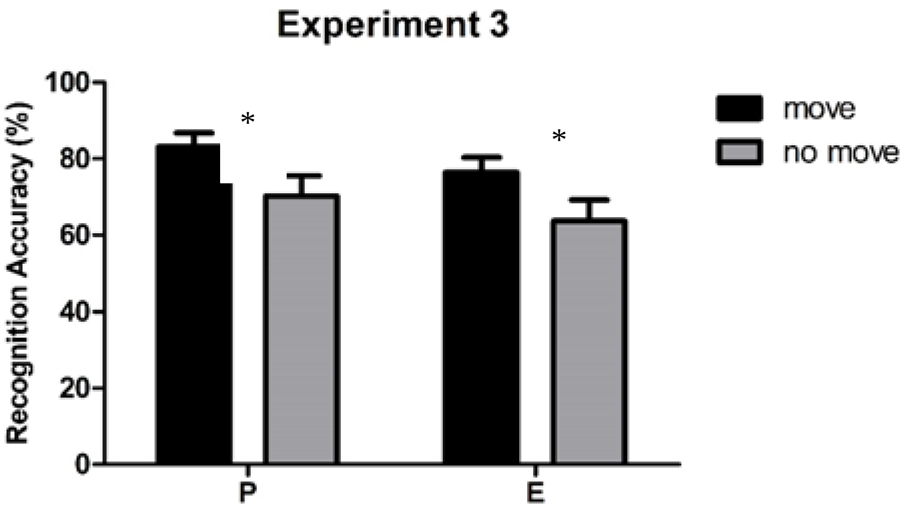


Figure 2.4. Experiment 3: Mean percentages of correct hits as a function of condition (move vs. no-move) and agent cue (P-cued vs. E-cued). Bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.1.5.4 Discussion: The results of Experiment 3 showed that, in a joint-action condition, both the arrays pointed by the participants and those pointed by the experimenter (the task-partner) were recognized better than passively-viewed arrays. Overall, these findings are in line with the idea that participants must actively engage in the task in order to represent the experimenter's pointing movements in the same way as self-performed movements (Sebanz et al., 2003, 2006; Tsai, Kuo, Jing, Hung & Tzeng, 2006). Interestingly, we also found that the move and no-move arrays encoded with the P cue (i.e., the arrays that were self-relevant) were recognized significantly better than the move and no-move arrays encoded with the E cue (i.e., the arrays that were relevant to the experimenter). This finding supports the conclusions reached by a series of previous studies showing that self-relevant materials are processed and remembered more efficiently than self-irrelevant materials (Humphreys & Sui, 2016; Sui, He, & Humphreys, 2012; Turk, Cunningham, & Macrae, 2008). It also echoes the results reported by Engelkamp and Zimmer (1997), who found that the recall of action phrases (such as "lift the pen") was significantly better if, during the encoding phase, they were performed by the subject, rather than by the experimenter.

2.1.5.5 Cross Experimental Analysis of Experiment 2 and Experiment 3:

To further determine whether the superior recognition of the arrays pointed by both the participants and the experimenter were due to the motor inhibition of no-move arrays, similarly to what was observed for self-pointed arrays in Experiment 1, the data from Experiment 3 were collapsed and analyzed together with the data from Experiment 2. Separate analyses were performed for P- and E-cued trials in Experiment 3. As concerns P-cued trials, a mixed 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (Experiment: Exp. 2 vs. P-cued trials of Exp. 3) ANOVA replicated the significant condition × experiment interaction observed in the comparison between Experiments 1 and 2 [$F(1, 38) = 13.35$, $p = 0.001$, $\eta_p^2 = 0.26$]. A follow-up analysis of simple effects confirmed that the recognition of move arrays did not differ between the two experiments [Experiment 2: $M = 84.81\%$ vs. Experiment 3: $M = 83.12\%$; $F(1, 38) = 0.15$, $p = 0.69$, $\eta_p^2 = 0.004$], whereas the recognition of no-move arrays was significantly lower in Experiment 3 ($M = 70.20\%$) than in Experiment 2 ($M = 87.54\%$) [$F(1, 38) = 9.86$, $p = 0.003$, $\eta_p^2 = 0.21$]. Thus, we again found that the higher recognition performance for self-pointed arrays was entirely explained by the motor inhibition of no-move arrays (Dodd et al., 2009). Similarly, regarding E-cued trials, a mixed 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (Experiment: Exp. 2 vs. E-cued trials of Exp. 3)

ANOVA revealed a significant interaction between condition and experiment [$F(1, 38) = 7.55$, $p = 0.009$, $\eta_p^2 = 0.16$]. The analysis of simple effects revealed a marginal tendency for the recognition accuracy of move arrays to be lower in Experiment 3 ($M = 76.45\%$) than in Experiment 2 ($M = 84.81\%$) [$F(1, 38) = 3.44$, $p = 0.07$, $\eta_p^2 = 0.08$]. The recognition accuracy for no-move arrays showed a robust, significant decrease in Experiment 3 ($M = 63.74\%$) as compared to Experiment 2 ($M = 87.54\%$) [$F(1, 38) = 17.44$, $p = 0.000$, $\eta_p^2 = 0.31$]. These findings suggest that the memory advantage observed in Experiment 3 for experimenter-pointed arrays was also accounted for by the negative effects produced by motor inhibition on no-move arrays. There was no evidence that the pointing movements performed by the experimenter facilitated spatial memory, over and above the baseline provided by Experiment 2: if anything, there was a small decrease which is consistent with the idea that, in Experiment 3, the encoding of E-cued trials was given less attentional resources than the encoding of self-relevant, P-cued trials (Humphreys & Sui, 2016; Sui et al., 2012; Turk et al., 2008). Collectively, these two analyses merge to suggest that experimenter-pointed arrays in Experiment 3 were processed in the same way as participants-pointed arrays in Experiment 1 and 3.

2.1.6 Experiment 4: Self- Performed and Experimenter- Performed Pointing in regular alternation.

2.1.6.1 Aim: We successfully established the ‘Joint- Pointing Effect’ in a shared pointing condition in Experiment 3 when participant and experimenter, the task-aprtner, took random turns to perform pointing movements towards the move arrays. It was suggested that E-cued trials were processed in the same way as P-cued trials due to the co-representation of experimenter’s pointing movements in participant’s action system, therefore no-move arrays in

E-cued trials suffered a similar inhibition as P-cued trials. It was speculated that participants must actively engage in the task in order to have a social interaction with the partner therefore representing the partner's actions as his own (Sebanz et al., 2003, 2005, Tsai et al., 2006). The joint pointing condition of Experiment 3 however involves two factors: a social interaction between the task-partners and a constant monitoring of their respective turns due to unpredictable distribution of agent-cue. Experiment 4, therefore, is designed to investigate whether the social interaction between the task-partners in a regular turn-taking condition (predictable distribution of agent-cue) can replicate the memory advantage obtained for the experimenter's pointing movements in the E-cued trials in Experiment 3.

2.1.6.2 Method:

2.1.6.2.1 Participants: Twenty three students (12 Females, M = 23.04 years, SD = 1.71 years) of the Faculty of Medicine and Psychology, Sapienza University of Rome, volunteered to participate in the experiment. All participants had either corrected or corrected to normal vision.

2.1.6.2.2 Design and materials: Experiment 1 followed a 2 (array order: first vs. second) x 2 (array size: 3 items vs. 4 items) x 2 (condition: move vs. no- move) x 2 (agent cue: P-cued vs. E-cued) within subject design.

Stimuli were the same 192 visuospatial arrays as used in Experiment 2.2.3.

2.1.6.2.3 Procedure: The general procedure mimicked that was used in Experiment 3 except that the distribution of agent-cue (P and E) was regular (i.e. predictable) instead of random (i.e. unpredictable). Pointing movements were performed in a regular alternation (e.g. P-E-P-E-P-E so on) by the participant and experimenter such that each agent knows exactly when it is her/his turn to perform pointing movements towards the move arrays. Each agent performs the pointing movements on an equal number of trials (i.e. 48 trials each). The memory task was always answered by the participant.

2.1.6.3 Results:

A 2 (array order: first vs. second) x 2 (array size: 3 items vs. 4 items) x 2 (condition: move vs. no- move) x 2 (agent cue: P-cued vs. E-cued) repeated measures ANOVA revealed (see **Figure 2.5**),

(a) A significant main effect of condition [F(1, 22) = 5.55, $p = 0.02$, $\eta_p^2 = 0.20$], indicating that the recognition accuracy for move arrays was found to be better (M = 83.15%) than the no-move arrays (M = 70.28%).

(b) A significant interaction was found between array order and array size [F(1, 22) = 5.20, $p = 0.03$, $\eta_p^2 = 0.19$], a follow up analysis showed that arrays with size 3 were remembered better when presented at last (M = 79.70%) as compared to when they were presented first (M = 69.92%), [F(1, 22) = 5.47, $p = 0.02$, $\eta_p^2 = 0.19$]. Arrays with size 4, were remembered better when presented first (M = 78.98%) than arrays with size 3 (M = 69.92%), [F(1, 22) = 4.60, $p = 0.04$, $\eta_p^2 = 0.17$].

(c) A marginally significant three way interaction was found among array order, condition and agent cue [F(1, 22) = 3.72, $p = 0.06$, $\eta_p^2 = 0.14$], follow up analysis revealed that arrays pointed by the participant were remembered better (M =

90.57%) than arrays pointed by experimenter (M = 83.33%) when they were presented at last [F(1, 22) = 4.86, *p* = 0.03, $\eta_p^2 = 0.18$].

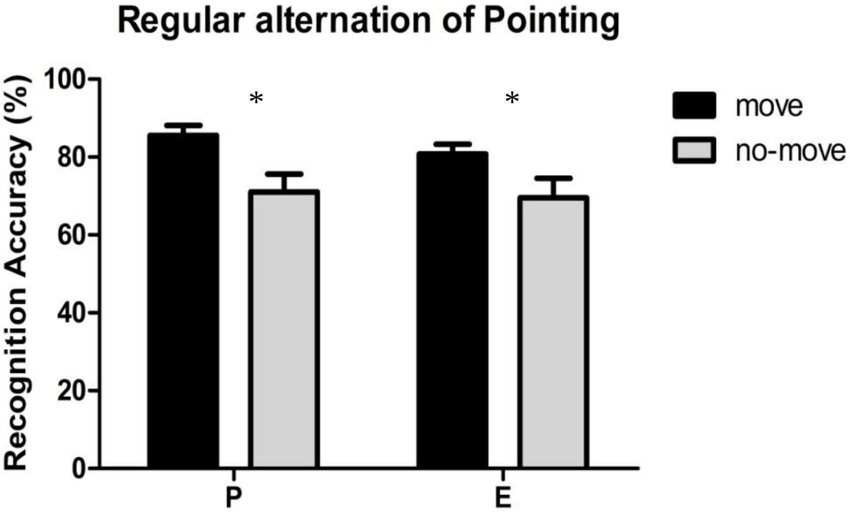


Figure 2.5: Experiment 1: An interaction between condition (move vs. no-move) and agent cue (P vs. E). Error bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.1.6.4 Discussion: Experiment 4 was aimed at investigating whether a predictable turn-taking condition to perform pointing movements towards the move arrays could replicate the memory advantage obtained for the experimenter's pointing movements in Experiment 3. The predictable turn-taking condition of pointing movements involves the social interaction between the task-partners but rule out the possibility of continuous monitoring of their respective turns as they know in advance when is their turn to point. The pattern of results suggests that the memory advantage for experimenter's pointing movements is observed even when the participant and experimenter take regular turns to perform pointing towards move arrays. The current findings support the assumption made by previous studies that in a joint-action condition participants work collectively in a social environment, therefore they represent each other's actions as their own (Sebanz et al., 2003).

2.1.6.5 Cross Experimental Analysis of Experiment 2 and Experiment 4:

To determine whether the no-move arrays in the trials pointed by both the participant and experimenter, suffered a similar motor inhibition as Experiment 3, the data from Experiment 4 were collapsed and analyzed together with the data from Experiment 2. Separate analyses were performed for P- and E-cued trials in Experiment 4. With regard to P-cued trials, a mixed 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (Experiment: Exp. 2 vs. P-cued trials of Exp. 4) ANOVA replicated the significant condition × experiment interaction [$F(1, 40) = 7.76$, $p = 0.008$, $\eta_p^2 = 0.16$]. A follow-up analysis of simple effects confirmed that the recognition of move arrays did not differ significantly between the two experiments; however being numerically higher in Experiment 4 [Experiment 2: $M = 84.81\%$ vs. Experiment 4: $M = 86.36\%$; $F(1, 40) = 0.20$, $p = 0.65$, $\eta_p^2 = 0.005$], whereas the recognition of no-move arrays was significantly lower in Experiment 4 ($M = 70.45\%$) than in Experiment 2 ($M = 87.54\%$) [$F(1, 40) = 10.68$, $p = 0.002$, $\eta_p^2 = 0.21$]. Thus, we again found that the higher recognition performance for self-pointed arrays was entirely explained by the motor inhibition of no-move arrays (Dodd et al., 2009). Similarly, regarding E-cued trials, a mixed 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (Experiment: Exp. 2 vs. E-cued trials of Exp. 4) ANOVA revealed a significant interaction between condition and experiment [$F(1, 41) = 4.46$, $p = 0.04$, $\eta_p^2 = 0.09$]. The analysis of simple effects revealed a marginal tendency for the recognition accuracy of move arrays to be lower in Experiment 4 ($M = 80.79\%$) than in Experiment 2 ($M = 84.81\%$) [$F(1, 41) = 1.41$, $p = 0.24$, $\eta_p^2 = 0.03$]. The recognition accuracy for no-move arrays showed a robust, significant decrease in Experiment 4 (*M* = 69.56%) as compared to Experiment 2 (*M* = 87.54%) [$F(1, 41) = 10.48$, $p = 0.002$, $\eta_p^2 = 0.20$]. These findings suggest that the memory advantage observed in Experiment 4 for experimenter-pointed arrays was also accounted for by the negative effects produced by motor inhibition on no-move arrays. Collectively, these two analyses suggest that experimenter-pointed arrays in Experiment 4 were processed in the same way as self-pointed arrays by participants.

2.1.7 General Discussion: The aim of the present study was to investigate whether, in a joint setting, pointing movements performed by a task-partner produce the same memory consequences as self-performed movements (Galantucci & Sebanz, 2009; Sebanz et al., 2003). To this purpose, we adopted a VSWM paradigm used in previous studies (Chum et al., 2007; Dodd et al., 2009) in which, for each trial, there is one array in which items should be observed and pointed to (the move array) and one array for which items should be passively observed (the no-move array). The usual finding for arrays with 3 or 4 elements (not with 5 or above) is that performance in a subsequent old/new recognition test is significantly better for move than for no-move arrays. We successfully replicated this result in Experiment 1. In Experiment 2 we examined whether passive observation of pointing movements performed by a task-partner, here the experimenter, produced the same memory advantage as that observed in Experiment 1 for self-performed pointing movements. Results showed that this was not the case, since recognition accuracy for the arrays pointed by the experimenter was no better than recognition accuracy for passively-viewed arrays. These findings suggested the possibility that the direct experience of the motor consequences of self-performed movements was necessary in order to represent the task-partner's task and thus observe a significant pointing advantage (Vesper et al., 2010). For this reason, in Experiment 3 and 4 we introduced a joint setting version of the task (Sebanz et al., 2003; Welsh et al., 2005), in which the participant and the experimenter took turns in making the pointing movements. In this joint condition, the analyses revealed that both the arrays pointed by the participants and those pointed by the experimenter were recognized better than passively-viewed arrays.

Our data are especially relevant for the understanding of the mechanisms through which pointing movements affect VSWM in the paradigm devised by Chum et al. (2007). Dodd and Shumborski (2009) had already pointed out that, when participants were required to point to items in one array while passively viewing the items of another array within the same trial, the recognition of passively-viewed arrays was inhibited, relative to trials in which all the items were either pointed to or passively viewed. In the present study, we took the memory performance in Experiment 2 as a baseline, because we observed no significant difference between the recognition of passively-viewed and experimenter-pointed arrays. We thus reasoned that, in this experiment, memory for passively-viewed arrays was uncontaminated by the negative effects of the motor inhibition reported by Dodd & Shumborski (2009). When the results of Experiments 1 and 2 were directly compared, we found that recognition accuracy for self-pointed arrays in Experiment 1 did not exceed the baseline level reached in Experiment 2. In contrast, we observed that recognition accuracy for passively-viewed arrays in Experiment 1 was significantly worse than the mean level observed in Experiment 2. This result points to the conclusion that the memory advantage produced by self-performed pointing movements in Experiment 1 was due to the motor inhibition of passively-viewed arrays resulting from the instruction to point to half of the items within a trial (Dodd & Shumborski, 2009). Importantly for the present purposes, we ran similar analyses comparing levels of performance in Experiment 2 and Experiment 3. These analyses demonstrated that the same mechanisms were at play in

Experiment 3 and could account for why the arrays pointed in turn by the participants and the experimenter were recognized better than passively-viewed arrays. A similar comparative analyses was also ran for Experiment 4, which showed similar patterns of motor inhibition for passively-viewed arrays in both types of trials, one pointed by participant and the one pointed by the task-partner, here the experimenter. It is worth noting that results from Experiment 3 and 4 provide strong support to the hypothesis that, in a joint setting, the experimenter's movements were represented just as the participant's movements and therefore produced comparable inhibitory effects on passively-viewed arrays (Sebanz et al., 2003): the pattern of behavioural findings when the participant completed the whole task alone (Experiment 1) is quite similar to that which emerges when participant and task-partner alternate in making the pointing movements (Experiment 3 and 4). These data help outline the processes by which self- and other-performed movements come to influence memory.

As mentioned above, one potential explanation for the recognition advantage of self-pointed arrays was that pointing movements enriched the visual memory trace with motor and kinesthetic information (Denis, Engelkamp, & Mohr, 1991; Engelkamp & Zimmer, 1989, 1997). Chum and colleagues (2007), in particular, assumed that self-performed pointing movements generated an egocentric (body-based) code which was integrated and retained together with allocentric (scene-based) information in the parietal cortex (Burgess, 2006; Save & Moghaddam, 1996; Zaehle et al., 2007). In the context of the present study, the idea that self-performed movements benefitted the encoding of move arrays was directly contradicted by the results of the cross-experiment comparisons. These analyses indicated that the pointing movements performed by participants in Experiment 1 and in the P-cued trials of Experiment 3 and 4 did not increase recognition accuracy beyond the baseline level provided by Experiment 2.

Importantly, our data also contradicts the hypothesis that the significant advantage of move arrays in the E-cued trials of Experiment 3 and 4 was due to the fact that participants simulated the pointing movements performed by the experimenter and this allowed them to encode the spatial locations by using both egocentric and allocentric representations (Chum et al., 2007). Indeed, there was no evidence that the pointing movements performed by the experimenter in the E-cued trials of Experiment 3 and 4 enhanced recognition memory above the baseline levels of Experiment 2. Taken together, these results do not support the general view that performing active movements, or simulating them, during the encoding phase lead to the formation of more recallable memory traces (Chum et al., 2007).

Furthermore, a strong version of the motor simulation hypothesis maintains that action observation, like imagination, should activate matching motor representations in the observers' mind (Decety & Grèzes, 2006; Grèzes & Decety, 2001; Jeannerod, 2001; Lindner et al., 2010, 2016). This account predicts that a significant memory advantage for the arrays pointed by the experimenter should have been obtained in Experiment 2, when the participants were passive observers – a prediction which is inconsistent with the present results. Our explanation is that participants in Experiment 2 did not represent the movements performed by the experimenter because the requirement to be completely passive implied that there was minimal personal involvement on the observer's side. In line with this proposal, previous studies have found that motor resonance for observed human actions is critically dependent on the extent to which the experimental condition can be understood as a social interaction (Hogeveen & Obhi, 2012; Kourtis, Sebanz & Knoblich, 2010, 2013). In particular, two electrophysiological studies by Kourtis and colleagues (2010) tested whether motor activation was stronger when people anticipated the actions of an interaction partner compared with the same actions performed by a ‘loner’ (i.e., a person who always acted alone). The participant, the interaction partner, and the loner were sat around a table and the task was to lift an object either returning it to its original position (individual action) or passing it to another person/receiving it from another person (joint action). The participant engaged with the ‘partner’ in joint actions, whereas the ‘loner’ always acted alone. As a measure of anticipatory motor simulation, Kourtis et al. (2010) focused on the amplitude of the contingent negative variation (CNV), a late component of the EEG which reflects activity in the supplementary and primary motor cortices. As predicted, the amplitude of the late CNV was more pronounced when participants anticipated the partner’s actions than when they anticipated the loner’s actions. Of special importance, the late CNV for the loner’s actions did not differ from the baseline motor activation in the no-go condition, in which all three individuals were instructed not to act. These results indicate that simulation of another person’s actions, as reflected in the activation of motor cortices, occurs only when the other is perceived as an interaction partner. The conditions of Experiment 2 in the present study were unlikely to be interpreted by participants as a meaningful social interaction, because they were instructed to be passive and not to interact with the experimenter. Rather, it seems plausible that the experimenter was perceived in the same way as the loner in the study by Kourtis et al. (2010): if this were the case, then the finding that our participants did not simulate the experimenter’s pointing movements comes to little surprise.

In sharp contrast, the results obtained in Experiment 3 and 4 indicate that the pointing movements performed by the experimenter in a joint-action condition produced a robust motor inhibition of passively-viewed arrays, which was comparable in size to that produced by self-performed pointing movements. In agreement with many previous studies, we propose that the arrays pointed by the experimenter were represented in the same functional way as self-pointed arrays, because participants experienced the joint task as a social interaction and therefore co-represented their partner's pointing movements as if they were their own (Galantucci & Sebanz, 2009; Sebanz et al., 2003, 2006). There is now good evidence to indicate that the generation of a co-representation involves the activation of the motor areas of the brain during the no-go trials in which the non-acting participant anticipates the task-partner's actions (Holländer, Jung, & Prinz, 2011; Kourtis et al., 2013; Ramnani & Miall, 2004; Sebanz et al., 2007). In some cases, this covert simulation has been shown to have positive consequences on memory. For example, Eskenazi and colleagues (Eskenazi, Doerrfeld, Logan, Knoblich & Sebanz, 2013) examined a phenomenon known as the *joint memory advantage*, in which the words responded to by a task-partner were recalled significantly better than non-task-relevant words (i.e., words to which neither the participant nor the co-actor responded). The authors proposed that this result was the consequence of motor simulation: that is, upon seeing the other-relevant exemplars, participants simulated the task-partner's motor responses, resulting in the formation of more recallable memory traces. As noted above, the data we obtained in Experiment 3 and 4 show a very different pattern of results, in which the simulation of the experimenter's movements produced a motor inhibition of passively-viewed arrays which in turn worsened the encoding and subsequent recognition of these stimuli (Dodd & Shumborski, 2009).

It should be noted that the inhibition of no-move array to which we refer in the present study is different from the way in which this construct has been typically used in joint-action studies. In that context, the usual assumption is that participants need to actively inhibit the motor representations of the actions performed by the co-actor if they want to prevent themselves from switching to the partner's task. Accordingly, an increasing number of studies have reported the involvement of monitoring and inhibition processes in joint action tasks (Sebanz et al., 2007;

Leynes & Kakadia, 2013; Tsai et al., 2006). Both Sebanz, Knoblich, Prinz and Wascher (2006) and Tsai et al. (2006) proposed that, when a stimulus requiring an action from the task-partner appeared on the screen, a related action representation was simultaneously activated in the participants' mind. Thus, additional control and monitoring was required to suppress the increased activation following action's anticipation (see also Sebanz et al., 2007). In the present study, the finding that recognition performance was better for move than for no-move arrays is apparently in contrast with the idea that the arrays pointed by the experimenter should be actively inhibited by participants. However, other results are more consistent with this view. First, the overall recognition of the move and no-move arrays preceded by the E cue was significantly worse than the recognition of the arrays preceded by the P cue, suggesting that the E-cued arrays suffered a global inhibition. Second, the cross-experiment comparison indicated that the recognition accuracy of move arrays tended to be lower in the joint conditions of Experiment 3 and 4 than in the passive observation condition of Experiment 2, again suggesting that the representations of these arrays were partially inhibited (although the magnitude of the inhibition was not as strong as that occurring on no-move arrays).

We might speculate about the origins of this motor inhibition. In particular, anatomical, functional and clinical data indicate that the dorsal visual system, classically associated with the processing of spatial information (Goodale & Milner, 1992), can be subdivided into two distinct sub-systems: the dorso-dorsal stream and the ventro-dorsal stream (Rizzolatti & Matelli, 2003). The dorso-dorsal stream is formed by areas of the superior parietal lobule and is responsible for the on-line control of actions such as reaching and touching objects, while the ventro-dorsal stream, formed by the visual areas of the inferior parietal lobule, is involved in space perception and action understanding. When, within the same trial, participants have to point to the move array but refrain from pointing to the no-move array, this might produce a conflict in the information processed in the two streams: the visual information concerning the no-move array (elaborated in the ventro-dorsal stream) does not fit the motor information concerning the move array (elaborated in the dorso-dorsal system), and this might led to the selective inhibition of the no-move array. As suggested by Dodd and Shumborski (2009), the underlying idea is that the motor encoding of half of the items in a trial necessitates the active inhibition of related competitors encoded via passive observation, making them less accessible at a later time. This is the same process that occurs in the retrieval-induced forgetting paradigm (Anderson, Bjork, & Bjork, 2000), where the act of recalling half of the items for some categories inhibits the recall of unpracticed, but encoded, items from these same categories (though note that inhibitory processes in our study are hypothesized to operate at encoding, rather than at retrieval). Since the order of the move and no-move arrays was counterbalanced across trials, the implication is that, in the trials in which the pointed-to array came second, inhibition could target a previously formed representation of the passively-observed array – just as, in the retrieval-induced forgetting paradigm, recalling some items in the practice phase can inhibit the recall of the representations of unpracticed items formed in a previous encoding phase. When interpreted in this perspective, the difference between the results of Experiments 1, 3 and 4 and those obtained in Experiment 2 might indicate that the inhibitory effects of a directly ‘experienced’ conflict were stronger than those of a ‘perceived’ conflict: this is because in the passive observation conditions of Experiment 2 the conflict was experienced by the task-partner who performed the pointing movements, but not by the participant who merely observed them.

In summary, the present study adds to the joint action literature by showing that, when the participants and the experimenter alternated in making pointing movements, a significant memory advantage for move arrays (over no-move arrays) was observed for both self- and experimenter-pointed arrays. On the other hand, the same advantage did not arise when participants were merely required to observe the pointing movements performed by the experimenter. On the basis of cross-experimental evidence, we propose that the benefit accruing to the stimuli pointed by the experimenter in the joint conditions of Experiment 3 and 4 did not result from covert motor simulation enriching the encoding of passively-viewed arrays with the addition of a motor code (Chum et al., 2007); instead, our data can be better accounted for by the hypothesis that, in a joint-action context, the pointing movements performed by the experimenter were represented and simulated by the participants as if they were their own and this resulted in a robust inhibition of passively-viewed arrays which hurt the recognition of these stimuli (Dodd & Shumborski, 2009). Because this is essentially the same mechanism which has been shown to underlie the memory advantage of self-pointed arrays (see Dodd & Shumborski, 2009, and the results of the cross-comparison between Experiments 1 and 2 in the present study), our data provide further support to the idea that working with others has a profound impact on the way in which we encode and remember the stimuli that are relevant to the task-partners (Sebanz et al., 2003; Sebanz, Bekkering, & Knoblich, 2006).

2.2 Your actions are my actions: An eye-movement study on the joint effects of pointing in visuo-spatial working memory

2.2.1 Abstract: Earlier experiments showed that pointing movements performed by a task-partner under joint condition benefit individual's visuo-spatial working memory similarly to self-performed pointing movements. In each trial, two consecutive spatial arrays were presented, where one array was encoded by mere visual observation (no-move array) while the other was encoded by observation and pointing movements (move array). In two eye-tracking experiments, we investigated the eye movement patterns for these two types of arrays (encoded by observation accompanied by pointing movements or encoded by mere visual observation). In Experiment 1, in the move arrays pointing movements were performed by either the participant or the task-partner, here the experimenter, in the joint condition, with participant and experimenter taking turns. In Experiment 2, pointing movements were performed only by the task-partner and participants were instructed to passively observe the pointing. The results of a later recognition task showed that move arrays pointed by both the participant and the experimenter were recognized better than no-move arrays in the joint condition of Experiment 1. In contrast, no difference between move and no- move arrays was found in Experiment 2. The patterns of eye movements are in line with the behavioral findings and showed that fixations to move arrays were higher in number and longer in duration than the fixations to no-move arrays in Experiment 1. No difference in eye movement patterns was found for the move and no-move arrays in Experiment 2. We propose that, in the joint-action condition, self and others' actions are coded at

the same representational level and their functional equivalency is reflected in a similar pattern of eye fixations.

2.2.2 Introduction: In the working memory model originally proposed by Baddeley and Hitch (1974), the visuospatial sketch pad (VSSP) represented the subsystem specifically devoted to the elaboration and maintenance of visual and spatial information (see Baddeley, 2012, for a review). Logie (1995) divided this sub-system into a passive visual cache and a movement-based inner scribe associated with rehearsal processes and called the whole system visuo-spatial working memory (VSWM). A large body of research has investigated how movements interact with encoding in VSWM and most studies using the dual-task paradigm reported significant detrimental effects when the performed movements were unrelated to the to-be-remembered stimuli (see Quinn, 2008, for a review). While the negative effects of unrelated movements on VSWM are well recognized in literature, other researchers have shown that pointing movements can have positive consequences on memory when they are directed towards the to-be-remembered stimuli and the pointing instructions are manipulated *within* trials. As illustrated in section 2.1, Chum, Bekkering, Dodd, and Pratt (2007) developed a novel paradigm in which each trial in the study phase involved the presentation of two consecutive spatial arrays of circles and squares; one array was encoded through passive visual observation (the *no-move* array), while the other was encoded through visual observation accompanied by pointing movements (the *move* array). After a brief interval, participants performed an immediate recognition task in which they had to decide whether a probe array of circles or squares matched or not one of the two arrays previously encoded. The results demonstrated that accuracy was significantly higher for move than for no-move arrays. To account for this advantage, Chum and colleagues (2007) proposed that pointing movements enriched visual memory with motor and kinesthetic information, leading to the formation of more recallable traces (Denis, Engelkamp, & Mohr, 1991; Engelkamp & Zimmer, 1989, 1997). More precisely, they assumed that self-performed pointing movements generated an egocentric (body-based) code which was integrated and retained together with allocentric (scene-based) information in the parietal cortex (Burgess, 2006; Save & Moghaddam, 1996; Zaehle et al., 2007). A later study by Dodd and Shumborski (2009) replicated the positive influence of pointing and additionally showed that, if a subset of items was not selected for action (i.e., if pointing instructions were manipulated *between* trials, such that participants either pointed to all items or passively observed them within a single trial), then pointing movements decreased VSWM performance, a finding confirmed by Rossi-Arnaud, Spataro, and Longobardi (2012). Most importantly for the present purposes, when Dodd and Shumborski (2009) directly compared the experiments in which pointing instructions were manipulated within or between trials, they found that the advantage for pointed-to arrays reflected the combination of two factors: an enhancement in the recognition of move arrays (from 66% when pointing instructions were manipulated between trials to 72% when pointing instructions were manipulated within trials), plus a decrease in the recognition of no-move arrays (from 75% when pointing instructions were manipulated between trials to 65% when pointing instructions were manipulated within trials). Taken together, these data demonstrate that a) pointing movements benefit VSWM only when a subset of items is actively selected for additional processing, and b) the requirement to point to a subset of items leads to the *motor inhibition* of no-move arrays, which are therefore recognized worse than move arrays (Dodd & Shumborski, 2009).

We, in earlier experiments (section 2.1), used the paradigm introduced by Chum et al. (2007) to determine whether the observation of the pointing movements performed by a task-partner, in our case the experimenter, could produce the same positive memory consequences as self-performed movements. Collectively, the results of this set of experiments indicated that the pointing movements performed by the task-partner were as beneficial to immediate recognition as self-performed movements when the two types of actions were alternated in a joint-action condition. In fact, we found that, when participants simply observed the pointing movements performed by the task-partner (Experiment 2: the *passive* condition), move arrays were recognized no better than no-move arrays. However, when the participant and the task-partner alternated in performing pointing movements (Experiment 3 and Experiment 4: the *joint-action* condition), then both the arrays pointed by the participant and those pointed by the task-partner were recognized better than no-move arrays. Furthermore, when these two experiments were directly compared with Experiment 2 having a passive observation condition, the results showed that the memory advantage accruing to the arrays pointed by the participant and the task-partner in the joint-action condition was almost entirely due to the motor inhibition of no-move arrays (for which recognition was significantly worse in the joint-action than in the passive condition). We proposed that the arrays pointed by the task-partner in the joint setting were represented in the same functional way as self-pointed arrays, because participants experienced that condition as a social interaction and, therefore, co-represented their partner’s pointing movements as if they were their own; this, in turn, produced comparable inhibitory effects on the no-move arrays. The results reported earlier are intriguing in that they echo previous studies demonstrating that working in a joint-action condition leads participants to form motor representations specifying what actions the task-partner will perform under specific conditions (Sebanz, Bekkering, & Knoblich, 2006; Sebanz, Knoblich, & Prinz, 2005). Interestingly, this co-representation occurs not only when the participant and the task-partner need to coordinate and synchronize their movements, but also when the participant has to perform tasks merely in parallel with, but independently of, the task-partner (as was the case in our previous experiments). In the first demonstration of this phenomenon, that we have previously mentioned, Sebanz, Knoblich and Prinz (2003) used a modified version of the Simon task in which participants responded to stimuli having both a task-relevant feature (i.e., the color of a ring) and a task-irrelevant feature (i.e., the direction of the finger wearing the ring). Participants working in the individual condition were asked to respond to the relevant feature by pressing either a right key (when the ring was green) or a left key (when the ring was red), while simultaneously ignoring the irrelevant dimension. In line with previous findings, Sebanz et al. (2003) found that the irrelevant feature (i.e., the finger's direction) was automatically processed and therefore activated the corresponding spatial representations: the consequence was that RTs were significantly delayed on trials in which the spatial direction suggested by the irrelevant feature was incongruent with the response dictated by the relevant feature. Importantly, the same incongruence effect emerged in a joint-action condition in which the task was distributed between two participants, such that the subject on the right side responded to the green ring, while the subject on the left side responded to the red ring.

According to Sebanz et al. (2003), the joint Simon effect occurs because participants co-represented the actions performed by their task-partner in the same way as their own actions. This means that when the finger (i.e., the irrelevant feature) pointed to the task-partner, the motor plan of the corresponding spatial action was activated not only in the task-partner's mind, but also in the participant's mind, and therefore interfered with the response required by the relevant feature. In other words, the finger pointing at the task-partner activated a representation of the other's action, which interfered with the planning of one's own action. The co-representation hypothesis is now supported by ample evidence indicating that observing others' actions can have similar functional and neural effects as performing the same actions on one's own (Prinz, 1997). For example, the *motor simulation account* assumes that the observation of others' actions generates an internal 'replica' that approximates the motor experience of performed actions (Jeannerod, 2001; Decety & Grézes, 2006). Similarly, the *direct matching hypothesis* proposes that action understanding relies on a mechanism that maps observed actions onto the motor representations of those actions (Iacoboni et al., 1999; Rizzolatti, Fogassi, & Gallese, 2001). Several findings confirm these views. First, researchers have identified a complex mirror neuron system in the parietal lobe that enables us to understand and imitate observed actions through an internal 'embodied' simulation that matches action observation with action execution (Enticott, Johnston, Herring, Hoy, & Fitzgerald, 2008; Iacoboni et al., 2005; Rizzolatti & Craighero, 2004; Rizzolatti, Fadiga, Fogassi, & Gallese, 1999). Second, it has been found that the motor cortex exhibits corresponding neural activity when participants perform specific actions and when they simply observe the same actions being performed by others (Buccino et al., 2001). Finally, both ERP and fMRI studies investigating the neural underpinnings of the joint Simon effect reported that portions of the motor cortex were activated by stimuli to which the task-partner, but not the participant, was to respond (Kourtis, Sebanz, & Knoblich, 2013; Sebanz, Knoblich, Prinz, & Wascher, 2006; Sebanz, Rebbechi, Knoblich, Prinz, & Frith, 2007; Tsai, Kuo, Jing, Hung, & Tzeng, 2006).

While earlier studies examining the issue of co-representation have typically used simple motor tasks coupled with behavioral measures of interference/conflict resolution (Atmaca, Sebanz, & Knoblich, 2011; Sebanz et al., 2003), more recent investigations have begun to explore the effects of motor simulation on memory encoding. For instance, as described in previous chapter (section 1.3), Eskenazi, Doerrfeld, Logan, Knoblich, and Sebanz (2013) had participants performing a categorization task either alone or together with a partner, followed by a surprise recall test. The words included in the encoding list belonged to three different categories, such that, in the joint encoding condition, each participant in a pair pressed a prespecified key when he/she saw words of one category (e.g., one person responded to animals, the other to household items). Words of a third category did not require a response and served as control items. The key finding reported by Eskenazi and colleagues (2013) was that both the words responded to by the participants and those responded to by their task-partner were recalled better than control words. The authors proposed that this '*joint memory effect*' resulted from the motor simulation of the task-partner's responses, which in turn led to the formation of more recallable memory traces. Another intriguing phenomenon which has been thought to originate from interpersonal motor simulation is the *observation inflation effect* (Lindner, Echterhoff, Davidson, & Brand, 2010; Lindner, Schain, & Echterhoff, 2016). In this paradigm, participants erroneously judge that they performed an action in a previous encoding phase while they had merely observed a confederate performing that action. According to the motor simulation account, such a confusion occurs because participants covertly imitate the actions performed by the task-partner; the consequence is that the memory traces of the observed actions will contain both visual and motor/proprioceptive information, making them virtually indistinguishable from the traces of self-performed actions (Lange, Hollins, & Bach, 2017).

Drawing on these theoretical backgrounds, the results reported earlier could be readily understood as being the result of the co-representation of the actions performed by the experimenter. Specifically, we proposed that, in the joint-action condition, the presentation of other-relevant arrays (i.e., the arrays to which the experimenter had to point) triggered the anticipatory motor simulation of the experimenter's pointing movements (Kourtis, Sebanz, & Knoblich, 2013). As stated above, this simulation produced a strong motor inhibition which worsened the encoding and subsequent recognition of no-move arrays. On the other hand, it was speculated that action co-representation and anticipatory motor simulation did not occur in the passive observation condition, because the experimental setting could not be regarded as a meaningful social interaction. Such a conclusion was based on previous EEG evidence showing that the way in which social situations were perceived modulated action simulation, such that anticipatory motor activation (as reflected in the amplitude of the contingent negative variation) occurred when participants expected a particular action to be performed by a partner with which they were actively interacting, but not when they expected the same action to be performed by a third person with whom they were not interacting (Kourtis, Sebanz, & Knoblich, 2010). In the previous study, participants working in the passive condition seated next to each other but were explicitly instructed to avoid any interaction with the experimenter. Thus, no anticipatory motor activation and no memory advantage were found for the arrays pointed to by the task-partner.

The goals of the present study were twofold. First, we aimed at replicating the results reported by earlier experiments in the joint-action (Experiment 1) and passive observation (Experiment 2) conditions. Second, we aimed at providing further evidence in support of the conclusions drawn by earlier experiments by providing eye movement patterns in the joint action (Experiment 1) and passive observation (Experiment 2) conditions. Previous studies have reported that the patterns of eye movements reflect sequential action planning and control (Land, 2006; Land and McLeod, 2000; Land and Furneaux, 1997; Land, Mennie and Rusted, 1999; Ballard, Hayhoe, Li and Whitehead, 1992; Hayhoe, Bensinger and Ballard, 1998). For instance, Ballard et al. (1992) and Hayhoe et al. (1998) involved their participants in a task where they had to copy the pattern of colored blocks on a computer screen, using the mouse to drag the blocks. They found that every action- choosing a block, checking its color and finding its proper position involved a new fixation, where eye movements generally precede the motor action. These proactive eye movement patterns are also shown in many daily activities such as tea-making and sandwich making (Land, Mennie, & Rusted, 1999; Hayhoe, 2000). It has been suggested that there is a strong coupling between gaze and hand movements, such that, for example, gaze proactively guides the hand towards the objects that must be grasped and subsequently towards the landing sites where the objects must be moved (e.g., Johansson, Westling, Bäckström, Flanagan, 2001; Land, Mennie, & Rusted, 1999; Hayhoe, 2000; Land, 2006, 2009; Land and Hayhoe, 2001). This coupling is so strong that the eye fixations supporting hand-movement planning and control have been regarded as being part of the overall motor programs underlying many common tasks (Land & Furneaux, 1997). Of greater importance, there is good evidence indicating that, during action observation, participants produce a pattern of eye fixations which is very similar to that produced when they perform the task on their own (Flanagan and Johansson, 2003; Rotman, Troje, Johansson and Flanagan, 2006). A seminal study by Flanagan and Johannson (2003) compared the eye fixations produced during the execution and observation of a block-stacking task in which either the participant or the task-partner were asked to stack three wooden blocks from the widest to the narrowest. The results showed that the spatial distribution of fixations and the gaze-hand coordination in the action observation task were very similar to those occurring in the action execution task. According to the Flanagan and Johannson (2003), these findings supported the direct matching hypothesis and confirmed that the understanding of the co-actor’s actions resulted from a mechanism that mapped observed actions onto the motor representations of those same actions (Iacoboni et al., 1999; Rizzolatti, Fogassi, & Gallese, 2001). Eye movements thus seem to be tightly coupled temporally and spatially, to the motor actions involved in a task.

Therefore, the implications for our study are as follows: First, in the joint action condition, similar eye movement patterns should be expected in both types of trials, one when the participants execute the pointing movement towards the move arrays and one where the participants observe pointing movement performed by experimenter. That is, if the presentation of other-relevant arrays in the joint-action condition triggers the anticipatory motor simulation of the experimenter's pointing movements, then it follows that the pattern of eye fixations during the experimenter-relevant trials should be very similar to that observed during self-relevant trials. In contrast, as speculated earlier, in the passive observation condition, experimenter’s pointing movements are unlikely to be simulated by the participants, the pattern of eye fixations for items pointed by experimenter in passive condition should be different from the patterns of eye fixations obtained for experimenter pointing trials in the joint condition. Second, given that, the eye-movements are highly related to the allocation of visuospatial attention (Chelazzi et al., 1995; Kowler, Anderson, Dosher and Blaser, 1995; Shepherd, Findlay and Hockey, 1986; Moore and Fallah, 2001; Corbetta et al., 1998), the memory advantage of move arrays in the joint-action condition of Experiment 1 should be reflected in the eye fixation measures. That is, participants should fixate the items in the move arrays significantly more than the items in the no-move arrays during both types of trials in the joint action condition. On the other hand, eye movement patterns for the move arrays and no-move arrays should not differ in the passive observation condition of Experiment 2 as these two types of arrays are expected to be processed similarly.

2.2.3: Experiment 1: Self-performed pointing and experimenter-performed pointing movements in random alternation

2.2.3.1 Aim: Experiment 1 used a joint-action condition in which the participants and task-partner (i.e. the experimenter) alternated in executing pointing movements. Behaviorally, we expected both self-pointed and experimenter-pointed arrays to be recognized better than no-move arrays. We also predicted that participants should fixate self-pointed arrays more than no-move arrays. Finally, if participants simulated the movements performed by the experimenter (as speculated earlier), then a similar pattern of eye-fixations should emerge also in the experimenter-pointed trials, that is, a significant difference in the fixations of move and no-move arrays should also emerge in the experimenter-relevant trials.

2.2.3.2 Method:

2.2.3.2.1 Participants. Twenty naive volunteers (16 males; age: $M = 23.35$ years, $SD = 2.20$ years) from the University of Hyderabad (India) participated in the experiment. All the participants reported to have normal or corrected-to-normal vision. The institutional ethics committee of the University of Hyderabad approved the study. In the Bhatia et al. (2019) study, the effect size associated with the significant main effect of Condition (showing the memory advantage of move arrays) in the joint-action condition of Experiment 3 was $\eta_p^2 = 0.41$, which corresponds to $f = 0.83$. Using the G*Power software (Faul, Erdfelder, Lang, & Buchner, 2007), we estimated that, with $N = 20$, $\alpha = 0.05$ and a medium correlation between the repeated measures ($r = 0.50$), the post-hoc power to achieve a within-subjects effect of Condition of a magnitude similar to that obtained in our earlier experiment (Experiment 2.1.3) exceeded 0.99 in a repeated ANOVA (F tests: ANOVA repeated measures, within factors).

2.2.3.2.2 Apparatus and stimuli. The stimuli were the same as those used in previous experiments. The whole set comprised 192 visuospatial arrays, containing 3 or 4 items arranged in a 5 × 5 matrix. Of these, 96 arrays contained only circles (48 for each length), whereas the other 96 arrays contained only squares (again, 48 for each length). Both the circles and squares were 2 cm × 2 cm in size, with two adjacent items separated by 1 cm. For each pattern, we constructed a test lure having all the items in common with the original array except one, which was shifted by one or two positions. All the stimuli were displayed in black against a grey background, including the fixation cross and the letter cue for the pointing instructions (displayed in Times New Roman; size: 40 pt). Responses were collected using a Cedrus RB-844 response pad (Cedrus Corporation, San Pedro, California, USA), that recorded RTs in the range of 2-3 milliseconds.

The experiment was conducted in a noise-free, dimly lit room. Participants sat at 40 cm from a desktop mounting a SR Research Eyelink 1000 eye tracker (SR Research Ltd, Ontario, Canada), which recorded eye-movements with a sampling rate of 1000 Hz and a spatial resolution less than 0.01°. A chin rest was used to stabilize head movements during the experiment. An HP display monitor (1280 × 1024 pixels; refresh rate: 60 Hz), controlled by Experiment Builder software (SR Research Ltd, Ontario, Canada), displayed the stimuli. To locate the positions of the eye fixations, the display screen was divided into a 5 × 5 matrix corresponding to that in which the stimuli were originally displayed, resulting in 25 regions of interest.

2.2.3.2.3 Design and Procedure. Experiment 1 followed a 2 (Array Order: first vs. second) × 2 (Array Size: 3 vs. 4 items) × 2 (Condition: no-move vs. move) × 2 (Agent Cue: P-cued vs. E-cued) within-subjects design.

The general procedure was modelled from Experiment 3 of section 2.1, which in turn was modelled after Chum et al. (2007) and Dodd and Shumborski (2009). Each trial comprised an encoding phase, in which *two* arrays were presented (one was designed as the move array, while the other was designed as the no-move array), and a recognition phase in which participants had to determine whether the locations of a probe array matched the locations of one of the two encoded arrays (see Figure 2.6). During the encoding phase, each trial started with a letter cue for 2000 ms, which signaled who had to point to the items in the move array (‘P’ for the participant, ‘E’ for the experimenter). The letter cue was followed by a fixation point (a cross at the center of the screen) for 1000 ms, after which two spatial arrays were consecutively presented, one after the other. Both arrays contained the same number of items (either 3 or 4) but differed in shape (either squares or circles). Each item was presented sequentially for 1000 ms, such that each subsequent circle (or square) appeared when the previous circle (or square) disappeared. The sampling of the items’ locations was completely random (with the constraint that, within a given trial, no item occurred in a location previously occupied by another item), so that there was no apparent regularity in the presentation. A blank screen appeared for 1000 ms at the end of each array, during which participants could maintain the encoded array in working memory. The inclusion of these blank screens was justified by a growing body of research showing that, when people want to keep information active in working memory, they often allocate visual attention to the empty locations in space in which the stimuli were presented during the encoding phase (see Theeuwes, Belopolsky, & Olivers, 2009, for a review). Moreover, it has been shown that these fixations to “nothing” can have facilitatory effects on the retrieval of spatial information (Johansson & Johansson, 2014).

The whole experiment included 48 trials for each array size, for a total of 96 trials. Array size was blocked and the order of the blocks was counterbalanced across participants: half of the participants performed the three-item trials before the four item trials, whereas the other half performed the task in the reverse order. A break of two minutes, including a drift correction, was given after one block. Each trial in the study phase contained a move and a no-move array. At the beginning of the experiment, participants were shown an illustration of the procedure and were instructed to memorize the location of the items in both arrays, because memory for one of the two arrays would have been tested at the end of each trial. In addition, participants were instructed to passively view the locations of the presented items for one array (the no-move array) and to move their hands towards the locations of each item until touching the screen for the other array (the move array). Half of the participants were instructed to tap only on the arrays containing the squares, while the other half tapped only on the arrays containing the circles (this information was provided together with the instructions, at the beginning of each experimental session). Overall, pointing movements were performed by the participants on 48 trials, whereas they were performed by the experimenter sitting next to them on the other 48 trials (these two types of trials were randomly alternated within each block). In half the trials, the no-move array was presented first, followed by the move array; for the other half of the trials, the order was reversed.

The test phase began immediately after the presentation of the second array, starting with a black screen for 150 ms. The purpose of the recognition task was to examine participant’s memory for the locations of the items presented either in the move or in the no-move array. Thus, each test screen contained either squares or circles. Participants did not know in advance which array was tested in each trial: thus, they had to memorize both the arrays presented during the study phase.

They were instructed to judge whether the locations of the items presented at test matched (or not) those shown at encoding, by pressing two large keys labelled “same” or “different”. Participants were required to make a response within 5000 ms.

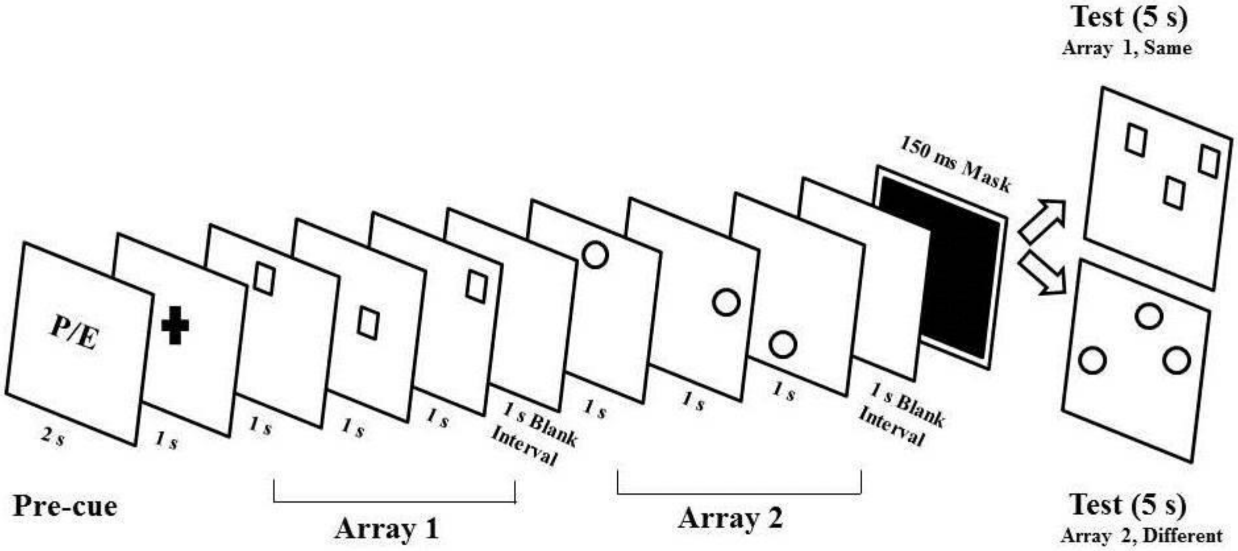


Figure 2.6. Schematic overview of the procedure used in Experiment 1. During the encoding phase, participants were initially presented with a letter cue indicating who had to perform the pointing movements (P for the participant, E for the experimenter). Then, they studied two consecutive arrays including three (or four) circles or squares: depending on the instructions, one array was designed as the no-move array, while the other array was designed as the move array. Each item was presented for 1 s. A mask (a black screen) followed the encoding phase for 150 ms, after which participants saw an array of squares or an array of circles: they had 5 s to decide whether the test array matched one of the two arrays presented at encoding.

2.2.3.2.4 Data Analysis. For recognition memory, accuracy was computed as the mean percentages of arrays correctly recognized as being equal to those shown during the encoding phase (this is the same dependent variable used in previous studies: Chum et al., 2007; Dodd & Shumborski, 2009). Our primary interest in the present study concerned the differences between move and no-move arrays. Since the effects of array size (3 vs. 4 items) and array order (first vs. second arrays) have been well characterized in previous studies (Chum et al., 2007; Dodd & Shumborski, 2009; Rossi-Arnaud et al., 2012), data were collapsed across these factors. The entire analyses are however reported in the Appendix I, where it can be noted that the main effects of arrays size and array order replicated the findings already reported in literature (namely, recognition was higher for 3- than for 4-item arrays, and higher for arrays presented as second than for arrays presented as first).

For eye movements, we computed the *relative fixation percentages* and the *relative gaze durations* for both the encoded arrays and the blank screen. By definition, a fixation represents a period during which the eyes remain relatively still and information is encoded into working memory, whereas gaze duration refers to the sum of all the fixations falling within an area of interest (AOI; Rayner, 2009). In the present study, the AOIs were obtained by dividing the display screen into a 5 × 5 matrix (a schematic figure is attached in Appendix I). Relative fixation percentages were computed by dividing the total number of fixations falling in a given AOI by the total number of fixations to the entire screen and multiplying the result by 100. Likewise, relative gaze duration was computed by summing the durations of all the fixations falling in a given AOI, dividing it by the cumulative duration of the fixations to the entire screen and multiplying the result by 100 (d'Ydewalle & De Bruycker, 2007; Chang & Choi, 2014; Georgescu et al., 2013).

2.2.3.3 Results and discussion:

2.2.3.3.1 Recognition accuracy. The mean percentages of arrays correctly recognized as being the same as those presented at encoding are shown in Figure 2.7. A 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 7.97$, $MSE = 518.07$, $p = 0.011$, $\eta_p^2 = 0.29$], indicating that recognition accuracy was higher for move ($M = 91.14\%$) than for no-move arrays ($M = 76.76\%$). The two-way interaction between Condition and Agent Cue approached the significance level [$F(1, 19) = 3.94$, $MSE = 43.13$, $p = 0.062$, $\eta_p^2 = 0.17$]: however, a follow-up analysis of simple effects indicated that move arrays were recognized better than no-move arrays both in the P-cued [$M = 89.16\%$ vs. $M = 77.70\%$: $F(1, 19) = 4.42$, $p = 0.049$, $\eta_p^2 = 0.19$] and E-cued trials [$M = 93.12\%$ vs. $M = 75.83\%$: $F(1, 19) = 11.31$, $p = 0.003$, $\eta_p^2 = 0.37$]. Lastly, the main effect of Agent Cue was not significant [$F(1, 19) = 0.24$, $MSE = 89.33$, $p = 0.63$, $\eta_p^2 = 0.01$].

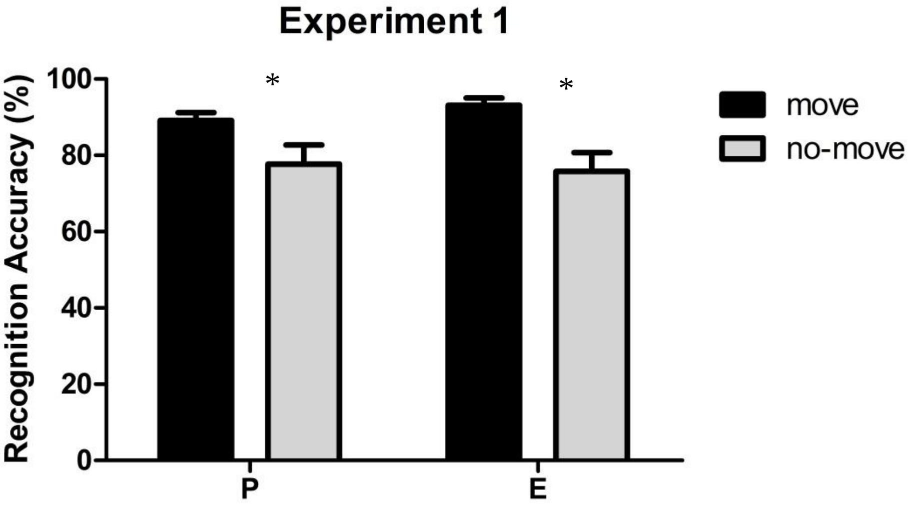
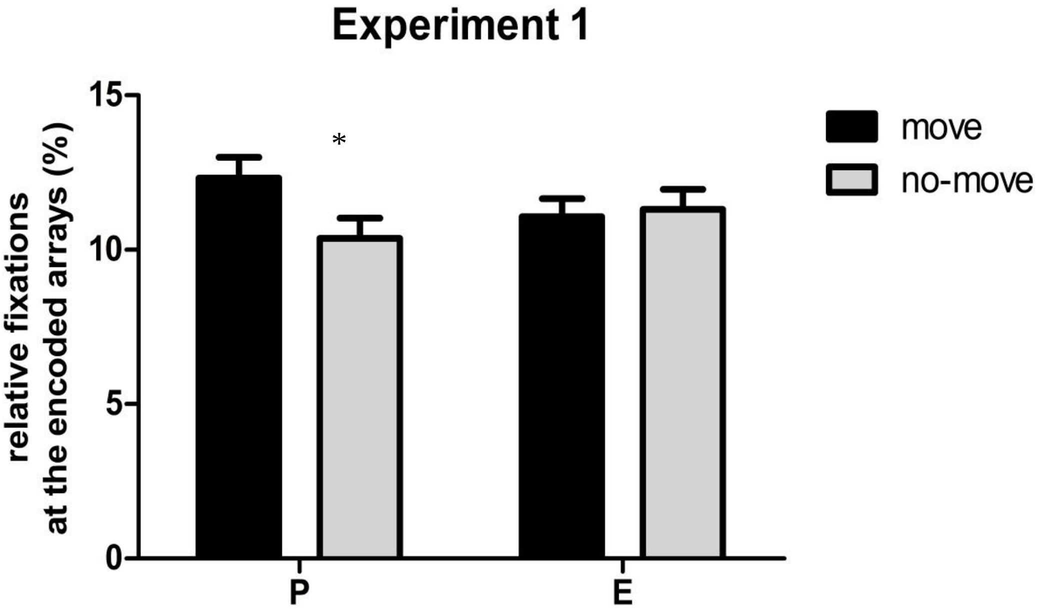


Figure 2.7. Experiment 1: Mean percentages of arrays correctly recognized as being the same as those presented at encoding, as a function of Condition (move vs. no- move) and Agent Cue (P-cued vs. E-cued trials). Bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.2.3.3.2 Relative fixation percentages for the encoded arrays. Relative fixation percentages for the encoded arrays are illustrated in the top panel of Figure 2.8. A 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 16.46$, $MSE = 0.88$, $p = 0.001$, $\eta_p^2 = 0.46$], indicating that relative fixation percentages were higher for move ($M = 11.69\%$) than for no-move arrays ($M = 10.84\%$), and a significant interaction between Condition and Agent Cue [$F(1, 19) = 19.38$, $MSE = 1.23$, $p < 0.001$, $\eta_p^2 = 0.51$]. A follow-up analysis of simple effects showed that relative fixation percentages were higher for move than for no-move arrays in the P-cued trials [$M = 12.32\%$ vs. $M = 10.37\%$: $F(1, 19) = 23.02$, $p < 0.001$, $\eta_p^2 = 0.55$], but not in the E-cued trials [$M = 11.07\%$ vs. $M = 11.31\%$: $F(1, 19) = 1.26$, $p = 0.28$, $\eta_p^2 = 0.06$]. The main effect of Agent Cue was not significant [$F(1, 19) = 0.46$, $MSE = 1.01$, $p = 0.50$, $\eta_p^2 = 0.02$].

2.2.3.3.3 *Relative fixation percentages for the blank screen*. The relative fixation percentages for the blank screen are illustrated in the bottom panel of Figure 2.8. A 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 10.63$, $MSE = 2.73$, $p = 0.004$, $\eta_p^2 = 0.36$], indicating that relative fixation percentages at the blank screen were higher for move ($M = 7.67\%$) than for no-move arrays ($M = 6.47\%$). The main effect of Agent Cue and the two-way interaction between Condition and Agent Cue were not significant [$F(1, 19) < 0.34$, $p > 0.56$].


Experiment 1
relative fixations
at the encoded arrays (%)
15
10
5
0
P
E
*
move
no-move

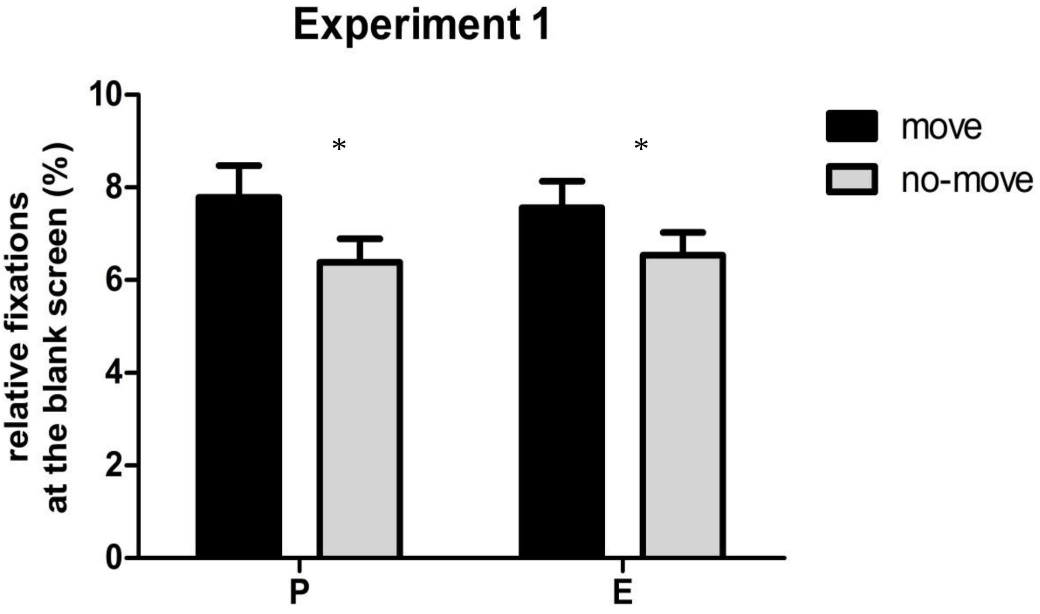
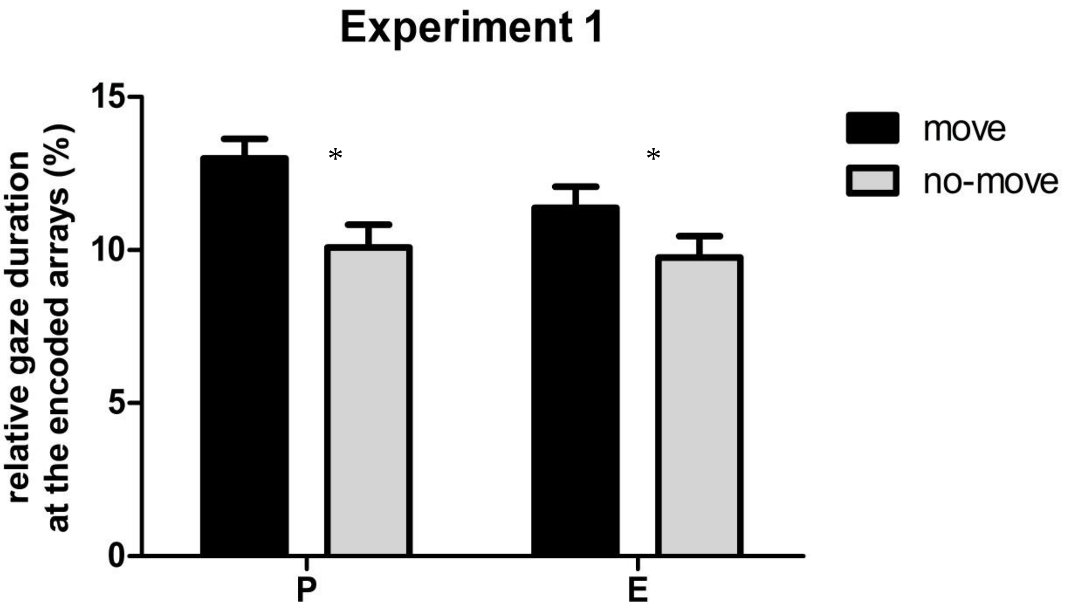

Experiment 1
relative fixations
at the blank screen (%)
10
8
6
4
2
0
P
E
*
*
move
no-move

Figure 2.8. Experiment 1: Mean relative fixation percentages for the encoded arrays (top panel) and the blank screen (bottom panel), as a function of Condition (move vs. no-move) and Agent Cue (P-cued vs. E-cued trials). Bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.2.3.3.4 Relative gaze durations for the encoded arrays. Relative gaze durations for the encoded arrays are reported in the top panel of Figure 2.9. A 2 (condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed significant main effects of a) Condition [$F(1, 19) = 22.88$, $MSE = 4.50$, $p < 0.001$, $\eta_p^2 = 0.55$], indicating that relative gaze durations were higher for move ($M = 12.18\%$) than for no-move arrays ($M = 9.92\%$); and b) Agent Cue [$F(1, 19) = 8.33$, $MSE = 2.27$, $p = 0.009$, $\eta_p^2 = 0.31$], indicating that relative gaze durations were higher in the P-cued ($M = 11.54\%$) than in the E-cued trials ($M = 10.56\%$). The interaction between Condition and Agent Cue was also significant [$F(1, 19) = 5.41$, $MSE = 1.54$, $p = 0.031$, $\eta_p^2 = 0.22$]: however, a follow-up analysis of simple effects indicated that the difference between the move and no-move arrays was significant in both the P-cued [$M = 12.99\%$ vs. $M = 10.08\%$: $F(1, 19) = 17.51$, $p = 0.001$, $\eta_p^2 = 0.48$] and E-cued trials [$M = 11.38\%$ vs. $M = 9.75\%$: $F(1, 19) = 22.10$, $p < 0.001$, $\eta_p^2 = 0.54$].

2.2.3.3.5 *Relative gaze durations for the blank screen.* Relative gaze durations for the blank screen are illustrated in the bottom panel of Figure 2.9. A 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 13.81$, $MSE = 4.39$, $p = 0.001$, $\eta_p^2 = 0.42$], indicating that relative gaze durations were higher for move ($M = 8.46\%$) than for no-move arrays ($M = 6.72\%$). The main effect of Agent Cue and the two-way interaction between Condition and Agent Cue were not significant [$F(1, 19) < 0.71$, $p > 0.41$].


Experiment 1
relative gaze duration
at the encoded arrays (%)
15
10
5
0
P
E
*
*
move
no-move

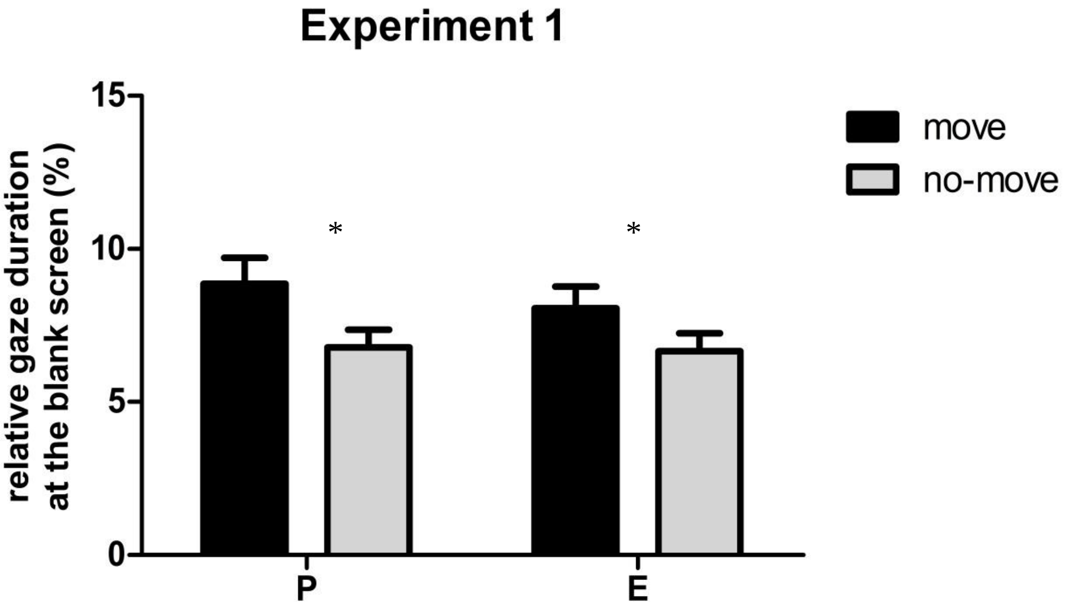

Experiment 1
relative gaze duration
at the blank screen (%)
15
10
5
0
P
E
*
*
move
no-move

Figure 2.9. Experiment 1: Mean relative gaze durations for the encoded arrays (top panel) and the blank screen (bottom panel), as a function of Condition (move vs. no-move) and Agent Cue (P-cued vs. E-cued trials). Bars represent standard errors. * indicates the significant different values, independent of the significance level.

Overall, the results of Experiment 1 confirmed the conclusions previously made, showing that, in a joint-action condition in which the participant and the experimenter alternated in making pointing movements, move arrays were recognized better than no-move arrays. Furthermore, the advantage accruing to move arrays was significant when the pointing movements were performed by the participant, but also when they were performed by the experimenter. Most importantly, the analysis of eye movements provided converging evidence, indicating that participants fixated the move arrays significantly more than the no-move arrays in both the P-cued and E-cued trials (the only exception being the relative fixation percentages for the encoded arrays, which did not differ between the move and no-move arrays in the E-cued trials).

2.2.4 Experiment 2: Experimenter-performed pointing movements

2.2.4.1 Aim: Experiment 2 used a passive condition in which pointing movements were solely observed by participants. The results previously reported in Experiment 2.1.4 suggest that observing the experimenter's pointing movements has no influence on the recognition of move arrays. Since Experiment 2 used the same stimuli and the same procedure as in Experiment 2.1.4, we expected to replicate their findings. We also expected no difference in the fixation and gaze duration measures between move and no-move items, suggesting that participants did not simulate the experimenter's movements in this passive condition.

2.2.4.2 Method:

2.2.4.2.1 Participant*s*. A new sample of twenty naive volunteers (13 males, mean age: *M* = 24.1 years, *SD* = 3.44) from the University of Hyderabad (India) students' community participated in the experiment. They all reported to have normal or corrected-to-normal vision. The study was conducted with the approval of the institutional ethics committee of the University of Hyderabad.

2.2.4.2.2 Apparatus and stimuli. The apparatus and the stimuli were the same as those illustrated in Experiment 1.

2.2.4.2.3 Design and Procedure. Experiment 1 followed a 2 (Array Order: first vs. second) × 2 (Array Size: 3 vs. 4 items) × 2 (Condition: no-move vs. move) within-subjects design.

The general procedure mirrored that adopted in Experiment 1, with the exception that, during the encoding phase of each trial, participants were instructed to passively observe the pointing movements performed by the experimenter towards the move arrays. At the beginning of experimental session, they were explicitly warned to maintain their hands still on the table. Half the participants observed the experimenter making pointing movements towards the squares, while the other half observed the experimenter making pointing movements towards the circles.

2.2.4.2.4 Data Analysis. Dependent variables were the same as those analyzed in Experiment 1 (percentage accuracy for recognition memory, relative fixation percentages and relative gaze durations for eye movements). As for Experiment 1, data were collapsed across array size and arrays order (see the Appendix for the full analysis of the recognition data).

2.2.4.3 Results and discussion

2.2.4.3.1 Recognition accuracy (%). A paired-sample *t*-test revealed that recognition accuracy did not differ between move (M = 76.91%) and no-move arrays (M = 75.78%), $t(19) = 0.50$, $p = 0.62$. This null result was confirmed by two types of follow-up analyses. First, we ran a Bayesian paired-sample *t*-test: in agreement with the above conclusion, the results showed that the null hypothesis was 5.19 times more likely than the alternate hypothesis (Rouder, Speckman, Sun, & Morey, 2009). Second, to determine whether we had sufficient power to detect a significant difference, we conducted a post-hoc power analysis. The to-be reached effect size was estimated from Experiment 1 by taking into account the means and standard deviations of the move and no-move arrays in the E-cued trials: $d_z = 0.85$. Using the software G-Power (Faul et al., 2007), we estimated that, with $N = 18$ and $\alpha = 0.05$, the power to detect an effect of Condition comparable to that observed in Experiment 1 was 0.95 (*t*-test for dependent means, two tails).

2.2.4.3.2 Relative fixation percentages for the encoded arrays and the blank screen. Two paired-sample *t*-tests showed that the relative percentages of fixations at the encoded arrays and the blank screen did not differ between move and no-move arrays: M = 10.34% vs. M = 10.16%, $t(19) = 0.65$, $p = 0.52$ for the encoded arrays and M = 7.36% vs. M = 7.47%, $t(19) = -0.25$, $p = 0.81$ for the blank screen, respectively. The Bayesian paired-sample *t*-tests confirmed that the null hypothesis was 4.78 and 5.69 times more likely than the alternate hypothesis.

.

2.2.4.3.3 Relative gaze durations for the encoded arrays and the blank screen. Two paired samples *t*-tests revealed no significant differences between move and no-move arrays in relative gaze durations: M = 10.72% vs. M = 10.40%, $t(19) = 1.05$, $p = 0.31$ for the encoded arrays and M = 7.92% vs. M = 7.97%, $t(19) = -0.11$, $p = 0.92$ for the blank screen, respectively. The corresponding Bayesian paired-sample *t*-tests indicated that the null hypothesis was 3.48 and 5.83 times more likely than the alternate hypothesis.

In summary, Experiment 2 replicated the behavioral findings reported in Experiment 2 in study 1 (passive observation condition), in that passively observing the pointing movements performed by the experimenter did not enhance recognition memory for the move arrays (relative to the no-move arrays). In agreement, the analysis of eye movements revealed no difference between the two types of arrays in terms of fixation percentages and relative gaze durations. Taken together, the results obtained in Experiments 1 and 2 support the hypothesis that participants simulated the movements performed by the task-partner in a mixed condition in which they alternated with the experimenter in performing pointing movements, but not in a passive observation condition.

2.2.5 General Discussion: The present study investigated the effects of pointing movements on VSWM in a joint-action condition, by simultaneously recording the participants' fixations during the encoding phase. We employed a novel paradigm introduced by Chum et al. (2007), in which each trial involved the presentation of two subsequent arrays: a n*o-move* array, encoded through passive visual observation, plus a *move* array, encoded through visual observation accompanied by pointing movements. The task was to determine whether a probe array, presented in a later test phase, matched or not one of the two previously encoded arrays. Behaviorally, the analysis of the recognition performance replicated the pattern reported in earlier experiments. When the participants and the experimenter alternated in performing pointing movements (Experiment 1), move arrays were recognized more accurately than no-move arrays, irrespectively of the agent who pointed (i.e., both the participant- and the experimenter-pointed arrays enjoyed a significant advantage). In contrast, when the participants passively observed the pointing movements performed by the experimenter (Experiment 2), move arrays were recognized no better than no-move arrays. Most importantly, the analysis of eye fixations revealed the expected differences between move and no-move arrays. In the joint-action conditions of Experiment 1, both the fixation percentages and the relative gaze durations were higher for move than for no-move arrays, with the advantage applying to participant- and experimenter-pointed arrays to the same extent. This pattern contrasted with the results obtained in the passive conditions of Experiment 2, in which neither the fixation percentages nor the gaze durations differed between move and no-move arrays.

Taken together, our data are consistent with the hypothesis, advanced in study 1, that participants working in the joint-action condition (but not those working in the passive condition) co-represented the task-partner's (i.e experimenter's) pointing movements as if they were on their own command (Atmaca et al., 2008, 2011; Sebanz et al., 2003, 2006). According to the *direct-matching hypothesis*, the observation of others' actions causes an automatic resonance in the observer's motor system, allowing him (or her) to understand the outcomes of these actions through simulation (Rizzolatti, Fogassi, & Gallese, 2001; Gallese et al., 2004). Starting from this account, previous studies demonstrated that the eye movements performed during action observation are strictly dependent on the activation of corresponding motor plans in the observer's mind (Flanagan & Johansson, 2003; Rotman, Troje, Johansson, & Flanagan, 2006). For instance, Flanagan and Johansson (2001) compared the pattern of eye movements during the observation and execution of a block-stacking task. Three wooden blocks having different width had to be stacked from the widest to the narrowest. In the task execution condition, participants fixated each forthcoming grasp and landing site well before the hand arrived; furthermore, their gaze exited the grasp and landing sites at about the same time as the hand. Importantly for the present purposes, this pattern of gaze-hand coordination was essentially equivalent to that exhibited in the passive observation condition. Flanagan and Johansson (2003) proposed that the similarity was due to the fact that participants observing others' actions implemented the same motor programs they used during action execution. Our results can be interpreted in the same way: the move arrays pointed by the experimenter were fixated more than the no-move arrays because participants working in the joint-action condition activated the same motor programs used to perform their own pointing movements.

As predicted, the results obtained in the passive conditions of Experiment 2 were very different from those reached in Experiment 1. First, the arrays pointed by the experimenter were recognized no better than no-move arrays. Second, the experimenter-pointed arrays did not enjoy a significant advantage in terms of fixation percentages or gaze durations. These results suggest that the way in which the arrays pointed by the experimenter were processed depended on whether the participant was actively involved in the task or not. In agreement, previous electrophysiological studies have shown that the social relationship between the participant and the task-partner plays a crucial role in the ability to anticipate others' actions (Hogeveen and Obhi, 2012). Specifically, anticipatory motor activation was found to be stronger when the participant expected a specific action to be performed by an interacting partner than by a third person they did not interact with (Kourtis, Sebanz and Knoblich, 2010, 2013). In the study by Kourtis, Sebanz and Knoblich (2010), the participant, an interaction partner and a non-interaction partner (i.e., a loner) sat around a table. Their task was lifting an object placed in the middle of the table and either returning it to its original position (individual action condition) or passing it to another person/receiving it from another person (joint action condition). The participants and the interaction partner engaged in the joint action condition, whereas the ‘loner’ acted alone and never interacted with others. The results showed that the contingent negative variation (a slow negative potential which reflects activity from the supplementary and primary motor cortices, and is therefore considered to be a marker of motor preparation: see Leuthold et al., 2004) was more pronounced when participants anticipated the partner’s actions than when they anticipated the loner’s actions. These findings indicate that participants were able to simulate the actions performed by the interaction partner, but not those performed by the loner (Tsai, Sebanz, & Knoblich, 2011; Vesper, Butterfill, Knoblich, & Sebanz, 2010). A similar explanation could apply to the passive observation conditions of Experiment 2, in which the absence of any personal involvement from the participant's side might have hindered their ability to perceive the experimenter as an interaction partner. If this were the case, then the fact that the pointing movements performed by experimenter were not represented in the participants’ cognitive system would come to little surprise.

The present study adds to the literature investigating the mechanisms by which observed and executed actions influence memory (Eskenazi et al., 2013). Our findings are indeed consistent with previous studies highlighting the importance of ‘selection’ in within-subject designs; when participants are required to select a subset of items for action, there is a distinct advantage for encoding these items in terms of perceptual and motor codes: the result is that these items stand out against the baseline established by passively-viewed items (Chum et al., 2007; Dodd & Shumborski, 2009). Similar positive effects of selection have been reported for many other experimental manipulations, including the perceptual-interference effect, the enactment effect, the production effect, bizarreness effects, word-frequency effects, and the effect of orthographic distinctiveness, among others (see Mulligan, Smith, & Buchin, 2019, for a discussion). Consider for instance the enactment effect, which refers to the improvement in the recall of action phrases when they are enacted by participants, as compared to when they are heard or performed by the experimenter (Engelkamp, 1998; Zimmer & Engelkamp, 1985). The enactment effect is robust when the study condition is manipulated within subjects, such that each participant encode both self-performed and experimenter-performed actions. In contrast, the effect disappears or decreases in size in between-subject designs in which each participant encodes only one type of action (Engelkamp & Dehn, 2000; Engelkamp & Zimmer, 1997). One potential explanation for these design effects is provided by the distinctiveness account. Under this account, enacted phrases benefit from distinctive processing (Hunt and Worthen, 2006) because they undergo extra encoding that results in a significant advantage over heard or experimenter-performed phrases. For example, in a typical within-subject design enacted phrases involve motor and kinesthetic processes that can be later used as additional cues to confirm that a phrase was studied (e.g., “I remember performing it, so I must have studied it”: Forrin, Ralph, Dhaliwal, Smilek, & MacLeod, 2019). By analogy, Chum et al. (2007) suggested that performing pointing movements induced participants to engage in a stronger form of egocentric (body-based) encoding, as compared to the allocentric (space-based) encoding used in the case of passively-viewed arrays. Allocentric pocessing is a function of the hippocampus and the parietal-cortex (DiMattia & Kesner, 1988; Save & Moghaddam, 1996) and is involved during the encoding of both pointed-to and passively-viewed arrays. Egocentric processing, on the other hand, is a function of the basal ganglia and is selectively involved when performing pointing movements (De Leonibus, Lafenetre, Oliverio, & Mele, 2003; De Leonibus, Oliverio, & Mele, 2005). Both the distinctiveness account and the Chum et al. (2007) hypothesis can be reconciled with the present results if we assume that participants in the joint-action condition covertly simulated the pointing movements performed by the experimenter: it is only in this case that the experimenter-pointed arrays would enjoy the same distinctive processing associated with participant-pointed arrays.

The above hypothesis is based on the idea that the covert simulation of the experimenter's pointing movements triggers additional encoding processes that can be later used to facilitate the recognition of the studied arrays. An alternative explanation comes from a series of studies showing that the activation of the motor programs corresponding to the experimenter's pointing movements requires a strong response inhibition which prevents participants from responding to the co-actor relevant stimuli (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, & Tzeng, 2006). Sebanz et al. (2006), for example, found that the amplitude of the No-go P300 (a component reflecting action control and response inhibition: e.g., Bokura, Yamaguchi, & Kobayashi, 2001) was larger in the joint-action condition than in the individual condition. A later fMRI study by Sebanz, Rebbechi, Knoblich, Prinz and Frith (2007) confirmed these findings by showing that the no-go trials in which the co-actor was expected to respond elicited greater activity in the inferior and superior parietal lobe as well as in the supplementary motor area (BA 6) in the co-action condition than in the individual condition. Sebanz et al. (2007) interpreted these results as an indication for increased demands on response inhibition during co-action. Based on these results, a different account for our data could be that the need to inhibit one's own responses during no-go trials led participants to devote more processing time to the experimenter-relevant stimuli: under this account, the higher recognition rates and the longer fixation times for the experimenter-pointed arrays would represent a byproduct of the selection conflict occurring in the joint-action condition (see Tsai et al., 2006, for a discussion). A similar hypothesis has been recently advanced by Elekes, Bródy, Halász, and Király (2016) to explain why the joint memory effect (i.e., the finding that other-relevant words are better encoded and recognized than non-task-relevant words) was significant even when the participants' task required non-motor responses (counting the target words). They proposed that, in the joint memory paradigm, it took longer for participants to decide whether they should respond to words from the other-relevant category (as it was associated with the co-actor's task) than to words from the non-task-relevant category, and that the longer processing time resulted in a deeper encoding of the former category. Clearly, our data cannot be used to discriminate between these alternative hypotheses: additional studies are needed to determine whether, in the joint-action condition, the presentation of the experimenter-relevant arrays triggers the activation of brain areas involved in the inhibition of motor responses.

To summarize, we successfully replicated the results already reported in study 1, showing that pointing movements were beneficial to recognition memory when the participant and the experimenter alternated in making pointing movements (Experiment 1), but not when the participant passively observed the movements performed by the experimenter (Experiment 2). Most importantly, we found that, in Experiment 1, both the number and the duration of eye fixations were higher for move than for no-move arrays, irrespective of the agent who performed the pointing movements; in contrast, no difference between the two types of arrays was obtained in Experiment 2. Taken together, these data are consistent with the co-representation account which proposes that, in the joint-action condition, self and others' actions are internally coded at the same representational level and their functional equivalency is reflected in a similar pattern of eye fixations. This co-representation might involve additional encoding (i.e., motor and kinesthetic) processes that can be later exploited to enhance recognition accuracy; alternatively, it might involve a longer encoding time due to the necessity to inhibit one's own responses to other-relevant stimuli. In both cases, the outcome would be a better recognition of the experimenter-pointed arrays (as compared to no-move arrays) in the joint-action condition.

2.3 Do I need to be involved? Determining the limiting conditions for the joint effects of pointing in visuo-spatial working memory

2.3.1 Abstract: Earlier studies suggest that performing pointing movements towards to-be-remembered target locations, result in a spatially-oriented perception of these target locations. In a VSWM task, participants were presented with two spatial arrays, one was encoded by visual observation while the other was encoded by visual observation accompanied by pointing movements. Earlier experiments suggest that pointing movements performed by a task-partner result in a similar memory advantage as obtained for self-performed pointing movements. In three experiments in the present study, the limiting conditions of the memory advantage obtained for pointing movements performed by task-partner are investigated. In Experiment 1, the frequency of turn-taking between the participant and task-partner for performing pointing movements is modulated (i.e. participants performed pointing movements towards move arrays six times less than the experimenter). In Experiment 2, the task-partner is replaced by an animated cartoon hand (i.e. pictures of a cartoon hand giving an animation effect to produce a movement to touch the target locations in the move arrays). Experiment 3 investigated whether the memory advantage obtained in Experiment 2 for the arrays pointed by the cartoon hand could be attributed to the mere presence of the cartoon hand. Findings show a) a memory advantage for the pointing movements performed by the experimenter similar to self-performed movements. It suggests that infrequent execution of pointing movements by the participants is sufficient to lead to the co-representation of partner's actions; b) a joint memory advantage also when the task-

partner (i.e. experimenter) is replaced by a moving cartoon hand; c) a lack of memory advantage when the movements were *only* carried out by the cartoon hand. The latter provides evidence that mere presence of the cartoon hand alone is not sufficient to contribute to the memory advantage. Altogether, the results suggest that the social-context between the task-partners while performing a task together modulates the task co-representation process.

2.3.2 Introduction: Earlier experiments investigating the influence of self-performed pointing movements on individual's visuo-spatial working memory (VSWM) have shown that movements performed towards to-be-remembered target locations benefit VSWM. Earlier experiments reported in the previous chapters, regarding pointing movements performed by a task-partner, have shown that partner's movements do not influence participant's memory in the same way in two pointing conditions, joint-action (i.e. turn-taking) and passive-observation conditions (i.e. when the participants are instructed to only passively observe partner's movements). Results showed that in the joint pointing condition when both agents perform the pointing movements by taking turns, observation of pointing movements performed by the task-partner benefited participant's memory similar to self-performed movements. In contrast, mere passive observation of pointing movements performed by the task-partner has no influence on the observer's spatial memory performance. It was suggested that, in a joint pointing condition, participants simulated and represented the pointing movements performed by the task-partner as their own movements, and that this, in turn, produced an advantage for the pointed-to arrays (move arrays) as compared to the merely visually-observed arrays (no-move arrays).

Moreover, the patterns of eye-movements provided by the eye-tracking experiments reported in section 2.2 have further strengthened this assumption by providing higher eye-fixations and gaze durations for the move arrays as compared to the no-move arrays in both types of trials, both when the participant executed the pointing movements and when the participants observed the experimenter perform the pointing movements towards the move arrays in the joint action condition. In contrast, in the passive observation condition, where participants were instructed to passively observe the pointing movements performed by the task-partner, eye-fixations and gaze durations did not differ between move and no-move arrays. The pattern of results, found so far, indicate that the task-partner's movements may be represented by participants as their own movements but only in the joint action condition when the participants performed the pointing movements alongside the task-partner. It was proposed that in the joint pointing condition, movements performed by the task-partner were represented in the same functional way as self-performed movements, because participants experienced this joint condition as a social interaction condition. Therefore, participants may have represented their partners' actions as their own, and this, in turn, produced similar eye movements patterns for self-performed and partner's performed pointing movements and subsequently had similar memory consequences for the movements performed by both agents. On the other hand, this was not the case for the passive observation condition, as participants were explicitly instructed to be passive and were allowed only to *observe* the pointing movements performed by the experimenter. It was suggested that the absence of a requirement to actively participate hindered the possibility of perceiving this condition as an effective social-interaction condition which subsequently led to the condition where the partner's movements were not processed similarly to self-performed movements.

The present study is designed to investigate the boundary conditions for the memory advantage obtained in the case of move arrays pointed by the experimenter in our previous experiments. As speculated in our previous experiments, the memory advantage obtained for experimenter-pointing movements is limited to a condition where both the participant and experimenter are actively involved in the task forming a joint-action condition. In other words, the memory advantage for experimenter-performed movements appeared only in a condition when the pointing movements were performed by the participants himself in half of the task. On the other hand, the memory advantage for experimenter-performed movements disappeared in a condition when participants were only passively observing the pointing movements. What the data suggests so far is how performing pointing movements towards target locations together with a task-partner affects an individual's memory for those spatial locations. The questions that remain open are (a) to which extent participants should be involved in performing pointing movements alongside a task-partner in order to actively simulate his movements and (b) whether the memory advantage obtained for observed pointing movements (i.e. experimenter’s movements in our previous experiments) is limited to a real life task partner. Therefore, the aims of the present study are twofold: (a) investigating whether a *minimum engagement* from the observer’s side (i.e. participant’s side) could produce the memory advantage for pointing movements performed by experimenter (b) investigating whether the memory advantage for pointing movements performed by the task-partner appears in a condition where the task-partner is replaced by an *animated cartoon hand*.

Earlier experiments (section 2.1 and 2.2) demonstrated that participants simulated the experimenter’s pointing movements only when they were actively engaged in the task themselves (i.e. joint action condition), not when they were passive observers of pointing movements performed by experimenter (i.e. passive observation condition). The findings so far are explained with regard to how the experimental condition is perceived by the participants. It has been suggested that participants co-represented experimenter's movements as their own considering that jointly performing pointing movements (i.e. by taking turns) created a condition of social interaction between them. In contrast, passively observing the experimenter's movements had no effect on the participants' performance on the memory task. These data were interpreted in terms of a condition of no social interaction between the two agents. Recently, the concept of *task co-representation* in joint-action partners had been investigated by using spatial compatibility tasks (e.g. Simon task) in an interactive context (Sebanz, Knoblich and Prinz, 2003; Tsai, Kuo, Jing, Hung and Tzeng, 2006). These sets of studies have shown that co-actors performing independent complementary tasks form representations of each-other's tasks, which interferes with their own performance (Sebanz, Knoblich and Prinz, 2003, 2005; Atmaca, Sebanz and Knoblich, 2011). In the case of joint Simon task (Sebanz et al., 2003), while co-acting with a task-partner, participants take their partner's actions into account and activate relative response codes in their action plans. However, what is unclear in the above-mentioned studies is the extent to which both task-partners should be involved in the task in order to co-represent each other's actions. In other words, it is not clear whether the joint-effects of performing a task together occur only when the task is equally distributed between the two agents. Thus, the present study introduced a condition where the participant is minimally engaged in performing pointing movements alongside the experimenter. In order to achieve this experimental condition, the joint-pointing condition introduced in experiment 2.1.5 (where the participant and experimenter took random turns to perform pointing movements; both agents performed pointing movements on 50% of the trials) is manipulated with respect to the frequency with which each task-partner performed the pointing movements towards move arrays. Therefore, the participant and experimenter took random turns in performing the pointing movements where the trials indicating participant's turn to perform pointing movements were dropped from 50% to 16.66%, i.e. six times less than the experimenter's turn. As speculated in our earlier experiments, no self-involvement from the participants' side (where participants were passive observers) led to a condition of no social interaction between the task-partners, therefore this condition did not affect participants' memory at all. The present study is sought to determine whether actively performing pointing movements towards move arrays in only one sixth of the total trials is sufficient to produce a memory advantage for the observed pointing movements (i.e. pointing movements performed by experimenter) by establishing a joint-action condition between the task-partners.

Experiment 2 in the present study is designed to test whether the memory advantage obtained when observing the experimenter's pointing movements in our previous experiments is limited to a real life task-partner. Previous studies have shown that mere belief of interacting with a task-partner is sufficient to elicit representations of partner's actions. For instance, Tsai, Kuo, Hung and Tzeng (2008) had their participants perform the social Simon task under the belief of having interacted with either a human partner (biological agent) or a computer program, where, in fact, all responses were generated by a computer. They found a social Simon like effect (comparison to compatible vs. incompatible trials) only when subjects believed they were interacting with a biological agent. They suggested that the believed intentionality of another person's actions is sufficient to elicit a social Simon effect. Similarly, a joint flanker effect emerged when the participants merely believed to be performing the task with a partner (Atmaca, Sebanz and Knoblich, 2011). A similar pattern of results have been obtained also in a picture-naming task (Kuhlen and Rahman, 2017). In this study, pictures of semantically related objects were named either individually by participants or in a joint condition where participant and partner took turns in naming the pictures. Naming latencies were increased when pictures were named by the partner in previous trials as compared to a condition when pictures were only visually presented but named by no-one. Interestingly, a similar interference effect was also observed when participants merely believed that their remotely located partner was naming the pictures. In a similar vein, studies have investigated how the animacy of task-partner affects the motor simulation process. Motor simulation is typically studied with a motor-priming paradigm, where participants are instructed to carry out responses while observing task-irrelevant congruent and incongruent actions (see, Brass, Bekkering, Wohlschläger and Prinz, 2000). Furthermore, a range of studies provide evidence that only actions performed by human agents lead to motor-priming effects, not when these actions are performed by non-human agents (Kilner, Paulignan and Blakemore, 2003; Tai, Scherfler, Brooks, Sawamoto and Castiello, 2004; Tsai and Brass, 2007). For example, Tsai and Brass (2007), had their participants perform a joint Simon task with either an animated human hand or a wooden hand performing the complementary part of the task. They found a Simon-like effect only when participants performed the task with an animated human hand. Authors proposed that the lack of joint-action effect when participants interacted with a wooden hand, suggests that the human co-representation system is biologically tuned. On the other hand, Müller et al. (2011) showed their participants a video fragment of a biological agent or of a non-biological agent, Pinocchio, before performing a joint Simon task with either an animated human hand or a wooden hand. The video was shown to participants to increase perceived animacy of these interacting agents. Participants who watched the video of biological agent showed a Simon-effect only when co-acting with animated human hand, whereas participants who watched video of Pinocchio, showed a Simon-effect only when co-acting with a wooden hand. They suggested that motor simulation strongly depends on higher-order processes. They further suggested that to co-represent partners' actions, participants do not necessarily need to be interacting with a real person. Instead, co-representation of actions may also occur when interacting with the characters of a video-game. Further studies reported by Dolk and colleagues (Dolk et al., 2011; Dolk, Hommel, Prinz and Liepelt, 2013) have shown Simon-like effect even in the presence of an attention-attracting event such as a Japanese waving cat or a metronome producing a ticking sound. They suggested that any attention-attracting event which provides a spatial reference for mapping one's own actions is sufficient to produce a Simon-like effect.

In another study, Liepelt, Prinz and Brass (2010), using a motor-priming paradigm, investigated whether agent-sensitivity of motor simulation depends on the types of observed actions. Participants were shown pictures of a human hand and a wooden hand depicting end positions of three types of actions: a communicative gesture (such as a hand-shaking gesture), a transitive (goal-directed action such as a fist holding an object) and an intransitive action (non-goal-directed such as a closed fist without an object). Interestingly, the motor-priming effect for communicating gestures was found for the human hand, but was absent for the wooden hand. On the other hand, motor-priming effects were found to be similar for transitive and intransitive actions. They suggested that biological tuning of motor-simulation is highly action-selective. Furthermore, neuroimaging studies have shown similar activation in the mirror system for human and non-human movement stimuli (Gazzola, Rizzolatti, Wicker and Keysers, 2007; also see Jansson, Wilson, Williams and Mon-Williams, 2007; Stanley, Gowen, and Miall, 2007). For instance, Gazzola and colleagues (2007) tested their participants for the motor activation of the mirror system for actions performed by humans and an industrial robot. Participants viewed video tapes of simple and complex goal-directed actions, performed by either a human or a robotic agent. They found that the mirror system was activated strongly by the sight of both human and robotic actions, with no significant differences between the two agents. The evidence presented so far with regard to the animacy of a task-partner involved in a joint task condition have produced mixed results therefore providing no clear conclusion whether the partner's actions are simulated and co-represented only when the task-partner is a conspecific.

Thus, having reviewed the evidence provided for the effects obtained by performing a task in a joint- action condition in previous literature, the present study is aimed at testing two hypotheses (a) minimum personal-involvement from the participant's side should be sufficient to create a condition of social interaction between the participant and experimenter, and this should be reflected in the memory advantage obtained for the move arrays over no-move arrays (b) memory advantage for the move arrays pointed by the task-partner should be replicated when the task-partner is replaced by a dynamic cartoon hand if the co-representation process is not modulated by the animacy of the task-partner.

2.3.3 Experiment 1: Self- Performed and Experimenter- Performed pointing movements in random alternation with least self-involvement from participant's side

2.3.3.1 Aim: Previous experiments reported that a joint memory advantage is observed only when participants and experimenter take turns in executing the pointing movements. The effect is attributed to the condition of social interaction between the two agents. We claimed that the lack of self-involvement from the participants' side in the pointing task, hindered the possibility of having represented this condition as an effective social interaction, which seems to be a requirement in order for a dyad to work jointly by representing partner's actions (Hogeveen and Obhi, 2012; Kourtis, Sebanz and Knoblich, 2010, 2013). If this were the case, a minimum personal involvement could be sufficient to create an environment of such a social interaction. In order to investigate the latter, in the present study, Experiment 1 introduces a condition of minimum self-involvement from the participants in the pointing task. The methodology is adopted from Experiment 2.1.5 where the participant and experimenter take turns to perform pointing movements and their respective turn is indicated by the presentation of a cue at the beginning of each trial. The critical exception is the percentage of trials in which participants were allowed to perform the pointing movements themselves. Pointing movements here were performed by participants in only one sixth of the total number of trials (i.e. 16 trials out of a total of 96 trials), whereas, pointing movements were performed by the experimenter in the rest of the trials (i.e. 80 trials out of a total of 96 trials). If personal involvement from the observer's side is a critical factor in the joint memory advantage obtained in earlier experiments having joint- pointing condition, we should replicate the results in Experiment 1, in the present study.

2.3.3.2 Method:

2.3.3.2.1 Participants: Twenty four students (15 female, 9 male, M = 23.08 years, SD = 1.47 years) of the Faculty of Medicine and Psychology, Sapienza University of Rome, volunteered to participate in the experiment. All participants had either corrected or corrected to normal vision.

2.3.3.2.2 Design and materials: Experiment followed a 2 (Array order: first vs second) x 2 (Array size: 3 items vs 4 items) x 2 (Condition: move vs no- move) x 2 (agent cue: P-cued vs. E-cued) within subject design.

Stimuli were the same 192 visuospatial arrays as used in our previous experiments.

2.3.3.2.3 Procedure: The general procedure mimicked the one used in the joint pointing condition (as in Experiment 2.2.5) with the exception of the frequency of the presentation of agent-cue indicating participant's turn ('P') to perform pointing movements: the agent-cue, 'P' for the participant's turn to execute pointing movements towards the item locations in the move array, appeared only in one sixth (16.66% of the total trials) of the total trials to minimize the active self- involvement from the participant's side. In the remaining 83.34% of the trials, the agent-cue, 'E' was presented indicating the experimenter's turn to perform pointing movements towards move arrays and participants were instructed to observe the movements. This percentage of trials for the participants' pointing movements was chosen so that they could perform pointing at least in one trial in each condition.

2.3.3.3 Results:

Recognition accuracy (i.e., the percentages of correct responses) for each condition is illustrated in **Figure 2.11**. A 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) x 2 (agent cue: P-cued vs. E-cued) repeated measures ANOVA found:

(a) A significant main effect of condition [$F(1, 23) = 25.95$, $p < 0.001$, $\eta_p^2 = 0.53$], indicating that move arrays (M = 76.64%) were recognized better than no- move arrays (M = 64.16%).

(b) A significant main effect of array order [$F(1, 23) = 35.30$, $p < 0.001$, $\eta_p^2 = 0.61$], indicating that recognition performance was higher for the arrays presented second (M = 83.12%) than the arrays presented first (M = 57.68%).

(c) A significant main effect of array size [*F*(1, 23) = 4.22, $p = 0.05$, $\eta_p^2 = 0.16$], indicating that recognition performance was higher for the arrays with size 3 (M = 73.22%) than the arrays with size 4 (M = 67.57%).

(d) A marginal interaction was found between condition and agent cue [*F*(1, 23) = 3.01, $p = 0.09$, $\eta_p^2 = 0.12$], follow up analysis showed that recognition performance for the move arrays was significantly higher than no- move arrays for both agent cues. At agent-cue, 'P', move arrays were recognized better (M = 78.12%) than no- move arrays (M = 60.41%)**,** [*F*(1, 23) = 13.21, $p = 0.001$, $\eta_p^2 = 0.36$]. Similarly, at agent-cue, 'E', move arrays were recognized better (M = 75.15%) than no- move arrays (M = 67.91%)**,** [*F*(1, 23) = 8.16, $p = 0.009$, $\eta_p^2 = 0.26$].

(e) A significant interaction was found between array order and condition [F(1, 23) = 11.02, $p = 0.003$, $\eta_p^2 = 0.32$], a follow up analysis revealed that the move arrays were recognized better (M = 69.94%) than no- move arrays (M = 45.41%) when they were presented first, [F(1, 23) = 28.80, $p < 0.001$, $\eta_p^2 = 0.55$].

No other effects or interactions were found significant.

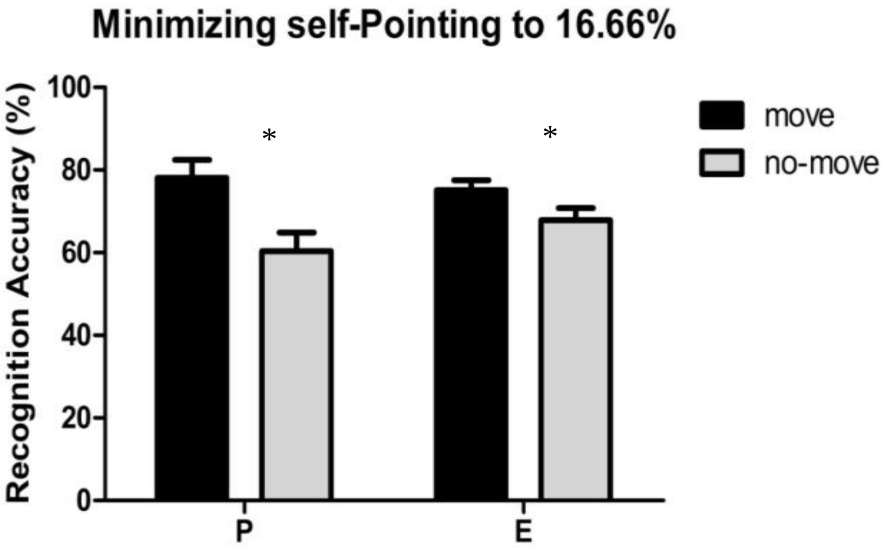


Figure 2.11: Experiment 1: Recognition accuracy as a fact or of condition (move vs. no-move) and agent cue (P vs. E). Error bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.3.3.4 Discussion: The results obtained in Experiment 1 in the present study replicate the pattern of results obtained in the joint- condition reported in chapter 2.1. Findings indicate that a significant memory advantage for the move arrays over no-move arrays is observed for both types of pointing movements, one where the participant performed the pointing movements and the other one where participant observed pointing movements being performed by experimenter. Moreover, both types of pointing movements elicited a similar memory advantage for the move arrays as compared to the no-move arrays. However, it is obvious to expect that the memory advantage for self-performed pointing movements should suffer when the trials with self-performed pointing movement were reduced. It was found that the recognition accuracy for the trials with self-performed pointing movements for the move arrays was 83.12 % and for no-move arrays 70.20% when participants performed pointing in half of the trials in Experiment 2.1.5 (i.e. 50% of the total trials). It was reduced to 78.12% for the move arrays and to 60.41%

for the no-move arrays in the present experiment when the percentage of self-performed trials was reduced to 16.66%. However, these differences in the recognition accuracy were only numerical and could not reach significance (in the case of move arrays: $F(1, 42) = 0.73$, $p < 0.39$, $\eta_p^2 = 0.01$ and in the case of no-move arrays: $F(1, 42) = 2.00$, $p < 0.16$, $\eta_p^2 = 0.05$). On the other hand, increasing the number of trials with pointing movements performed by experimenter (from 50% to 83.34%) did not increase recognition accuracy. Recognition accuracy for the move arrays was found to be 76.45% and for no-move arrays 63.74% when experimenter performed pointing movements in half of the trials, and it was found to be 75.15% for the move arrays and 67.91% for no-move arrays when they performed pointing movements on 83.34% of the trials. Similarly, no significant differences were found (in the case of move arrays: $F(1, 42) = 0.08$, $p < 0.77$, $\eta_p^2 = 0.002$ and in the case of no-move arrays: $F(1, 42) = 0.49$, $p < 0.49$, $\eta_p^2 = 0.12$).

2.3.4 Experiment 2: Do I need a real-life task partner? Self-performed and cartoon hand performed pointing movements in random alternation.

2.3.4.1 Aim: Previous studies have shown that the effect of co-representation of partner's actions emerges even when the participants are not interacting with a real person (Atmaca, Sebanz and Knoblich, 2011; Tsai et al., 2008; Tsai and Brass, 2007; Müller et al., 2011; Dolk et al., 2011, 2013). Therefore, Experiment 2 in the present study is aimed at investigating whether the memory advantage obtained for the experimenter's pointing movements in the joint-pointing condition reported in our earlier experiments, is limited to a real life task-partner. Thus, the partner is replaced with an image of a cartoon hand, animated to touch the item locations in the move arrays.

2.3.4.2 Method:

2.3.4.2.1 Participants: Twenty six students (25 females, 1 male, M = 24.23 years, SD = 2.47 years) of the Faculty of Medicine and Psychology, Sapienza University of Rome, volunteered to participate in the experiment. All participants had either corrected or corrected to normal vision.

2.3.4.2.2 Design and materials: Experiment 3 followed a 2 (Array order: first vs. second) × 2 (Array size: 3 items vs. 4 items) × 2 (Condition: move vs. no- move) × 2 (agent cue: P-cued vs. H-cued) within subject design.

Stimuli were the same as used in other experiments.

2.3.4.2.3 Procedure: The procedure was exactly the same as used in Experiment 2.1.5 with the critical manipulation of task- partner being replaced by a dynamic cartoon hand. In the case of no-move arrays, participants were instructed to look carefully at the locations of the presented items, while in the case of move arrays they were instructed to make pointing movements or to watch the animated cartoon hand making pointing movements towards the locations of each item during its presentation. Pointing movements were performed either by the participant or by the cartoon hand, depending on the agent-cue presented at the beginning of each trial indicating the turn to point towards the move arrays: in the case of agent-cue ‘P’, the participant had to perform pointing movements towards the move arrays. Whereas in the case of agent-cue ‘H’ (for hand, in original ‘M’ for ‘Mano’, in Italian), an animated cartoon hand appeared on the screen and touched the locations occupied by items in move arrays (see Figure 2.8). In the case of agent-cue ‘H’, participants were instructed to look carefully at the pointing movements performed by the cartoon hand. Overall, the agent-cue ‘P’, was presented in half of the trials whereas the agent-cue ‘H’, was presented in the other half of the trials. The presentation of these two types of cues was randomly alternated. To mimic the pointing movement performed by a real person, two images of a hand were included to produce an animation for the effect of an opening index finger; a closed fist for first 200 ms first appeared followed by an open index finger touching the item location for 800 ms in the first item presented in the move array, having the total duration of item presentation as 1000 ms. As observed with the real person, during the presentation of move array, the steps in performing pointing movements includes: (a) closed fist transitions into an open index finger (b) index finger touches the location of first item presented from the move array (c) index finger stays open to touch the locations for following item locations. The experimental setting virtually mimicked the real life interactive-context. A sample trial is shown in Figure 2.12.

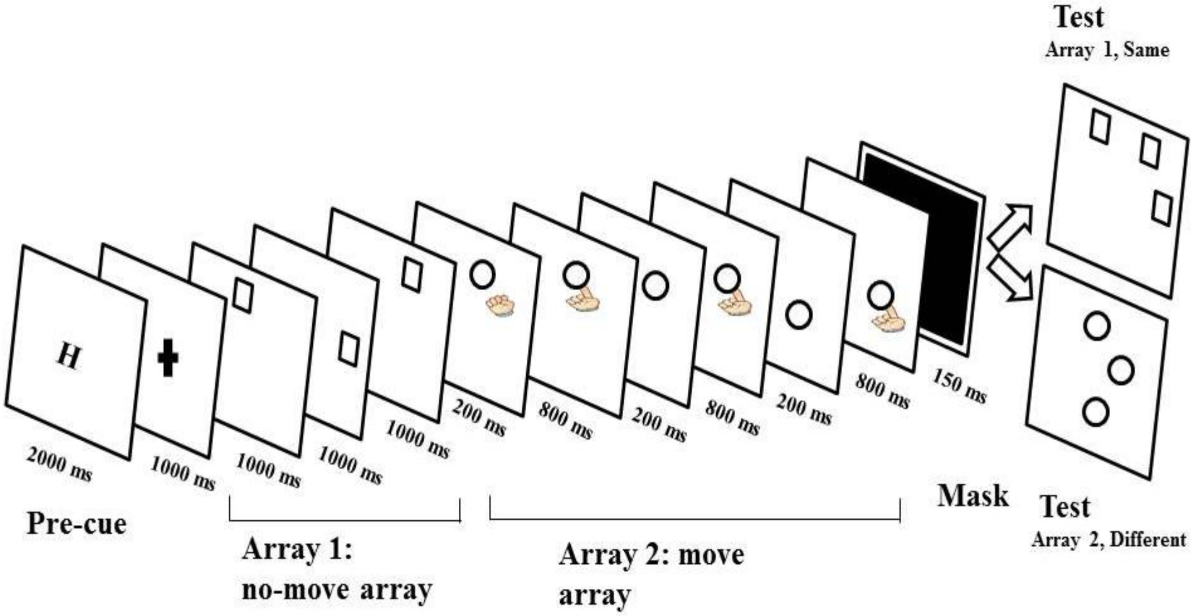


Figure 2.12: Experimental trial depicting the sequence of events in Experiment 2. An agent-cue was presented indicating the turn to perform pointing movements followed by the presentation of move and no-move arrays. In the case of 'Hand' cued trials, an image of a close fist first appeared for 200 ms followed by an open index finger touching to the target location for 800 ms. The index finger then stays open for the following target locations in a given trial to mimic the movements produced a real task-partner.

2.5.4.3 Results:

Data were analyzed with a 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) × 2 (agent cue: P-cued vs. H-cued) repeated measures ANOVA, which found (see **Figure 2.13**):

(a) A significant main effect of condition [$F(1, 25) = 7.09$, $p = 0.013$, $\eta_p^2 = 0.22$], indicating that move arrays (M = 84.45 %) were recognized better than no-move arrays (M = 71.55 %).

(b) A significant main effect of array order [*F*(1, 25) = 30.63, $p < 0.001$, $\eta_p^2 = 0.55$], indicating that arrays presented second were recognized better (M = 85.57 %) than arrays presented first (M = 70.43 %).

(c) A marginal significant main effect of agent cue [*F*(1, 25) = 3.91, $p = 0.059$, $\eta_p^2 = 0.13$], indicating that arrays pointed by the participant were recognized better (M = 79.80 %) than the arrays pointed by the animated cartoon hand (M = 76.19 %).

(d) A marginal interaction was found between array order and condition [*F*(1, 25) = 3.24 , $p = 0.084$, $\eta_p^2 = 0.11$]: a follow-up analysis of simple effects showed that move arrays were recognized better (M = 79.16 %) than no move arrays when presented first (M = 61.69 %), [F(1, 25) = 11.28, $p < 0.01$, $\eta_p^2 = 0.31$].

No other effects or interactions could reach significance [*F*(1, 25) < 2.04, $p > 0.166$].

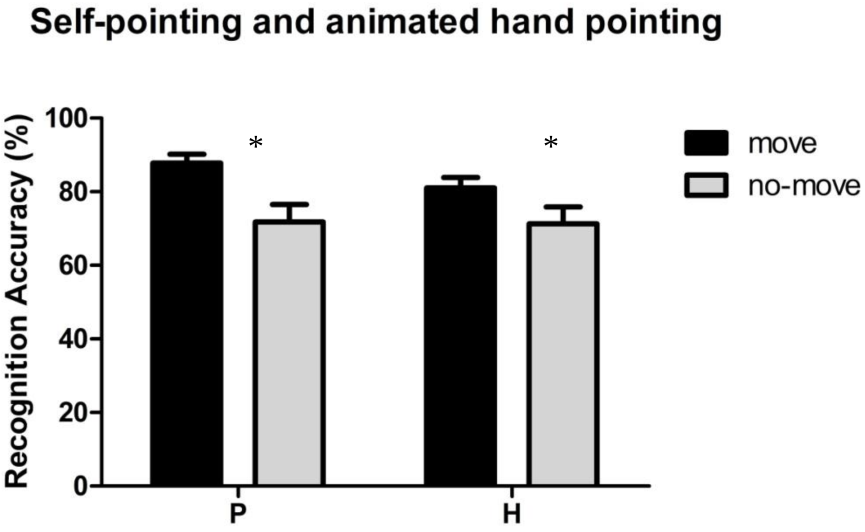


Figure 2.13: Experiment 2: Recognition accuracy as a function of condition (move vs. no-move) and agent-cue (P vs. H). Error bars represent standard errors. * indicates the significant different values, independent of the significance level.

2.5.4.4 Discussion: The pattern of results obtained in Experiment 2 in the present study is in line with previous studies suggesting that in a joint-action condition co-representation of partner's actions appears also when the task-partner is not a real person. The memory advantage obtained for the move arrays pointed by an animated cartoon hand supports the idea that participants represent not just their own actions but also their partner's actions even when the complementary part of the task is carried out by an animated cartoon hand.

2.3.5 Experiment 3: Animated cartoon hand as a task partner and no self-involvement from observer's side: Do I still get my memory benefit?

2.3.5.1 Aim: A set of studies have shown that neither the animacy of the agent nor the social context between the partners is responsible for the Simon effect (or joint Simon effect) obtained in previous joint-action studies (Dolk, Hommel, Prinz and Liepelt, 2013; Dolk et al., 2011, 2014). The joint Simon effect is obtained in a go-nogo version of the classical Simon task where the task is distributed between two participants such that each one of them operates one of the two responses (Sebanz et al., 2003). Dolk and colleagues proposed a referential coding account according to which any sufficiently salient event that provides a reference for spatially coding one's own actions, is able to produce a Simon like effect. By varying the salient nature of reference-providing events in an auditory go-nogo version of the Simon task, they demonstrated that a Simon effect emerged when the reference providing event was a non-social Japanese waving cat (Experiment 1) or even when a movement feature was provided by a clock (Experiment 2). In a series of experiments, Dolk et al. (2013) showed that as long as attention attracting events are providing a spatial reference for self-performed actions, events seem to be co-represented irrespective of the agent or objects producing them. Experiment 2 in the present study showed that the memory advantage obtained for the move arrays was observed even when the pointing movements were executed in an alternation with an animated cartoon hand. It was suggested that the social partner does not need to be a real person in order for the participant to co-represent the partner's actions in a joint task. However, the presence of memory advantage for the arrays pointed by the animated cartoon hand, could also be due to the mere presence of the hand on the screen. Therefore to test whether the memory advantage obtained in Experiment 2 was due to a) the 'socialness' formed from jointly performing the task, by taking turns in performing pointing movements, or b) the increased saliency of the move arrays with the mere presence of the hand which was absent for the no-move arrays, in Experiment 3 we included a condition where the pointing movements were performed by the animated cartoon hand in all the trials.

2.3.5.2 Method:

2.3.5.2.1 Participants: Twenty participants (10 Females, M = 25 years, SD = 2.90 years) from the Faculty of Medicine and Psychology, Sapienza University of Rome, were recruited in the experiment. All participants had either corrected or corrected to normal vision.

2.3.5.2.2 Design and materials: The experiment had 2 (array order: first vs. second) x 2 (condition: move vs. no-move) x 2 (array size: 3 vs. 4) within subjects design. The stimuli were the same 192 visuo-spatial arrays used in Experiment 2.2.3.

2.3.5.2.3 Procedure: The general procedure mimicked the one used in Experiment 2 except that the participants were not allowed to make pointing movements themselves.

Pointing movements in all the trials were executed by the animated hand, thus no agent-cue was presented at the beginning of the trials. Therefore, the move arrays were always pointed and touched by the cartoon hand in all the trials, and participants were instructed to passively observe the pointing performed by the cartoon hand. The no-move arrays were encoded by visual observation only. The encoding phase was followed by a recognition test.

2.3.5.3 Results:

A 2 (array order: first vs. second array) × 2 (array size: 3- vs. 4-item array) × 2 (condition: move vs. no-move array) repeated measures ANOVA revealed only a significant main effect of array order [$F(1, 19) = 19.29$, $p < 0.001$, $\eta_p^2 = 0.50$], indicating a recency effect for arrays being remembered better when they were presented at last (M = 82.91%) than when presented first (M = 73.33%). The effect of condition was not found significant [$F(1, 19) = 0.73$, $p = 0.40$, $\eta_p^2 = 0.03$] as well as other effects and interactions (see **Figure 2.14**). The Bayesian paired-sample *t*-test was performed to estimate the null effect of condition which confirmed that the null hypothesis was 3.62 times more likely than the alternate hypothesis (Rouder et al., 2009).

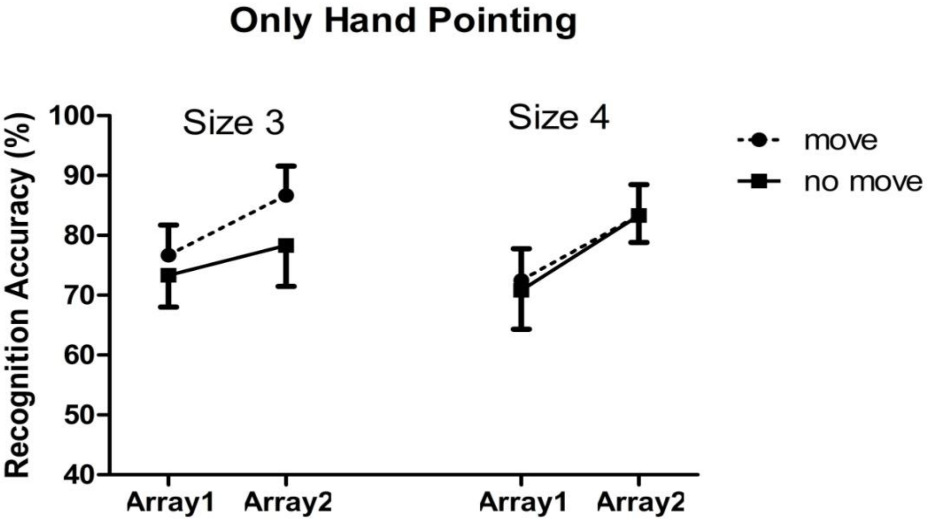


Figure 2.14: Experiment 3: Recognition accuracy as a function of array order (1 vs. 2), condition (move vs. no-move) and array size (3 vs. 4). Recognition Accuracy for move arrays was found to be 79.78% and for no-move arrays 76.45% (*p* = 0.40). Error bars represent standard errors.

2.3.5.4 Discussion: Experiment 3 did not produce a memory advantage for the move arrays over the no-move arrays, instead the recognition accuracies for two types of arrays did not differ at all. The pattern of results do not support the hypothesis that the mere presence of the hand, was sufficiently salient to produce a memory advantage, instead, suggests that participants' self-involvement in the motor task is a prerequisite for the memory advantage obtained in the case of move arrays pointed by a task-partner.

2.3.6 General Discussion: The present study had two aims: (a) investigating whether a minimum involvement on the participant's side in the pointing task could produce a memory advantage for the arrays pointed by the experimenter (b) investigating whether an animated cartoon hand could successfully replace a real life task-partner. Experiment 1 successfully replicated the memory advantage obtained in earlier experiments for experimenter-pointing movements performed in the joint- pointing condition even with minimum self-involvement from participant's side (i.e. one sixth of the total trials). Moreover, the results from Experiment 2 indicated that jointly performing the task with an animated cartoon hand replicated the results obtained while interacting with a real person. Experiment 3 in the present study, was designed to investigate whether the memory advantage obtained for the pointing movements made by animated cartoon hand could be accounted for the increased saliency of the move arrays as compared to the no-move arrays due to the presence of the hand. The pattern of results obtained in Experiment 3, however rule out this possibility, and additionally suggest that the social-context between the task-partners is an important factor for the memory advantage obtained from joint effects of pointing.

The patterns of results are in line with previous studies suggesting that people co-represent their partner's actions in ways that are functionally equivalent to their own actions, when the task is distributed between them (Sebanz et al., 2003). Sebanz and colleagues developed a go-nogo version of the Simon task which was divided between two individuals such that each one takes care of one response alternative. A task irrelevant feature was added by employing a picture of pointing index finger which either pointed to the left, or to the right, or to the middle. Participants had to respond to the color of the ring present on the index finger. One participant responded to the red color with left button press while the other responded to the green color with right button press. It was found that participants were faster to respond to their ring color when the finger pointed towards them, and slower when the finger pointed towards their task-partner, therefore activating a representation of partner's action. This interference effect, which is used as an index of co-representation of partner's actions, stems from the automatic activations of joint action plans (Sebanz et al., 2006). This finding indicated that even when coordination between the task-partners is not required, they take each other's actions into account and include each other's actions in their action planning.

Furthermore, in an EEG study, Sebanz, Knoblich, Prinz and Wascher (2006) investigated how performing a task together with a partner affects action planning and control. Participants performed a go-nogo version of the Simon task either individually or together with a partner. An action selection conflict was observed when individuals had to respond to a stimulus referring to the partner's action. They also showed that incompatible stimuli evoked larger P3 in the joint condition than in the individual condition suggesting that the partner's action was represented in a way similar to one's own. These results suggested that two action alternatives (i.e. relevant to one's own and partner's action) were represented in a functionally equivalent way. Furthermore, the analysis of nogo trials revealed that P3 amplitude was larger in the joint condition indicating that the response inhibition was stronger in the joint condition. Increased response inhibition in the joint condition was required for participants to prevent themselves from responding on nogo trials requiring partner's response. The study therefore revealed the cognitive mechanisms underlying simple looking joint action conditions. Furthermore, studies have shown how jointly performing a task is modulated by the social interaction between the task- partners. For instance, Kourtis, Sebanz and Knoblich (2010, 2013) investigated whether modulating psychological relationship between the task-partners would affect the motor simulation process during partner's actions. A participant, an interaction partner and a loner (who always acted alone) sat around a table. Their task was to lift an object placed in the middle of the table and either returning it to its original position (individual action) or pass/receive it to/from another person (joint action). In the joint action condition, the participant always interacted with the partner and the loner always acted alone. They found larger CNV (contingent negative variation) amplitude and stronger decrease in beta activity when participants anticipated an interaction partner to perform an action than when they expected the same action to be performed by the loner (third person they did not interact with). More relevant to the present context, in a control experiment performed by Kourtis and colleagues (2010), the loner was replaced by an infrequent partner. The EEG participant was now interacting with both partners but the frequency of interactions was modulated. The EEG participant now interacted with the frequent partner six times more than with the infrequent partner. Interestingly, no difference in the CNV amplitude for anticipating the actions performed by the frequent and infrequent partner was found this time. Authors proposed that social context between the task partners formed during the performance of a joint action task modulates action simulation. The replication of the pointing advantage obtained for experimenter's movements in Experiment 1 in the present study supports the assumption that the social interaction between the task-partners modulates the motor simulation process. The results obtained in Experiment 1 point out that participants simulate their partner's actions and take their actions into account when planning their own actions as long as they are interacting with them and perceiving the condition as a social interaction between them. The results support the claim that shared task- representations emerge as a consequence of how a social condition is conceptualized (Sebanz, Knoblich and Prinz, 2005).

Furthermore, the memory advantage for the move arrays pointed by the task-partner was replicated even when the task-partner is replaced by an animated cartoon hand. A significant number of evidence has been provided to show that co-representation occurs in the jointly performed tasks even when the task-partner is not a real-person (Tsai et al., 2008; Atmaca, Sebanz and Knoblich, 2011; Kuhlen and Rahman, 2017; Tsai and Brass, 2007; Müller et al., 2011). As mentioned earlier in the introduction, an interesting study by Tsai and Brass (2007), investigated whether co-representation occurs only when individuals engage in an interaction with a biological agent or could also be demonstrated when co-acting with non-biological co-actors. They had their participants perform the go-nogo version of the Simon task alongside either a static/dynamic human hand or a static/dynamic wooden hand. The compatibility effect was found to be specific to the dynamic human hand that is participants were faster in responding to the compatible stimuli and slower when responding to the incompatible stimuli only when co-acting with a dynamic human hand. Such an effect was absent for the static human hand, static wooden hand and dynamic wooden hand conditions. However, they found that RTs in the dynamic wooden hand condition were similar to the RTs for incompatible trials in the dynamic human hand condition. They suggested that movements performed by non-biological co-actors exert an unspecific influence on the observer's motor system. Therefore, they suggested that specific neural mechanisms facilitate human's social interactions with conspecifics.

Similarly, in the present study, a memory advantage obtained for the arrays pointed by the animated (i.e. dynamic) cartoon hand similar to the memory advantage for the arrays pointed by the participant could also be accounted for by the interactive-context formed while co-acting with a biological agent. Therefore, it can be suggested that while co-acting with a dynamic cartoon hand, participants co-represented its pointing movements as their own, which led to a similar memory advantage for the move arrays over no-move arrays. However, one can argue that participants not necessarily co-represented the pointing movements performed by the cartoon hand, but mere presence of the hand in the case of move arrays elicited a better encoding of these arrays as compared to the no-move arrays. The results obtained in Experiment 3 in the present study, however, do not support this possibility. Experiment 3 in the present study was designed to investigate whether the increased saliency of move arrays due to the presence of animated hand could be accounted for the memory advantage obtained for the arrays pointed by the animated hand. As reported by Dolk, Hommel, Prinz and Liepelt (2013), any sufficiently salient event that provides a spatial reference for one's own actions, can result in a Simon effect. In Experiment 1, they showed that a nonsocial co-actor (a Japanese waving cat) produced a go-nogo Simon effect by providing a spatial reference for the participants' own actions (i.e. cat presents on the left side of the screen and participant sits on the right side). Further, to rule out the possibility that the cat's face and arm movement could provide a degree of socialness/biological feature, which might be sufficiently similar to the presence of a human being, they replaced the cat with a clock. In Experiment 2, exchanging biological with non-biological movement feature also revealed a significant Simon effect suggesting that salient nonsocial events can produce a Simon effect. It was speculated that the visible movement feature could result in the significant Simon effect, therefore, in Experiment 3, the visible movement feature of the clock was replaced by an audible clicking sound (but no visible movement feature was present) of metronome. Experiment 3 still produced a reliable Simon effect. In a follow-up experiment they replicated Experiment 3, but the metronome did not produce any sound, to test whether the objects must possess some sort of event character to produce a Simon effect or whether the mere presence is sufficient. Interestingly, they did not find Simon effect in the absence of attention-attracting event suggesting that the dynamic events are responsible for the Simon effect. These sets of experiments, therefore, suggested that the social setting of an experimental situation alone is unlikely to modulate the task representations in a go-nogo task. In fact, as long as an event character (such as a hand movement, clock sound etc.) attracts attention in a bottom-up fashion, Simon like effect emerges. The current patterns of results, however, do not completely support this claim. The lack of memory advantage for the move arrays pointed solely by the animated cartoon hand, in Experiment 3, in the present study, suggests that the event character (the hand's dynamic nature) or mere presence of the hand alone was not sufficient to produce the memory benefit in the case of move arrays over no-move arrays. This further strengthens the argument that the social relationship between the task-partners is important to have a beneficial effect of pointing movements performed by the task-partner. However, due to the biological appearance of the cartoon hand used in the current study, the current findings are not sufficient to completely rule out the possibility that the dynamic events irrespective of their social or biological nature are sufficient to produce effects similar to the effects obtained while co-acting with a human being. In fact, collectively, the patterns of results indicate that social interaction between the task-partners modulates the process of task representations and subsequently influence memory for items relevant to a task-partner. The pattern of results is in line with previous studies suggesting that the co-representation of partner's actions emerges as a result of social interaction between the task-partners formed while performing a task together (Sebanz et al., 2006).

2.4 Effect of self and other's pointing movements on the encoding of spatial locations: An Event Related Potentials study

2.4.1 Abstract: Participants performed a joint visuo-spatial working memory task where two spatial arrays were encoded, one through pointing movements performed by the participant and the task-partner in alternation and the other through mere visual observation. A memory advantage was obtained for the arrays pointed by both agents as compared to the visually observed arrays. The electrophysiological findings revealed that early inhibitory control was applied during the presentation of move arrays which required to be pointed by the task-partner. Moreover, it was found that higher visual attention was allocated to the processing of move arrays pointed by both agents compared to no-move arrays. Further, higher attentional resources were found to be allocated to the item of move arrays pointed by the participant her/himself whereas attention shifted towards the movements performed by the task-partner instead of the item locations. The latter suggests that participants relied on the partner's actions to create more recallable memory representations in the case of move arrays pointed by the task-partner. Altogether, the data suggest that people make use of additional movement information available at encoding in order to create better memory representations.

2.4.2 Introduction: Experiments reported earlier (in section 2.1, 2.2 and 2.3) have demonstrated a consistent memory advantage for the arrays pointed by a task-partner in the joint version of the visuo-spatial working memory task compared to arrays encoded through mere visual observation. As shown in the earlier parts of the experimental chapter, when the participant and the task-partner alternate in performing pointing movements, the benefit on the participant's visual working memory is comparable for both the pointing movements performed by the task-partner and those performed by the participant her/himself. To recap the task, in the joint condition, the visuo-spatial working memory task is divided between the participant and the task-partner such that the pointing movements were either performed by participant or by the task-partner in a turn-taking context. Their respective turns to perform pointing movements in a given trial were instructed by a pre-cue presented at the beginning of each trial. As a general result, it was shown that participants' memory was better for the arrays which were pointed (move arrays) as compared to the arrays which were not pointed (no-move arrays). Most importantly, a similar memory advantage for the arrays pointed by the task-partner was also consistently observed when participants and the task partner performed the pointing movements in a joint pointing condition (i.e. by turn-taking). In contrast, in the passive observation condition i.e. when participants passively observed pointing movements performed by the task-partner, no memory advantage for the move arrays was observed. It was suggested that participants simulated their partner's pointing movements and represented them as their own, but only when they themselves were actively involved in the pointing task (i.e. the joint pointing condition). This in turn produced a similar memory advantage for pointing movements performed by both agents.

In the present study, we investigated the electrophysiological basis of the memory advantage obtained in the case of move arrays when they were pointed by participant her/himself and the task-partner over no-move arrays. With regard to how an individual's performance is affected while performing a task together with a partner, previous studies have demonstrated a joint compatibility effect using a joint go-nogo version of the Simon task (Sebanz, Knoblich and Prinz, 2003). As mentioned in earlier chapters, Sebanz and colleagues (2003) showed a joint-compatibility effect in the joint go-nogo version (i.e. where two people perform the task together, each one taking care of one response alternative) of the Simon task which was absent in the individual version (i.e. where one person performed his part of the task while the partner's chair remained empty) of the task. In other words, when the task in performed in a shared-context, where each participant was responsible for a complementary part of the task, a spatial-compatibility effect was observed (i.e. joint-compatibility effect). Sebanz and colleagues (2003) proposed that such an effect emerged because participants activated a representation of the partner's actions. This effect was not observed in a control condition when participants performed their part of the task alone (individual go-nogo condition; Sebanz et al., 2003). Furthermore, in an electrophysiology study, Sebanz, Knoblich, Prinz and Washer (2006) found a larger go-P3 in the incompatible trials (i.e. when a stimulus requiring one's response pointed towards the other person) in the joint condition than in the individual condition reflecting to a similar representation of the stimulus referring to the partner's actions as one's own, whereas no such representations were processed in the individual condition. They also found stronger nogo P3, in the joint condition as compared to the individual condition reflecting stronger response inhibition required in the former condition. Moreover, with regard to the ERP components related to action inhibition, Tsai et al., 2006 found N2 at about 200 ms after stimulus onset which was maximal at frontal sites (Falkenstein et al., 1999; Bruin and Wijers, 2002). N2 is typically larger on the nogo than on the go trials reflecting response inhibition (Kok, 1986; Falkenstein, 1995, 1999; Lavric et al., 2004). Tsai and colleagues had their participants perform go-nogo task in three conditions: joint go-nogo condition, individual go-nogo condition and being observed go-nogo condition. They found nogo N2 in all conditions indicating that action inhibition occurred in the nogo trials irrespective of the social context of the condition. On the basis of these findings it was suggested that perceiving a stimulus relevant to the co-actor’s actions leads to the activation of representational structures involved in one’s own execution of these actions which need to be inhibited in order to perform the self-relevant task successfully.

Evidence for these interpretations come from several studies suggesting that perception and execution of an action have common representational bases. Moreover, the findings of mirror neurons have stimulated extensive research on the relationship between perception and action (Rizzolatti and Craighero, 2004). Neurons in the ventral premotor and inferior parietal cortex in macaque monkeys were found to be active during both execution and observation of the same action (Gallese et al., 1996; Rizzolatti et al., 1996). Several neuroimaging evidence suggests that the observation of another person’s actions evoke activity in motor related areas in humans (Grèzes and Decety, 2001; Blakemore and Decety, 2001; Buccino, Binkofski and Riggio, 2004; Grèzes, Passingham and Frith, 2004; Rizzolatti and Craighero, 2004). These findings suggest that observed actions are directly mapped onto motor representations of the same actions in observers (Iacoboni et al., 1999). According to the ideomotor theory (James, 1890; Prinz, 1997) the direct matching of observed and executed actions result from an overlap of sensory and motor representations. A number of studies have provided evidence of this common-coding by suggesting that during observation of an action a corresponding representation in the observer’s action production system is activated (Brass, Bekkering and Prinz, 2001; Brass, Bekkering, Wohlschläger, & Prinz, 2000; Press, Bird, Flach, & Heyes, 2005; Stürmer, Ascherschleben, & Prinz, 2000). Furthermore, it has been suggested that mirror system helps to understand an action through an internal embodied simulation of the action (Gallese & Goldman, 1998; Rizzolatti, Fogassi, & Gallese, 2001).

2.4.2.1 Motor Preparation related ERP components: Preparatory processes for a motor response are generally studied using a precuing paradigm (for a review, see Leuthold, Sommer and Ulrich, 2004). A precue is provided with the information about the required motor response for the upcoming imperative stimulus after a brief preparatory period (i.e. foreperiod). During the foreperiod two ERP components arise indicating different preparatory processes: Contingent negative variation (CNV) and Lateralized readiness potential (LRP). A lateralized increase in the activity of primary motor cortices due to one side limb movement over the contralateral hemisphere causes a difference in electrical potential between the electrodes placed upon the contra and ipsilateral motor cortices, producing LRP. The LRP during the foreperiod can be served as an index of lateralized movement preparation and motor activation (Coles, 1989; de Jong, Wierda, Mulder and Mulder, 1988). The foreperiod LRP reflects preparatory processes on a motor stage and the CNV indicates preparatory processes occurring prior to lateralized motor preparation. The CNV is a sustained negative potential shift recorded at central electrodes and develops after the cue stimulus (Walter, Cooper, Aldridge, McCallum and Winter, 1964). CNV has been shown to be linked with motor and non-motor preparatory processes (Ikeda et al., 1996; van Boxtel and Brunia, 1994a, 1994b), anticipation of upcoming stimulus (Cohen and Walter, 1966; Gaillard, 1977; Loveless and Sanford, 1974; Tecce, 1972) and response preparation (Leuthold et al., 2004; Ulrich, Leuthold and Sommer, 1998; Wild-Wall, Sangals, Sommer and Leuthold, 2003).. LRP reflects response preparation and commences before the muscle contraction begins and therefore, can occur in the absence of an overt response (Galdo-Alvarez and Carrillo- de-la-Pena, 2004; Miller and Hackley, 1992; Minelli et al., 2007).

2.4.2.2 Attention and memory related modulations of ERP components:

ERP modulations typically begin within 100 ms of stimulus onset which provide strong evidence for the proposal that attention operates at an early stage to improve sensory processing (Mangun et al., 1993). Previous studies have implicated an interactive fronto-parietal network of cortical areas that mediates the on-line storage of location information (for a review, see, Awh and Jonides, 1998; Awh, and Anllo-Vento and Hillyard, 2000). It has been suggested that the active maintenance of location-specific representations are mediated by shifts of spatial attention to the memorized locations (Awh et al., 1995; Smyth and Scholey, 1994). Several ERP studies have shown that stimulus selection during spatial attention occurs at an early level of processing in visual areas and involves an amplitude modulation of sensory-evoked brain activity beginning at about 80 ms after stimulus onset (Hillyard and Anllo-Vento, 1998; Anllo-Vento and Hillyard, 1996). Specifically, stimuli at attended locations elicit enhanced positive (P1, latency 80-130 ms), and negative (anterior/frontal and posterior/parietal N1, latency 150-200 ms) ERP components as compared to stimuli at unattended locations suggesting that spatial rehearsal involves selective spatial orienting towards locations held in working memory (Awh, and Anllo-Vento and Hillyard, 2000; Gonzalez et al., 1994). Therefore, visuo-spatial attention has been found to be modulating subsequent stimulus processing, reflecting in early visual-evoked potentials P1 and N1 (Mangun and Hillyard, 1991). Several studies suggest that these effects may reflect different aspects of attention, P1 reflecting inhibition of processing of unattended stimuli whereas N1 reflecting amplification of processing of attended stimuli (Couperus and Mangun, 2010; Freunberger et al., 2008; Luck et al., 1994).

Moreover, P2 peak evoked around 200 ms after the onset of visual stimulus is also modulated by the attention paid to the stimulus. The amplitude of P2 changes as the subjects pay attention to the visual stimuli in right or left hemispace (Maeno et al., 2004; Johannes et al., 1995). Anterior P2 component, constitutes part of the normal sensory-evoked response to a visual stimulus and has been linked with higher order perceptual and attentional processing (Luck and Hillyard, 1994). Further, it has been suggested that P2 may occur when a stimulus is being compared with mental representations already stored in memory (Evans and Federmeier, 2007; Federmeier and Benjamin, 2005). Moreover, P2 component is found to be associated with sensory functions and is thought to reflect automatic identification and classification process (Johnstone, Barry, Anderson and Coyle, 1996; Oades, Dittmann-Balcar and Zerbin, 1997).

P3 wave comprises a frontally maximal P3a component and parietally maximal P3b component (Squires, Squires and Hillyard, 1975). Both are elicited by unpredictable, infrequent changes in the stimuli but the P3b (commonly referred to as P3) is present only for the task-related changes in the stimuli (Luck, 2014). Regarding the cognitive processes reflected by P3, Donchin (1981) proposed that it is related to a process of ‘context updating’. However, Donchin (1981) referred to the context representations in terms of broad representations of overall state of the environment, not specific to individual items or task, but it has been often interpreted that P3 reflects the updating of working memory (Luck, 2014; Luck, 1998a; Vogel, Luck and Shapiro, 1998; Vogel and Luck, 2002). Furthermore, Donchin (1981) suggested that the amplitude of the P3 elicited by a stimulus is predictive of its later memory. Moreover, it has been shown that P3 amplitude is larger when subjects devote more effort to a task suggesting that P3 is a measure of resource allocation (Isreal, Chesney, Wickens and Donchin, 1980). Thus, if a task is more difficult, this might increase P3 amplitude by encouraging subjects to devote more effort to the task.

Furthermore a *Late Positive Potential* (LPP) has been found to reflect information maintenance in a working memory task (Ma et al., 2019; Weinberg and Hajcak, 2011; Ruchkin, Ray Johnson, Canoune and Ritter, 1990). Moreover, it has been shown that LPP is sensitive to the attentional demands of the WM tasks (Corbetta et al., 1993; Yantis et al., 2002). Studies have also shown that the amplitude of LPP is smaller when attention is focused on an additional mental activity (Gevins et al., 1996; Mcevoy, Smith and Gevins, 1998; Ma et al., 2019).

2.4.2.3 Inhibition related modulations of ERP components: Inhibition refers to the ability to actively suppress a response, interrupt an activated response and avoid interference (Harnishfeger, 1995; Pfefferbaum, Ford, Weller and Kopell, 1985). Brain imaging studies provide evidence for the involvement of prefrontal areas of the cortex in response inhibition (Kawashima et al., 1996; Rubia et al., 2001). Inhibitory processing is generally studied with the go-nogo task, which has a go stimulus requiring a response (such as a button press) and a nogo stimulus requiring withholding or inhibition of a response. Go and nogo stimuli generally contain early ERP components as P1 and N1 reflecting early processing of the stimulus. Nogo stimuli further elicits two later components: N2 and P3 which are sensitive to manipulations of inhibition. N2 is a negative component occurring at about 200 ms after stimulus onset maximally at frontal areas (Bruin and Wijers, 2002) and is larger in nogo than go (Eimer, 1993; Falkenstein et al., 1995; Kopp et al., 1996; Pfefferbaum et al., 1985; Kok, 1986). P3 is a positive component related to inhibition which is fronto-central maximum for nogo stimulus and a centro-parietal maximum for go stimulus (Bruin and Wijers, 2002; Falkenstein et al., 1995; Kopp et al., 1996; Pfefferbaum et al., 1985; Kok, 1986). However, the role of inhibitory processes reflected by N2 and P3 components are well established in go-nogo studies, studies investigating early inhibitory processes are not very common. Nevertheless, Nakata et al. (2004, 2005a, 2006) found an early inhibition component, N140, occurring at around 140 ms after nogo stimulus onset. They had their participants perform a go-nogo task and recorded ERPs in three conditions: control (when subjects rested quietly with no task, count (when they had to count the number of go stimuli), and movement (when they had to respond to the go stimulus with their right hand). They showed that during the nogo trials, when subjects decide not to count or respond, a difference in potential between the go and no-go stimuli is evoked in this interval. They suggested that the enhancement of the amplitude in nogo stimuli is related to the cortical processes for active inhibition of the subject's response. Several other studies have reported the recruitment of early inhibitory processes occurring within 200 ms of stimulus onset. Kirmizi-Alsan et al. (2006) studied the markers of response inhibition in a visual go-nogo task. They observed a significant amplitude increase in the nogo N1 than go N1. They suggested that participants must recruit inhibitory processes to withhold the nogo response and N1 is an indicator of response inhibition despite its early occurrence. In another study by Thomas and colleagues (2009), the level of inhibition required to withhold the nogo response was manipulated by varying the number of preceding go trials. Thus, a higher number of consecutive go trials before a nogo trial increased the inhibitory load. They found increased latency of N1 and P2 in the first 200 ms in the nogo trials. With these findings they supported a potential relationship between these components and inhibition (Fallgatter and Strik, 1999). Furthermore, Bekker et al. (2005) found a larger N1 in successful compared with failed stops in an auditory stop-signal task. In a similar vein, N1 was found to be an index of automatic inhibition that would be recruited when objects right after their presentation would trigger the actions they are associated to (Touzel et al., 2018; Franca et al., 2012).

2.4.2.4 The current study: The present study was aimed at examining the electrophysiological basis of the processes by which self and other's performed movements come to influence individual's visuo-spatial memory. Two spatial arrays differing in shape (one made up of circles and the other made up of squares, presented as a black outline against a white background) containing three items each were presented to participants. Participants were instructed to encode one array by visually observing the item locations while the other array was encoded by observation accompanied by pointing movements towards the item locations. The pointing movements were either performed by the participant or by the task-partner in a turn-taking condition. Simultaneous EEG was recorded from the participant, but not from the task-partner. A recognition test followed the encoding phase which was always answered by the participants. We first expected to replicate the behavioral results obtained in our earlier experiments involving the joint pointing condition, that is, the memory advantage for move arrays should appear over no-move arrays in both types of trials, one when the move arrays were pointed by participants themselves and the other when they were pointed by the task-partner. Second, we expected motor preparatory processes during the self-performed pointing movements reflected in the LRPs and also in CNV during the agent-cue presentation. Third, we expected to better understand the processing of item-locations, both for those encoded through pointing movements and for those encoded through mere visual observation. We predicted an inhibition related ERP component such as N1 or N2 to appear during the move arrays in the confederate-pointing trials as the participants would require inhibiting the motor representations activated upon the presentation of partner's relevant stimuli. Moreover, attention and cognitive processing related ERP components such as P2 and P3 are expected to appear for the move and no-move arrays during both types of trials, one with self-performed and the other with confederate performed pointing movements. Therefore, using ERPs we aimed at investigating whether item-locations pointed by self and others are processed similarly during encoding, which would further help us to understand how others movements come to benefit individual's memory.

2.4.3 Method and materials:

2.4.3.1 Participants: 23 right- handed pair of participants (27 Female, 19 Male, mean age = 24.9 years, SD = 4.9 years) were recruited. Out of 23 pair of participants, 3 male EEG participants were paired with male confederates, 7 female EEG participants were paired with female confederates, 9 female EEG participants were paired with male confederates and 4 male EEG participants were paired with female confederates. All participants either had normal or corrected to normal vision. All the participants provided their written consent after full explanation of the study. The study was approved by the United Ethical Review Committee for Research in Psychology **(**EPKEP**)** (Ref. No. 2019/43).

2.4.3.2 Experimental set-up: The experiment was run in a quiet, normal illuminated room. Participants were sitting side by side and EEG was recorded from one of them while the other played the confederate. An LED touch-screen monitor (DELL S2240T, 54.6 cm, screen resolution = 1920 x 1080) was used to present stimuli to the participants. The responses for the memory test were recorded by a response box (The Black Box Toolkit Ltd., Sheffield, UK). The experimental sessions were monitored by a video camera (HERO4, GoPro Inc.) to ensure that the participants were correctly carrying out the task.

2.4.3.3 Design and Materials: The experiment followed a 2 (Array order: first vs. second) x 2 (Condition: move vs. no-move) x 2 (Pointing agent: participant vs. confederate) x 2 (Pointing hand: Left vs. Right) design where all variables were manipulated within participants.

Stimuli were 64 visuo-spatial arrays, containing 3 items each, arranged in a 5 x 5 matrix which was invisible to participants. Of these, 32 arrays contained white-filled circles whereas the other 32 contained white-filled squares. Both circles and squares were 2 x 2 cm in size, with two adjacent items separated by 1 cm (See Rossi- Arnaud, Spataro and Longobardi, 2012). For each visuo-spatial array, a test lure was constructed, having all items at the same position as the original array except one, which was shifted by one or two positions. All the stimuli were presented as black outlines on a white background.

2.4.4 Procedure: The general procedure mimicked the original joint- pointing experiment reported earlier (see Experiment 2.1.5). Each trial comprised an encoding phase and a test phase. In the encoding phase, two visuo-spatial arrays, made up of white-filled squares and circles, containing 3 items each, were presented to the participants. In the test phase, the memory for one of the arrays locations was tested. Each item in the array was presented for 1000 ms, in a sequential order such that each subsequent item appeared when the previous item disappeared. The sampling of locations for each item was completely random, such that, no item appeared on a previously occupied location in a given trial. The order of presentation of arrays was counterbalanced within participants: arrays made up of circles were presented first for half of the trials followed by the arrays made up of squares, the opposite was true for the other half of the trials.

The main participant and the confederate sat next to each other, in front of the computer screen. EEG was recorded from the main participant who carried out the memory task throughout the experiment. Participants were shown an illustration of the procedure at the beginning of each experimental session and instructed to pay attention to the locations occupied by both the arrays because their memory for any one of the two arrays will be tested at the end of each trial. Each trial in the encoding phase contained both a move and a no-move array. Participants were instructed to point until touching to the locations of move arrays in addition to visually observing them whereas the no-move arrays were encoded only through visual observation. Therefore, move arrays were encoded through both a motor and a perceptual code whereas the no-move arrays were encoded only through a perceptual code. The shape occupied by move and no-move arrays was communicated at the beginning of each experimental session. For half of the participants, squares were move-arrays and the circles were no-move arrays whereas the opposite was true for the other half of the participants, which was communicated at the beginning of the experiment. During the encoding phase, each trial was preceded by a cue (a single uppercase letter: L for the person sitting on the left side and R for the person sitting on the right side) presented for 1000 ms, which indicated who had to touch the items for the move array in the following trial. Both the L-cued and the R-cued trials had both a move and a no-move array. Thus, it was instructed that the cue is to indicate their respective turns to point to the items in move array for a given trial; although the memory task was always answered by the EEG participant irrespective of who pointed to the items in the move arrays. The participant and the confederate pointed to the same types of items (either squares or circles) which were communicated at the beginning of each experimental session. The presentation of the cue L and R was completely random. The participant and the confederate both executed pointing movements on the half of the trials (L-cue was presented for 16 trials whereas R-cue was presented on other 16 trials in a random order). Each experimental block contained 32 trials followed by a break, comprising a total of 256 trials, divided into 8 experimental blocks of about 7.5 min. The EEG participant was seated on the Left side of screen and the confederate was seated on the Right side for first 4 blocks; they were instructed to make pointing movement towards the move array with the index finger of their right hand. The EEG participant was instructed to make the response with his left hand for the memory task. Their positions were switched in later 4 blocks; such that the EEG participant was sitting on the Right side of the screen and the confederate was sitting on the Left side, they were now instructed to make pointing movement towards move array with the index finger of their left hand. The EEG participant was instructed to make the response to the memory test with his left hand in the later 4 blocks. Blocks with the Left-Hand pointing condition were performed first followed by the blocks with the Right-Hand pointing condition for the half of the participants and the order was reversed for the other half of the participants. The presentation of the agent-cue was followed by a start button on the screen which would be touched by participants to move on. A blank interval of 500 ms was added at the beginning of the presentation of items in move array to have a brief baseline period for the motor activity obtained for the move arrays. The inter-trial-interval was varied between 1500- 1800 ms.

The test phase began immediately after the presentation of the second array, starting with the blank screen appearing for 1000 ms. In the test phase, participants' memory for the locations of the items from either the move or the no-move array was tested. Thus, each test screen contained all the items from one of the two presented arrays (either squares or circles). Participants did not know in advance which array will be tested; therefore, they were instructed to memorize both arrays at the encoding phase. They were instructed to judge whether the locations of the items at test match with what has been shown at encoding, by pressing the keys

‘S’ and ‘D’ for ‘Same’ and ‘Different’ respectively. The response keys were counterbalanced across blocks. The time limit for responses was set to 5000 ms. On half of the trials, the locations of the items in the test screen, matched those presented during the encoding phase, making a ‘Same’ response, while, on the other half of the trials, one item was shifted towards a previously unoccupied location, marking a ‘Different’ response. A sample trial is shown in **Figure 2.15** below.

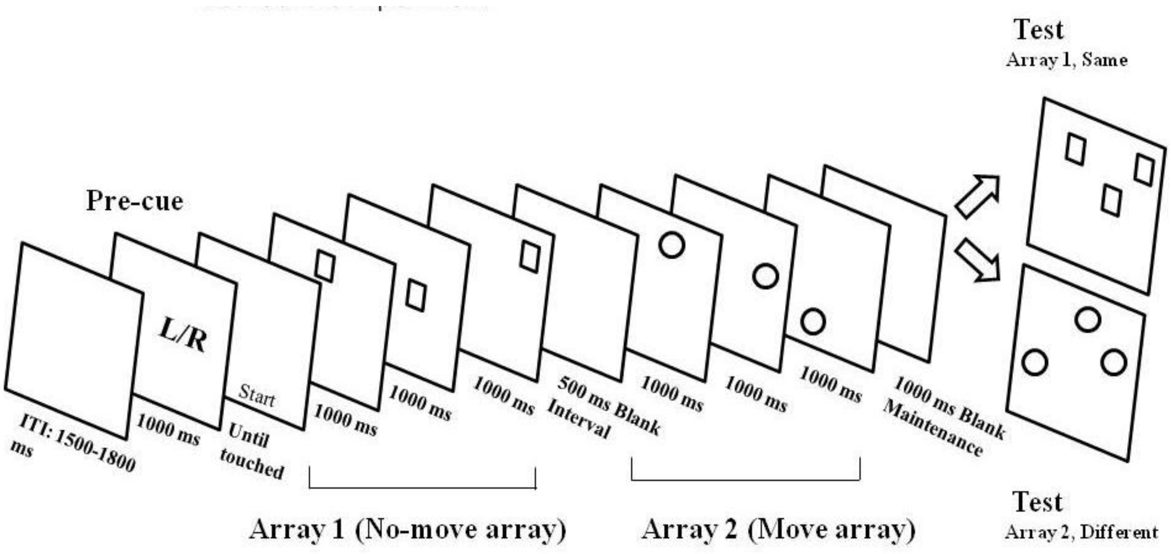


Figure 2.15: Illustration of a sample trial of EEG experiment. Participant and confederate were seated next to each other in front of a computer touch-screen. Two spatial arrays containing three items each are presented in a sequential presentation. One array is encoded through visual observation only (called no-move array) while the other is encoded through visual observation accompanied by pointing movements until touching. An agent-cue is presented for 1s to indicate who will perform the pointing movement towards the items in move array in the following trial. Intertrial interval varied between 1500 to 1800 ms. A blank interval of 500 ms is added before the presentation of move array. The encoding phase is followed by the recognition test after a 1s of maintenance phase. Participant made a same-different judgement by pressing a key on the response box.

2.4.4.1 EEG Data Acquisition: EEG was recorded from one participant using 64 active electrodes (Acticap, Brain Products, GmbH, Germany) arranged according to the standard 10-20 electrode placement system using the carefully positioned nylon cap. All electrodes were referenced to the right mastoid during the recording. Horizontal and Vertical eye movements were recorded using a pair of electrodes: one positioned lateral to the left and right eyes and one placed on the top and bottom of the left eye. Impedance for all the electrodes was kept below 25kΩ. EEG and ocular signals were amplified with a band-pass filter of 0.01-250 Hz by two BrainAmp DC amplifiers (BrainProducts, GmbH, Germany) and sampled at 500 Hz. The ground electrode was placed on the forehead.

2.4.5 EEG Data Processing and analyses: Trials during which a participant did not perform the pointing movements or the action onset was not recorded in the given time (e.g. if the pointing was not performed within 1000 ms), were removed from the encoding phase analysis. Action onset (or Reaction Time) was defined by the time participants took to touch the target item in the move array after its presentation. The encoding phase EEG analysis was focused on the 6500 ms time period of the presentation of move and no-move array during which the pointing movements towards the move arrays were either performed by the participant or observed being performed by his partner. The time period before 200 ms of stimulus onset was considered as the baseline. No pointing movements were performed during the presentation of the no-move arrays. EEG data were first re-referenced to the mean of both mastoid electrodes. Ocular correction with independent component analysis was performed using the brain vision analyser software (version 2.1, Brain Products, GmbH, Germany). The data were then filtered using a low cut-off filter of 0.01 Hz and a high cut-off filter of 40 Hz to remove the low and excessive high frequency noise, respectively. The notch filter was applied at 50 Hz to remove the line noise. After the segmentation, semi-automatic artifact rejection was performed before averaging in order to remove trials containing remaining vertical eye movements or other EEG related artifacts. An epoch was rejected when the difference between the maximum and minimum value at a single channel exceeded 100 µV.

2.4.5.1 ERP components

2.4.5.1.1 Contingent Negative Variation: The CNV is analyzed during the agent-cue presentation in the present study. The waveforms from electrodes C1, C2, FCz and Cz were pooled together for agent (self and confederate) and hand (right and left) conditions. The amplitudes from the averaged waveforms were exported for the last 200 ms of the presentation of agent cue (Kourtis, Knoblich and Sebanz, 2010, 2013).

2.4.5.1.2 Lateralized readiness potential (LRP): The data was segmented 1000 ms prior and 800 ms after the touch (i.e. the moment participant's finger touched the screen) and processed. LRP was calculated from the grand averaged waveforms locked to the touch for self and confederate pointing condition separately for the order of presentation of move array (first and second). LRP was calculated from the electrodes over motor cortex by subtracting the ipsilateral activity from the contralateral activity (see, Coles, 1989). For the right hand pointing condition, the activity for electrode C4 was subtracted from the activity for electrode C3 and similarly activity for C6 from activity of C5 (C3 minus C4 and C5 minus C6). For the left hand pointing condition, the activity for electrode C3 was subtracted from the activity for electrode C4 and similarly activity for C5 from activity for C6 (C4 minus C3 and C6 minus C5). 200 ms prior to the onset of stimulus was taken as the baseline period. The LRP was averaged for the order of presentation. The amplitudes for the LRPs obtained for the self-pointing and confederate pointing were exported for the 500 ms prior to the touch during which the motor activity is expected for the move arrays. The data was averaged across the presentation order and scalp sites (averaged for LRPs obtained as C4 minus C3 and C6 minus C5).

2.4.5.1.3 N1, P2, P3 and LPP: The following six anterior sites were selected for statistical analyses for anterior N1, P2 and LPP: F3, FC3 (Left); Fz, FCz (middle/midline); F4, FC4 (Right) (Luck and Kappenman, 2011). Time windows for anterior N1, P2 and LPP were 120-200 ms, 180-250 ms and 400-600 ms after stimulus onset (Benikos et al., 2013; Ma et al., 2019). Posterior N1 and P3 were derived from the following posterior sites: P3, PO2 (Left); POz, Oz (middle); P4, PO4 (Right) (Ma et al., 2019; Wronka, Kaiser and Coenen, 2012). Time windows for N1 and P3 were 160-220 ms (Ma et al., 2019) and 230-400 ms (Wronka et al., 2012) respectively.

2.4.6 Results and discussion:

2.4.6.1 Behavioral Results: One person’s data was removed from further analysis due to very low memory performance (below chance). Recognition Accuracy (proportion of hits) was analysed with a 2 (Order: first vs. second) x 2 (Condition: move vs. no-move) x 2 (Pointing agent: self vs. confederate) x 2 (Pointing hand: left vs. right) repeated measures ANOVA which revealed:

(a) A significant main effect of Condition [$F(1, 21) = 11.55$, $p < 0.01$, $\eta_p^2 = 0.35$], suggesting that the move arrays were remembered better (M = 78.56%) than the no-move arrays (M = 71.46%) (see **Figure 2.16**).

(b) A significant effect of Order [$F(1, 21) = 32.35$, $p < 0.001$, $\eta_p^2 = 0.60$], suggesting that the array presented lastly were remembered better (M= 81.83%) than the array presented firstly (M= 68.20%).

(c) A significant effect of Pointing agent [$F(1, 21) = 9.11$, $p < 0.01$, $\eta_p^2 = 0.30$], suggesting that the arrays relevant to self-pointing were remembered better (M = 77.69%) than the arrays relevant to confederate-pointing (M = 72.33%).

(d) A significant interaction between Order and Condition [$F(1, 21) = 8.51$, $p < 0.01$, $\eta_p^2 = 0.28$], a follow-up analysis showed that the effect of condition was significant only at order 1, i.e. participants remembered the move arrays higher (M = 75.08%) than the no-move arrays (M = 61.33%) only when they were presented first [$F(1, 21) = 16.27$, $p = 0.001$, $\eta_p^2 = 0.43$]. In contrast, no difference was found between the recognition accuracy for move (M = 82.05%) and no-move arrays (M = 81.6%) when they were presented second [$F(1, 21) = 0.03$, $p = 0.87$, $\eta_p^2 = 0.001$].

(e) A significant interaction between Order and Pointing agent [$F(1, 21) = 8.43$, $p < 0.01$, $\eta_p^2 = 0.28$], a follow-up analysis suggested a self-pointing advantage at order 1, i.e. participants remembered the items better when they touched them themselves (M = 72.80%) than when they were touched by the confederate (M = 63.60%) when these items were presented first [$F(1, 21) = 12.65$, $p = 0.002$, $\eta_p^2 = 0.37$]. In contrast, at order 2, no difference in the recognition accuracy was found when the items were touched by participants themselves (M = 82.59%) and when they were touched by the confederate (M = 81.06%), [$F(1, 21) = 0.74$, $p = 0.39$, $\eta_p^2 = 0.03$].

Higher order interactions were found and are reported in Appendix III (a significant three-way interaction between Order, Condition and Pointing hand [$F(1, 21) = 11.14$, $p < 0.01$, $\eta_p^2 = 0.34$], a significant higher order interaction between Order, Condition, Pointing agent and Pointing hand was observed [$F(1, 21) = 14.19$, $p = 0.001$, $\eta_p^2 = 0.40$].

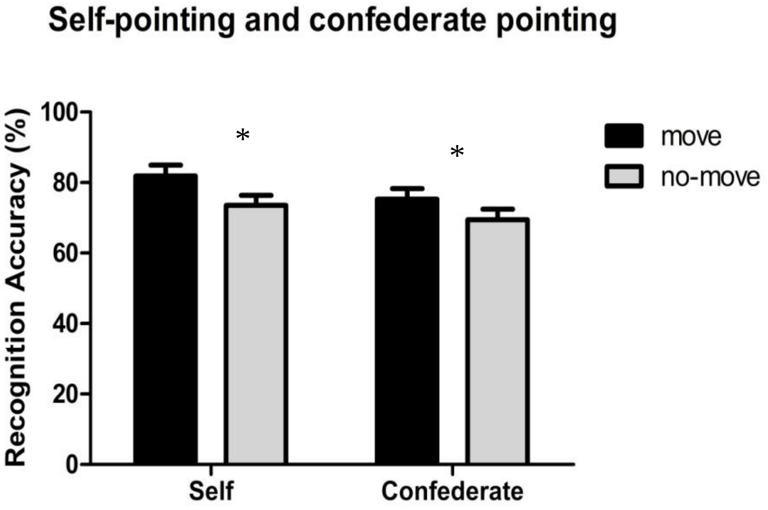


Figure 2.16: Recognition accuracy (%) as a function of condition (move vs. no-move) and pointing agent (self vs. confederate). Error bars represent standard error. * represents the significant differences between the two conditions, independently of the significance level.

The current findings first replicated the behavioral pattern of results as obtained in our earlier experiments involving a joint pointing condition. The main effect of condition indicated that the move arrays were remembered better than the no-move arrays. Most importantly, the interaction between the condition and pointing agent was not found to be significant, [$F(1, 21) = .80$, $p = 0.38$, $\eta_p^2 = 0.037$] suggesting that the memory advantage in the case of move arrays over no-move arrays was significant in both types of trials, one when the move arrays were pointed by participants themselves [$F(1, 21) = 11.56$, $p < 0.01$, $\eta_p^2 = 0.35$], and also when the move arrays were pointed by the confederate [$F(1, 21) = 5.07$, $p < 0.05$, $\eta_p^2 = 0.19$]. Moreover, the main effect of pointing agent was found significant indicating a self-pointing advantage over observed-pointing advantage.

2.4.6.2 ERP Results:

2.4.6.2.1: Motor preparation during self-performed pointing movements:

2.4.6.2.1.1 Contingent negative variation (CNV): The data was submitted to 2 (agent-cue: self vs. confederate) x 2 (pointing hand: left vs. right) repeated measures ANOVA which yielded a significant main effect of agent-cue [$F(1, 21) = 187.16$, $p < 0.001$, $\eta_p^2 = 0.899$]. Negative potential was found only for the presentation of agent-cue 'self' (M = -3.62) (see **Figure 2.17**).

2.4.6.2.1.2 Lateralized readiness potential (LRP): The data was submitted to a paired sample t-test which yielded a significant difference in the amplitudes for the LRPs obtained for self-pointing (M = -.35, SD = .81) and confederate pointing (M = .25, SD = .51) conditions, t(21) = -3.19, $p < 0.01$. (see **Figure 2.17)**

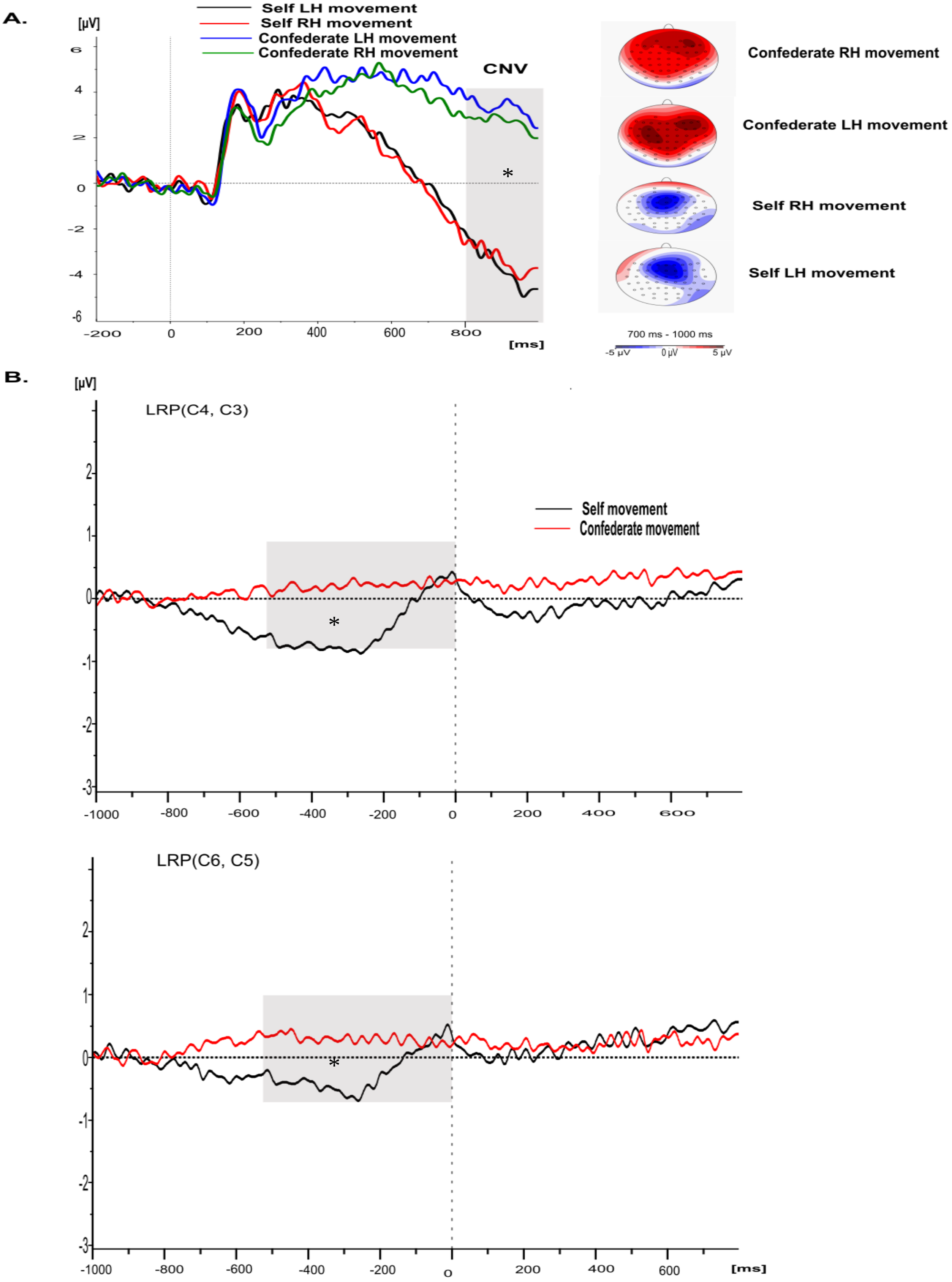

A.
[µV]
Self LH movement
Self RH movement
Confederate LH movement
Confederate RH movement
CNV
*
6
4
2
0
-2
-4
-6
-200
0
200
400
600
800
[ms]
Confederate RH movement
Confederate LH movement
Self RH movement
Self LH movement
700 ms - 1000 ms
-5 µV
0 µV
5 µV
B.
[µV]
LRP(C4, C3)
Self movement
Confederate movement
*
2
1
0
-1
-2
-3
-1000
-800
-600
-400
-200
0
200
400
600
LRP(C6, C5)
*
2
1
0
-1
-2
-3
-1000
-800
-600
-400
-200
0
200
400
600
[ms]

Figure: 2.17: (A.) Color coded, grand averages CNV waveforms derived from pooled electrode sites (C1, C2, FCz and Cz) and scalp topo-graphies from last 300 ms during the agent-cue presentation indicating the turn to perform pointing. (B.) Color coded LRP waveforms locked to the touch, derived for self and confederate pointing condition. * represents the significant differences between the two conditions, independently of the significance level.

2.4.6.2.2 N1: The data was submitted to 2 (agent: self vs. confederate) x 2 (condition: move vs. no-move) x 2 (pointing hand: left vs. right) x 3 (laterality: left vs. middle vs. right) repeated measures ANOVA. At the anterior sites, a significant interaction between the agent and condition was found [$F(1, 21) = 11.99$, $p < 0.01$, $\eta_p^2 = 0.36$], further analysis showed that the amplitude for N1 was significantly higher for the move arrays pointed by the confederate (M = -0.40 μV) as compared to the participant (M = 0.57 μV) [$F(1, 21) = 20.63$, $p < 0.001$, $\eta_p^2 = 0.49$]. Main effect of laterality was found [$F(2, 42) = 27.98$, $p < 0.001$, $\eta_p^2 = 0.57$] indicating a significantly negative potential over the right scalp sites (M = -0.15 μV) than the left (M = 0.45 μV) and middle (M = 0.79 μV) scalp sites. At the posterior sites, higher amplitude for move arrays (M = -0.93 μV) as compared to no-move arrays (M = 0.20 μV) was found, [$F(1, 21) = 34.92$, $p < 0.001$, $\eta_p^2 = 0.62$]. An interaction between the condition and laterality was found, [$F(2, 42) = 4.81$, $p < 0.05$, $\eta_p^2 = 0.187$], suggesting that significantly higher amplitude is found over right scalp sites than the midline scalp sites. An interaction between the pointing hand and laterality was found, [$F(2, 42) = 24.68$, $p < 0.001$, $\eta_p^2 = 0.54$], suggesting that significantly higher amplitude for N1 was found at the contra-lateral side to the pointing hand. A three way interaction among condition, laterality and pointing hand was found, [$F(2, 42) = 5.93$, $p < 0.01$, $\eta_p^2 = 0.22$]. No other effects were found significant.

2.4.6.2.3 P2: A 2 x 2 x 2 x 3 ANOVA with the same factors as above was conducted. Main effect of agent was found; P2 was more pronounced for the self-pointing movements (M = 2.24 μV) than confederate pointing movements (M = 1.55 μV), [$F(1, 21) = 13.70$, $p = 0.001$, $\eta_p^2 = 0.39$]. Main effect of condition was found; P2 amplitude was found to be higher for no-move arrays (M = 2.12 μV) than move arrays (M = 1.68 μV) [$F(1, 21) = 7.21$, $p = 0.01$, $\eta_p^2 = 0.25$]. The main effect of laterality was found; P2 amplitude was higher for midline electrodes than left and right, [$F(2, 42) = 20.30$, $p < 0.001$, $\eta_p^2 = 0.49$]. A significant interaction of agent and condition was found, [$F(1, 21) = 26.23$, $p < 0.001$, $\eta_p^2 = 0.55$]. P2 was more pronounced for the move arrays pointed by participant himself than by confederate [$F(1, 21) = 29.49$, $p < 0.001$, $\eta_p^2 = 0.58$], and for no-move arrays than move arrays for confederate pointing trials [$F(1, 21) = 34.43$, $p < 0.001$, $\eta_p^2 = 0.62$]. An interaction between hand and laterality was found, [$F(2, 42) = 4.51$, $p < 0.05$, $\eta_p^2 = 0.17$], which indicated that P2 amplitude was pronounced at central (midline) scalp sites as compared to left or right for both right and left hand pointing condition. A three way interaction between condition, hand and laterality was found, [$F(2, 42) = 6.74$, $p < 0.01$, $\eta_p^2 = 0.24$] and wasn't analyzed further.

2.4.6.2.4 P3: A 2 x 2 x 2 x 3 ANOVA with the same factors as above was conducted. Main effect of agent was found, higher amplitude for P3 was found for self-pointing trials (M = 1.68 μV) than for confederate pointing trials (M = 0.83 μV), [$F(1, 21) = 22.44$, $p < 0.001$, $\eta_p^2 = 0.51$]. Main effect of laterality was found, [$F(2, 42) = 3.39$, $p < 0.05$, $\eta_p^2 = 0.139$]; P3 amplitude for midline and left hemisphere was higher than the right hemisphere. A significant interaction of agent and condition was found, [$F(1, 21) = 8.69$, $p < 0.01$, $\eta_p^2 = 0.29$]. More pronounced P3 was found for move arrays from self-pointing trials than the move arrays from confederate pointing trials, $F(1, 21) = 22.59$, $p < 0.001$, $\eta_p^2 = 0.51$]. In the confederate pointing trials, no-move arrays elicited larger P3 than move arrays, $F(1, 21) = 10.33$, $p < 0.01$, $\eta_p^2 = 0.33$]. An interaction between condition and laterality was found, $F(2, 42) = 5.27$, $p < 0.01$, $\eta_p^2 = 0.20$], indicating significant higher distribution at midline than right scalp sites in the case of move arrays, and on the left scalp sites than midline and right sites in the case of no-move arrays. An interaction between pointing hand and laterality was found, $F(2, 42) = 3.57$, $p < 0.05$, $\eta_p^2 = 0.14$], indicating that higher distribution is found at midline and left sites than the right sites for both the right and the left hand pointing conditions.

2.4.6.2.5 Late positive potential: A 2 x 2 x 2 x 3 ANOVA with the same factors as above yielded a significant main effect of condition. No-move arrays elicited larger LPP (M = 1.06 μV) than move arrays (M = 0.05 μV), $F(1, 21) = 26.11$, $p < 0.001$, $\eta_p^2 = 0.55$].

No other effects or interactions were found significant.

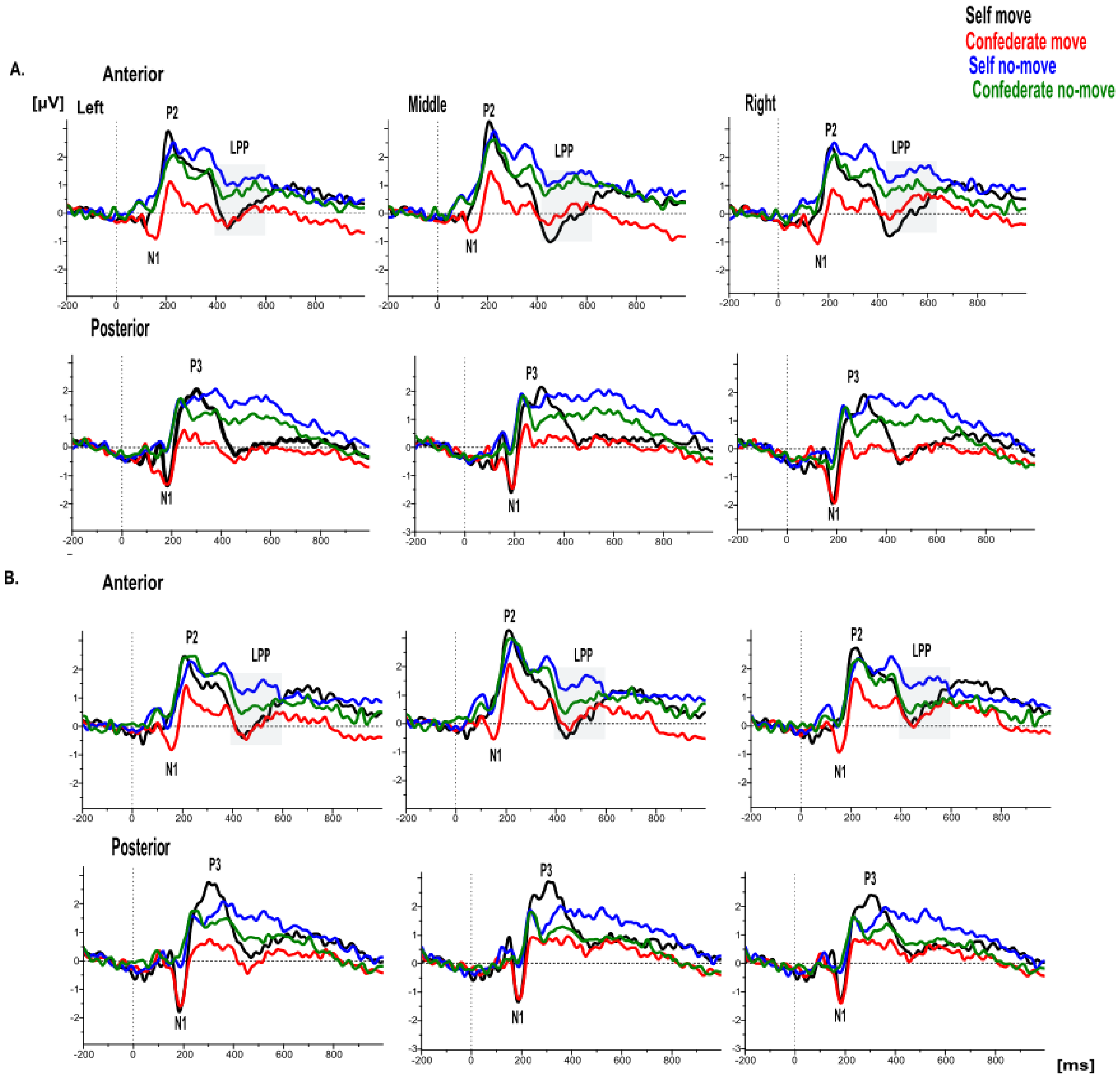


Figure 2.18: Color coded, ERP waveform for demonstrating anterior and posterior N1, anterior P2, posterior P3 and anterior LPP components. A. Left hand pointing condition B. Right hand pointing condition.

2.4.7 General Discussion: The present study was aimed at investigating the event related potential components related to the encoding processes of spatial locations in two arrays presented to the participants. Two visuo-spatial arrays containing three items each were presented in a sequential presentation, one of which was encoded by visual observation only (the no-move array) while the other was encoded by performing pointing movements towards target locations (the move array). Pointing movements towards the target locations in the move arrays were performed by the participant and the task-partner in a turn-taking context. An agent-cue was presented at the beginning of each trial indicating their respective turns to perform pointing movements in a given trial. The encoding phase was followed by a recognition test, where either the same or different item locations as shown in encoding were presented; and participants had to make same-different judgments.

With the recording of event related potentials, we were interested in examining the electrophysiology of a) how item locations pointed by participants themselves and by the task-partner are processed and also b) how the processing of item locations encoded by pointing differs from those encoded by only visual observation.

We firstly successfully replicated the behavioral findings obtained in our earlier experiments involving a joint pointing condition where participant and a task-partner take turns to perform pointing towards the move arrays. Their respective turns were indicated by the presentation of an agent-cue presented at the beginning of each trial, indicating who will execute the pointing movements in the following trial. Therefore, in each trial, one array was encoded through pointing movements either self-performed or performed by the task-partner, and the other array was encoded through mere visual observation. We found a memory advantage for the move arrays over no-move arrays in both types of trials, when the move arrays were pointed by the participant her/himself and also when they were pointed by the task-partner and participants were observing the pointing movements performed by other. The findings indicated that move arrays encoded through pointing movements both performed and observed elicited better encoding than arrays which were merely visually observed. Furthermore, a memory advantage was found for both move and no-move arrays in the self-pointing trials as compared to task-partner pointing trials. This finding is in line with several previous studies suggesting that self-relevant materials are processed and remembered better than self-irrelevant materials (Humphreys and Sui, 2016; Sui, He and Humphreys, 2012).

Secondly, we analyzed the motor preparation for the pointing movements performed by the participants towards the target locations in the move array. The LRPs reflected a motor preparation before the participant touched the item locations in the self-pointing trials. Moreover, the presentation of the agent-cue indicating participant’s turn to perform pointing movements in the following trial elicited a negative potential which was absent when the agent-cue indicated partner’s turn to point. This clearly reflects that perceiving the cue related to self-movement elicited preparatory processes whereas knowing the other’s turn to perform pointing did not commence preparation for a response. As previously mentioned in the introduction, the LRPs can be considered as an index of lateralized movement preparation and motor activation (Coles, 1989; de Jong, Wierda, Mulder and Mulder, 1988). The LRP reflects preparatory processes on a motor stage and the CNV indicates preparatory processes occurring prior to lateralized motor preparation. CNV has been shown to be linked with motor and non-motor preparatory processes (Ikeda et al., 1996; van Boxtel and Brunia, 1994a, 1994b) and response preparation (Leuthold et al., 2004; Ulrich, Leuthold and Sommer, 1998; Wild-Wall, Sangals, Sommer and Leuthold, 2003). The current data suggest that upon perceiving a stimulus related to self-movements, participants start preparing their response as reflected by the patterns of results obtained in CNV and LRPs.

Thirdly, an early component, N1 at anterior scalp sites was found only for the move arrays pointed by the confederate which was evident only at the right scalp sites. The finding that the N1 peak at the anterior sites was found only in the move arrays pointed by the task-partner may reflect that early inhibitory processes (Kirmizi-Alsan et al., 2006; Thomas et al., 2009; Nakata et al., 2004, 2005, 2006) were recruited during the presentation of move arrays which required an action from the task-partner. As speculated in our earlier experiments, where in the joint condition participants and the task-partner take turns to perform pointing movements, observing pointing movements performed by the partner might activate corresponding motor representations as self-performed movements. Such motor representations of partner's action plans will therefore require a strong response inhibition to prevent participants from performing pointing towards partner's relevant move arrays (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, & Tzeng, 2006). Previous studies have reported the N1 component reflecting an early recruitment of inhibitory processes in stop-signal and go-nogo tasks. Moreover, it has been reported that a more pronounced N1 component during nogo trials is obtained as compared to go trials in a go-nogo task (Kirmizi-Alsan et al., 2006; Thomas et al., 2009; Nakata et al., 2004, 2006a, 2006b; Benikos, Johnstone and Roodenrys, 2013) and stop-signal task (Bekker et al., 2005). Studies have suggested that inhibitory control can be exerted either proactively or reactively (Cai et al., 2011; Majid et al., 2013). Proactivity refers to a top-down signal transfer whereas the reactivity refers to a more bottom up activation of neural circuits (Kenemans, 2015). Studies have revealed right prefrontal ERP activity in response to the stop signal in stop-signal paradigms (Schmjuk et al., 2006; Jha et al., 2015). Schmjuk and colleagues using letter A and B as go stimuli and S for stop signal, reported a larger right frontal activity for the successful stops compared to failed stops. In another study, by Brown and Bever

(2005), right inferior frontal gyrus was found to be activated in a stop task variant where cues were used which preceded the go stimulus indicating that the stopping could be expected to be difficult or less difficult because of the duration of go-stop interval. Therefore, there is a possibility that early proactive inhibition was applied by participants to prevent themselves from performing pointing movements towards partner's action-relevant move arrays due to the activation of representational structures involved in their own execution of pointing movements.

As mentioned in the Introduction, such inhibition has been reported in earlier studies when participants perform a go-nogo task together with a partner. For instance, Sebanz et al., 2006 found larger amplitude of nogo-P3 in the joint condition of the go-nogo version of the Simon task than in the individual version. They suggested that perceiving a stimulus relevant to task-partner's actions leads to the activation of representational structures involved in one's own execution of this action. Therefore, increased response inhibition was required to prevent oneself from responding on nogo trials in the joint condition. Similarly Tsai et al., (2006), in addition to a larger nogo-P3 in the joint condition as compared to the individual condition, found nogo-N2 reflecting the response inhibition in the nogo-condition. Similarly, in the present study, the occurrence of N1 component only in the move arrays requiring pointing response from the task-partner may reflect that inhibitory processes were at play during the presentation of move arrays required to be pointed by the task-partner. This difference in the inhibitory component in our study and studies by Sebanz et al. (2006) and Tsai et al. (2006) could be because of the experimental design as the current study was not designed to focus on a specific ERP component but to better understand how the arrays pointed by self and partner are processed.

Next, we found interesting ERP components reflecting the stimulus processing and maintenance in the visuo-spatial working memory task used in the current study. The ERP components obtained in the present study provide further insights into the mechanisms involved in the encoding of move and no-move arrays action- relevant to both agents: to the participant her/himself and to the task-partner. We first found, higher amplitude of posterior N1 component for the move arrays as compared to the no-move arrays in both types of trials, when the move arrays were pointed by the participant and when the move arrays were pointed by the task-partner. It has been shown that visuo-spatial attention modulates subsequent stimulus processing, reflecting in early visual-evoked potentials such as P1 and N1 (Mangun and Hillyard, 1991; Mangun, 1995; Hillyard and Anllo-Vento, 1998). Several studies have suggested that these effects may reflect different aspects of attention, with P1 reflecting inhibition of processing of unattended stimuli and N1 reflecting amplification of processing of attended stimuli (Couperus and Mangun, 2010; Freunberger et al., 2008; Luck et al., 1994). The finding that the N1 amplitude is much higher for the move arrays pointed by both agents, the participant and the task-partner, suggests that move arrays differ from no-move arrays in early processing. It could be suggested that the processing of move arrays was facilitated compared to no-move arrays, reflecting that higher visuo-spatial attention was initially allocated to the move arrays pointed by both agents, possibly because they were action-relevant stimuli to self and partner in contrast to the no-move arrays which were action irrelevant (Vogel and Luck, 2000; Hopf, Vogel, Woodman, Heinze and Luck, 2002). This finding is consistent with the behavioral data showing that higher recognition accuracy is found for the move arrays pointed by both agents (participant and task-partner) as compared to the no-move arrays. Moreover, a laterality effect was found suggesting that the N1 attention effect was more pronounced over contralateral posterior brain regions that processes information from the relevant hand (for similar effects, see Slagter, Prinssen, Reteig and Mazaheri, 2016).

Second, we found a P2 component which was significantly larger for the self-relevant arrays (trials requiring self-pointing towards move arrays) than the task-partner relevant arrays (trials requiring pointing from task-partner towards move arrays). P2 component has been linked with higher order perceptual and attentional processing of visual stimulus (Luck and Hillyard, 1994). Moreover, it has been found that the amount of attention paid to the visual stimuli modulates the amplitude of P2 (Maeno et al., 2004; Johannes et al., 1995) which explains the presence of higher P2 amplitude in the self-relevant trials than in task-partner relevant trials: self-relevant trials were given significantly higher attention than the task-partner relevant trials. Furthermore, the interaction between the condition (move vs. no-move) and the pointing agent (self vs. confederate) revealed that in the case of move arrays, self-pointed arrays showed higher P2 amplitude than the move arrays pointed by the task-partner. This is intriguing because it may suggest that participants paid attention to the item locations in the move arrays when they had to also point to them (i.e. self-pointing condition), but did not pay attention to the item locations when they were pointed by the task-partner, instead they shifted their attention to the pointing movements performed by the partner. The fact that the move arrays pointed by the partner were still remembered better suggests that participants made use of the observed pointing movements to create better memory representations of these item locations.

On the contrary, no-move arrays in the partner pointing trials, elicited higher P2, which may suggest that the no-move arrays were given higher attention because they did not involve pointing movements that had to be observed as in the case of move arrays. Therefore, participants required higher attentional resources in the case of no-move arrays to form comparable memory representations. The effect of laterality in the case of P2, suggested that the distribution was more towards the midline scalp sites as compared to the left or right sites.

Third, the current findings showed that the self-pointing trials elicited larger P3 than the task-partner's pointing trials. Posterior P3 has been linked with the evaluation of stimuli with respect to action planning as shown in the go-nogo studies (Sebanz et al., 2006; Tsai et al., 2006; Kok, 2001). Moreover, it has also been shown that P3 amplitude is larger when subjects devote more effort to a task, suggesting that P3 is a measure of resource allocation (Isreal et al., 1980). Therefore, if a task is more difficult, this might increase P3 amplitude by encouraging subjects to devote more effort to the task. Likewise, self-pointing trials required higher allocation of cognitive resources because of the requirement to point towards the item-locations of move arrays compared to the partner's pointing condition. Further, we found that in the trials where move arrays were pointed by the task-partner, no-move elicited higher P3 amplitude than the move arrays, suggesting that since no-move arrays had to be encoded merely through visual observation they required higher cognitive resources than the move arrays. This is in line with the idea that participants rely on pointing movements performed by the partner and that the move arrays encoded by observing partner's movements therefore need less resource allocation; which subsequently leaves more cognitive resources available for the no-move arrays in task-partner pointing trials.

Finally, it was found that the no-move arrays from both pointing trials elicited larger late positive potential which evoked around 400 ms after stimulus onset compared to move arrays. Previous studies have suggested that the late positive potential (LPP) reflects information maintenance in a working memory task (Ma et al., 2019; Weinberg and Hajcak, 2011; Ruchkin, Ray Johnson, Canoune and Ritter, 1990). Therefore, it can be suggested that the no-move arrays required higher maintenance than move arrays because no-move arrays were only encoded by the visual observation, thus requiring higher maintenance in memory to be correctly recognized in the memory task.

In sum, the current electrophysiological findings shed light upon the processes involved in the encoding of spatial information encoded by visual observation or by performing pointing movements towards to-be-remembered spatial locations. Moreover, the current findings also reflect the processes through which the spatial locations pointed by self and task-partner are encoded which subsequently result in a similar memory performance. Altogether, the patterns of results suggest that self-performed pointing movements recruit higher cognitive and attentional resources, due to the additional requirement to perform pointing movements towards target locations, which leads to better encoding of these target locations as compared to visually observed target locations. On the contrary, target locations pointed by the task-partner require less cognitive resources (because there is no requirement to perform pointing). Not only this, but participants also shift their attention to the pointing movements (i.e. pointing finger) instead of the target locations which in turn, helps them form better memory representations of these target locations. Furthermore, since these target locations have better memory representations than the memory representations of visually observed target locations, they also need less maintenance for a recognition task.

The current study, to the best of our knowledge, is the first one showing how people process self and partner’s actions and integrate these actions into the encoding of spatial information. We conclude that people take partner’s actions into account while performing a pointing task together and that seems to subsequently benefit their memory performance.

Chapter 3: General Discussion

3.1 Introduction: The aim of the current dissertation was to investigate whether actions (specifically pointing movements) performed by a task-partner could influence individual's memory similar to self-performed movements. To this purpose, we conducted 10 experiments investigating how and when a task-partner's movements influence an individual's memory; findings are extended from behavioral to physiological data. We devised a joint-version of a visuo-spatial working memory task originally developed by Chum et al. (2007) (see Figure 1.1) in which the participant shared the pointing task with a task-partner. Two arrays were presented sequentially, one of which was encoded by visual observation only (named no-move array), whereas the other was encoded by visual observation accompanied by pointing movements (named move array). In the shared (or joint) version of this task, pointing movements were performed either by participant her/himself (named self-performed pointing movements) or by the task-partner (i.e. co-actor or confederate) by taking turns. An agent-cue was presented at the beginning of each trial indicating who will perform pointing movements in the following trial. The encoding of arrays was followed by a recognition test, which was always answered by the participants.

3.2 Behavioral findings: First, we replicated the experiment reported by Chum et al. (2007) indicating a self-performed pointing advantage in the case of move arrays over no-move arrays. Second, given that the observation of an action activates the same neuronal population as involved in the same action execution (Grèzes and Decety, 2001), we investigated whether passively observing the pointing movements performed by a task-partner would benefit participants' memory similarly to self-performed movements. Pointing movements were always carried out by the task-partner towards the move arrays and participants were instructed to passively observe pointing performed by the task-partner. The no-move arrays were encoded through visual observation only. It was expected that participants would simulate the pointing movements performed by the task-partner, which, in turn, would lead to a memory advantage for the move arrays over no-move arrays in the later array recognition task.

Contrary to our expectations, we found that passive observation of pointing performed solely by the task-partner did not influence the observers' (i.e. participants) memory. In this condition, no memory advantage for the move arrays pointed by the task-partner was found over the no-move arrays. This finding led to the possibility that the direct consequences of self-performed pointing movements are necessary in order to simulate co-actor's movements (Vesper et al., 2010). In other terms, participants must be actively engaged in the pointing task in order to have a memory advantage for the arrays pointed by the task-partner. Therefore, in a further experiment, a joint pointing condition was introduced where both the participant and the task-partner performed pointing movements alternatively. Participant and task-partner were sat next to each other and both were allowed to perform pointing movements in half of the task in a turn-taking condition. Their respective turns were indicated by an agent-cue (either P= participant or E= experimenter) presented at the beginning of each trial. Depending on the agent-cue presented at the beginning of each trial, either the participant or the task-partner had to perform pointing movements towards the move arrays in a given trial. Participants were also instructed to observe the pointing movements performed by the task-partner. The no-move arrays were always encoded through visual observation only. The memory task was always answered by the participants. Results of Experiment 3 (a random turn-taking condition) and Experiment 4 (a regular turn-taking condition) in Study 1 suggest that in such joint pointing conditions participants actually simulate and represent task-partner's movements as their own and this, in turn, led to a memory advantage for the move arrays performed by the task-partner (similar to that obtained with self-performed pointing movements) over no-move arrays.

Further, the memory performances between the experiments were compared: the memory performances for self-pointing and task-partner pointing condition in both Experiments Exp 3 and Exp 4 (joint pointing conditions) were compared against the baseline provided by the Exp 2 (where participants were instructed to passively observe the pointing movements performed by the task-partner). This cross-experimental analyses suggested that the memory advantage in the joint pointing conditions obtained in the case of move arrays was actually due to a motor inhibition of no-move arrays because of the instructions of no-pointing towards these arrays. Such an inhibition in the case of no-move arrays was found in both self-pointing and task-partner's pointing condition. It was suggested that in a joint pointing condition, the task-partner's movements were represented just as the participant's own movements and therefore produced comparable inhibitory effects on no-move arrays. Moreover, the findings from Study 1 (section 2.1) did not provide evidence in support of the explanation provided by Chum at al., (2007) that performing pointing movements towards an array enriched the visual memory trace with motor and kinesthetic information (Denis, Engelkamp, & Mohr, 1991; Engelkamp & Zimmer, 1989, 1997). The data from Study 1 did not provide evidence that performing active movements, or simulating them, during the encoding phase lead to the formation of more recallable memory traces (Chum et al., 2007). Instead, the data supported the view that the requirement to perform pointing movements towards a subset of items lead to an inhibition of the items not selected for an action (Dodd and Shumborksi, 2009). Moreover, the finding that a memory advantage in the case of move arrays over no-move arrays was found when the move arrays were pointed by the task-partner, similarly to what was shown with self-performed pointing movements, but only when the participant and the task-partner took turns to perform pointing movements, supported the view that motor resonance for observed human actions is critically dependent on the extent to which an experimental condition can be understood as a meaningful social interaction (Hogeveen & Obhi, 2012; Kourtis, Sebanz & Knoblich, 2010, 2013). Therefore, we proposed that the arrays pointed by the task-partner in the joint pointing condition were represented in the same functional way as self-pointed arrays, because participants experienced this condition as a social interaction and, therefore, co-represented their partner's pointing movements as if they were their own; this, in turn, produced comparable inhibitory effects on the no-move arrays.

Next, we dug a little further in the current dissertation to understand the factors which influence the memory advantage resulting from the pointing movements performed by the task-partner. First, we manipulated the frequency with which the participants performed the movements themselves, as opposed to the task-partner. By limiting the involvement on the participant's side, we aimed at understanding to what extent self-involvement from participants side is required in order to create a condition of social-context between the task-partners. The same joint (i.e. shared) version of the visuo-spatial working memory task was performed by the participant and a task-partner with the exception that the participants' turns to perform pointing movements were six times less than the task-partner's turn. We still found a memory advantage for the move arrays pointed by both the participant and the task-partner, over no-move arrays. Results from this experiment show that even a very limited number of trials with self-performed movements is sufficient to produce the memory advantage shown for movements performed by other. Results are also in line with the idea that a social-context is formed between the task-partners only when they perform the task together. This supports the conclusion that, in joint pointing conditions, participants represent pointing movements performed by the task-partner as their own, thus producing a memory advantage in the case of move arrays pointed by the task-partner similar to when they were pointed by participants themselves.

We then turned to analyze whether such an effect is unique to real life task-partners. Therefore, in a further experiment we replaced a real human task-partner with an image of a cartoon hand, giving it an animation effect to produce a movement to touch the target locations of items in the move arrays. Several studies have evidenced that the participants co-represent partner's actions even when they are interacting with non-biological agents (Tsai et al., 2008; Müller et al., 2011; Dolk et al., 2011, 2013). To investigate whether the memory advantage obtained in the case of move arrays pointed by the task-partner in our experiments is specific to a condition when the task-partner is a real life person (i.e. human), or it is a general effect which emerges as a consequence of social-context formed between the partners while performing a task in turns, we replaced the task-partner with an animated cartoon hand. In other terms, the participants performed the joint version of visuo-spatial working memory task together with a cartoon hand, where move arrays were either pointed by the participant her/himself or by the animated cartoon hand in a turn taking condition. The no-move arrays were always encoded by visual observation. Interestingly, a memory advantage occurred for the move arrays over no-move arrays when they were pointed by the cartoon hand, similar to the memory advantage obtained for the move arrays which were pointed by participants themselves.

Further in a control experiment, it was tested whether the memory advantage obtained for the move arrays pointed by the cartoon hand, was just a matter of increased saliency of these arrays due to the presence of the hand with respect to the no-move arrays. The same visuo-spatial working memory task was performed by the participants but pointing towards the item locations in the move arrays was always performed by the cartoon hand. Participants were instructed to passively observe pointing performed by the cartoon hand and try to remember the item locations in both move and no-move arrays. In this case, we did not find a memory advantage for the move arrays pointed by the cartoon hand over no-move arrays. In fact, the results were similar to the one obtained in the condition we had in our earlier experiment in Study 1, where the task-partner (i.e. the experimenter) performed the pointing movements towards move arrays in all the trials and participants were instructed to observe pointing performed by the experimenter. This suggests that the animacy of the task-partner does not seem to really matter because the similar patterns of memory performance were found when participants were instructed to passively observe pointing movements performed by the task-partner irrespective of whether the task-partner was a real life task-partner or a cartoon hand. Similarly, in the turn-taking condition, the patterns of memory performance were found to be similar when the move arrays were pointed by participant himself and the task-partner irrespective of the animacy of the task-partner. A plausible explanation is provided by the task-co-representation account which suggests that participants take their partner’s actions into account when performing a task together and represent them as their own actions and such task-co-representations are modulated by the social-context formed between them while performing a task together. Such task co-representations have been found to be formed when performing a task together with a biological agent (if not a real person) (Tsai and Brass, 2007; Müller et al., 2011) and even when the participants merely believed to be performing the task together with a partner (Atmaca, Sebanz and Knoblich, 2011; Kuhlen and Rahman, 2017). It is worth noting that the cartoon hand used in the current study had a resemblance to a human hand and was not like a wooden or Pinocchio’s hand, as used in several studies which claimed not to have an effect when participants perform a task together with a non-biological agent (Kilner, Paulignan and Blakemore, 2003; Tai et al., 2004; Tsai and Brass, 2007). It is possible that the cartoon hand used in the current study, due to its biological appearance, was able to produce motor representations similar to the motor representations of self-performed movements. Moreover, it is therefore possible that a social-context was formed between the participant and the cartoon hand while performing the task together due to the cartoon hand's biological appearance. Therefore, to completely rule out the possibility that the animacy of the task-partner doesn't really matter in a shared version of a VSWM task as used in the current studies, further studies are needed when the task partner should be replaced by a non-biological agent such as a wooden hand or a Japanese waving cat. As if now, evidence suggests that the critical factor in order to have a representation of the partner's actions similar to own motor representations, is the social-context between the task-partners in which the personal involvement in the pointing task from the participant's side is a prerequisite.

3.3 Eye-tracking and Event related potential findings:

We replicated the joint pointing condition (as in Experiment 3 in section 2.1) and the passive observation condition (as in Experiment 2 in section 2.1) with eye-tracking. The patterns of eye-movements for the move and no-move arrays were investigated in two conditions: Experiment 1 where the pointing movements were performed by both the participant and the task-partner in a random turn-taking condition, and Experiment 2 where *only* the task-partner performed the pointing movements and participants were instructed to passively observe them. Previous studies have reported that the patterns of eye movements reflect sequential action planning and control (Land, 2006; Land and McLeod, 2000; Land and Furneaux, 1997; Land, Mennie and Rusted, 1999; Ballard, Hayhoe, Li and Whitehead, 1992; Hayhoe, Bensinger and Ballard, 1998). Moreover, it has been demonstrated that the eye movements performed during action observation are strictly dependent on the activation of corresponding motor plans in the observer's mind (Flanagan & Johansson, 2003; Rotman, Troje, Johansson, & Flanagan, 2006). With the series of behavioral findings, we speculated that participants co-represent partner's movements as their own and that in turn, influences their visuo-spatial memory similarly to their own movements in a joint pointing condition. We further suggested that the personal involvement from the participant's side in the pointing task is a prerequisite for participants to co-represent the partner's actions as their own as it creates a social-context between the task-partners which is only formed while performing a task together in a turn-taking condition. Based on these assumptions, with eye-tracking we expected highly similar patterns of eye movements in both types of trials: when the participant performed pointing movements and when the task-partner performed pointing movements, but only in the joint pointing condition (i.e. turn-taking condition) of Experiment 1. That is, higher number of eye-fixations and longer gaze durations should be obtained for the move arrays as compared to the no-move arrays for the pointing movements performed by both agents. On the contrary, in Experiment 2, when participants were passive observers of pointing movements performed by the task-partner, we expected no differences in the patterns of eye movements for the move and no-move arrays because it was assumed that participants did not represent partner's movements in this condition. Interestingly, the results obtained from the eye-tracking study, were in line with our expectations. The patterns of eye movements found for the self-pointing condition were similar to the patterns of eye movements obtained for the task-partner pointing condition but only when participant and task-partner took turns in performing pointing towards the item-locations in move arrays. We found that eye-fixations were higher in number and gaze-durations were longer for the move arrays compared to the no-move arrays in both conditions, when the pointing towards move arrays was performed by the participant and when the pointing was performed by the task-partner. In contrast, the eye-fixations and gaze durations were found to be similar for the move and no-move arrays when the pointing movements were only performed by the task-partner and passively observed by participants. The patterns of results obtained in Study 2 provided further evidence in support of the assumptions made in Study 1, that is, arrays pointed by the task-partner in the joint pointing condition were represented in the same functional way as self-pointed arrays, because participants experienced this condition as a social interaction and, therefore, co-represented their partner's pointing movements as if they were their own. A point worth noting is that no robust inhibition for no-move arrays was found when a cross experimental analysis was conducted between the memory performances obtained in Experiment 1 (joint pointing condition) and Experiment 2 (passive observation condition). Instead we found that the memory advantage obtained in the case of move arrays was a result of facilitation of move arrays not of an inhibition of no-move arrays (see Appendix II for cross experimental analysis). This is in direct contrast with the findings obtained in section 2.1 where the memory advantage in the case of move arrays was entirely attributed to the inhibition of no-move arrays. Further studies are needed to clarify this issue and to better understand whether the memory advantage obtained in the case of move arrays is due to the facilitation of move arrays or an inhibition of no-move arrays or a combination of these two factors.

We further went on to investigate electro-physiologically the processes involved in the encoding of move and no-move arrays and whether the processing for move arrays pointed by participants themselves was similar to the processing of move arrays pointed by the task-partner.

We found that the encoding processes of move arrays differ from the processes involved in the encoding of no-move arrays. But also, we found that the encoding of move arrays pointed by the participants themselves were not exactly similar to the encoding of move arrays pointed by the task-partner. The event related potential (ERP) data revealed a negative peak at the right-frontal sites only for the move arrays pointed by the task partner suggesting that early inhibition was applied to inhibit the motor representations activated from the appearance of move arrays which required a pointing response from the task-partner. This further indicates that the target (i.e. item locations in the move arrays) requiring a response from the task-partner activates motor representations similar to those activated during self-performed pointing movements (Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Jing, Hung, & Tzeng, 2006).

Further the ERP data revealed that the amplitude of P2 and P3 components were significantly higher for move arrays pointed by participants themselves as compared to the move arrays pointed by the task-partner as they required higher attentional and cognitive resources because of the additional requirement to perform pointing movements towards the target locations. On the contrary, when the participants observed pointing movements performed by the task-partner towards the move arrays in the partner's pointing condition, it was found that participants required less attentional and cognitive resources to encode these arrays. Moreover, participants shifted their attention towards the pointing movements performed by the task-partner instead of the item locations. This was indicated by the lower amplitude of P2 and P3 components found in the case of move arrays pointed by the task-partner. Furthermore, with regard to the no-move arrays, it was found that they were maintained longer in memory than the move arrays, as indicated by the late positive potential. These findings suggest that pointing movements performed either by participants themselves or by the task-partner yielded better memory representations than the no-move arrays. One possible explanation is that performing pointing movements and observing movements being performed by the task-partner activated common motor representations which further enriched the visual memory with motor and kinesthetic traces, resulting in better recallable memory representations as compared to passively-viewed (no-move) arrays (Chum et al., 2007). We disregarded this explanation earlier with the findings obtained in Study 2.1 where the pointing movements did not really enhance the memory for item locations above the baseline level provided by the passive observation condition; instead the memory for items not selected for pointing (no-move arrays) was impaired. But as we said earlier, the replication of these experiments with eye-tracking yielded a significant enhancement in the memory for pointed-to arrays when compared to the baseline provided by the passive-observation condition. This discrepancy needs further investigation. The idea that performing pointing movements and observing pointing being performed by the task-partner activated common motor representations, thus yielding better memory traces for the move arrays, is also supported by the eye movements data reported in Experiment 2.2.3 (provided in the current dissertation). It has been suggested that observing an action performed by another person recruits eye movement programmes which are driven by observer's own motor representation of the same action (Flanagan and Johnasson, 2003; Rotman et al., 2006). There is evidence that the motor representations formed through action execution and action observation selectively interact with each other so that participants are found to be faster when performing an action compatible to an observed one and slower when they perform an action incompatible to the observed one (Brass, Bekkering and Prinz, 2001; Kilner, Pauligan and Blakemore, 2003). Similarly, Constantini and colleagues (Constantini, Ambrosini and Sinigaglia, 2012) investigated participants' gaze behavior when they held an object either with a compatible grip or an incompatible grip, or rested their hand freely, while observing an actor grasping a small or a large object (Pre-shape condition). In a control condition the actor merely touched the object (No shape condition). They found that when participants freely rested their right hand on the table, the proactivity (i.e. anticipating actions rather than just responding to stimuli (Land and Furneaux, 1997) of their gaze behavior was significantly higher when observing a pre-shaped hand grasping the target than when observing a closed fist merely touching it. On the contrary, their gaze behavior was less proactive when they observed the actor grasping the object with an incompatiable grip as to what their own right hand used to hold an object.

This finding that performing an incompatible grip (i.e. prehension) affects the patterns of eye movements to observe a hand action suggests that effective observation of an action might depend on how readily the motor representation of that action can be recruited (Busiello, Costantini, Galati and Committeri, 2011). It has been suggested that target-specific proactive eye movements are driven by the motor representations of an action during both observation and execution of a hand action (Ambrosini et al., 2011, 2012; Flanagan and Johansson, 2003; Rotman et al., 2006). In a similar vein, a recent study by Manzone, Cole, Skarratt and Welsh (2016) showed that the social inhibition of return (sIOR) effect was not found when one observes a response which is different from one's to-be-performed action. Pair of participants executed different actions – one participant executed key press while the other executed aiming movements to the same set of stimuli in a turn-taking condition. In the full vision condition participants were allowed to observe both the target stimuli and the response of the partner. In the partial vision condition participants were allowed to observe only the response of the partner but not the target stimuli. It was found that sIOR emerged in the full vision condition but not in the partial vision condition. The authors suggested that in the absence of attention capturing events and attentional shifts to the target location, the observation of a different response is not sufficient to activate the mechanisms leading to sIOR. In line with the common-coding and response co-representation approach, they suggested that due to the different observed action from the to-be-executed action, participants could not use the representation of partner's response to simulate their own performance of the task, therefore could not activate the processes leading to sIOR.

One could therefore assume that during the observation of pointing movements performed by the task-partner, participants recruited their own motor representations of the pointing movements, which yielded similar patterns of eye movements during self-pointing and partner's pointing condition, thus enriching the memory representations of move arrays with motor and kinesthetic information. On the contrary, in the passive observation condition, due to the requirement of not performing pointing movements, there is a fair possibility that participants could not match the motor representations of observed actions, which therefore yielded no differences in the patterns of eye movements during move and no-move arrays. In sum, the collective pattern of data obtained in four studies suggest that motor representations of pointing movements performed by the task-partner are mapped onto the similar motor representations activated during the self-pointing movements and together these motor representations form more recallable memory representations than visually-observed arrays. Collectively, the four sets of studies provide a reboust evidence of how spatial arrays are encoded using different modalities such as visual observation and pointing movements performed by self and the task-partner.

3.4 Limitations of the current studies: In this part, a number of limitations of the present study will be outlined and discussed. First the ERP data reported here suggests that move arrays were processed somewhat differently when they were pointed by participants themselves and when pointed by the task-partner. With self-performed pointing movements, participants allocated their attentional and cognitive resources towards the item locations but they shifted their attention towards the pointing movements, instead of the item locations, when the pointing movements were performed by the task-partner in a joint pointing condition.

We cannot provide evidence that this shift of attention towards the pointing movements performed by the task partner occurs only when they are performing the task together in a turn-taking context, and not when they passively observe the pointing performed by the task-partner. However, it seems unlikely that participants will not direct their attention towards the pointing movements performed by the task partner when they are passive observers of partner's movements because in both conditions they were instructed to observe the movements performed by the partner. So what needs to be explained is why in the condition when only the task-partner performs the pointing movements (and the participant passively observes), this do not lead to a memory advantage for pointed-to arrays. A possible explanation is that, even if participants attend the movements performed by the task-partner in the passive observation condition, they do not have motor representations that can be matched with the motor representations of the observed pointing movements, since they have never performed the pointing themselves. This is linked to the idea suggested that the participants' personal involvement in the pointing task seems a prerequisite for the memory advantage obtained in the case of move arrays pointed by the task-partner. The latter aspect can be examined in a future experiment replicating the passive observation condition with ERPs and we will not expect an inhibition component during the presentation of arrays requiring partner's movements. As speculated earlier in this chapter, future experiments will also need to be performed to examine how the animacy of the task-partner influences the motor representation system by manipulating the biological features of the task-partner.

Another interesting aspect concerns contingent negative variation. Previous studies have shown that contingent negative variation, a late ERP component also reflects action anticipation (Kourtis et al., 2010, 2013; Cohen and Walter, 1966; Gaillard, 1977; Loveless and Sanford, 1974; Tecce, 1972). It would be worth investigating whether participants anticipate an upcoming action to be performed by the task-partner. As the current data stands, we did not observe a negative going potential when participants anticipated the partner's actions indicated by the presentation of an agent-cue. The limitation of the paradigm used in current experiments may be a reason, as the current paradigm did not include trials where no-one points. The addition of trials with a no-pointing condition (where neither the participant nor the task-partner performs the pointing movements, but both arrays in a trial will be encoded by visual observation) would provide a baseline against which the negative rising potential can be analyzed indicating action anticipation for an upcoming action.

3.5 Conclusion:

In conclusion, the current dissertation by using different methodologies (behavioral, eye-tracking and EEG/ERP) present interesting data to carefully explore the research question addressed: how action (e.g. pointing movements) performed by self and others influence visuo-spatial memory. Using a joint/shared version of the visuo-spatial working memory task the current dissertation provided robust evidence that pointing movements performed by a task-partner influences one's visuo-spatial memory similar to self-performed movements. Taken together, the current dissertation has shown that people working together in a joint/shared context take partner's actions into account and that in turn, can influence their memory similar to their own actions. In a broader view, the current studies collectively add to the literature of memory and joint action by providing good evidence upon how self and other's movements are processed and influence one's visuo-spatial memory.

References

Allport, D. A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. *Perspectives on perception and action*, *15*, 395-419.

Ambrosini, E., Costantini, M., & Sinigaglia, C. (2011). Grasping with the eyes. *Journal of neurophysiology*, *106*(3), 1437-1442.

Ambrosini, E., Sinigaglia, C., & Costantini, M. (2012). Tie my hands, tie my eyes. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 263.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic bulletin & review*, *7*(3), 522-530.

Anisfield, M. (1979). Response to Meltzoff and Moore (1977). *Science*, *205*, 214.

Anllo-Vento, L., & Hillyard, S. A. (1996). Selective attention to the color and direction of moving stimuli: electrophysiological correlates of hierarchical feature selection. *Perception & psychophysics*, *58*(2), 191-206

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research , 211* (3-4), 371-385.

Atmaca, S., Sebanz, N., Prinz, W., & Knoblich, G. (2008). Action co-representation: The joint SNARC effect. *Social Neuroscience , 3* (3-4), 410-420.

Awh, E., Anllo-Vento, L., & Hillyard, S. A. (2000). The role of spatial selective attention in working memory for locations: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *12*(5), 840-847.

Awh, E., & Jonides, J. (1998). Spatial working memory and spatial selective attention.

Awh, E., Smith, E. E., & Jonides, J. (1995). Human rehearsal processes and the frontal lobes: PET evidence.

Bäckman, L., & Nilsson, L. G. (1985). Prerequisites for lack of age differences in memory performance. *Experimental Aging Research*, *11*(2), 67-73

Bäckman, L., Nilsson, L. G., & Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition*, *14*(4), 339-346.

BÄCKMAN, L., NILSSON, L. G., & NOURI, R. K. (1993). Attentional demands and recall of verbal and color information in action events. *Scandinavian journal of psychology*, *34*(3), 246-254.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual review of psychology*, *63*, 1-29.

Baddeley, A. D., Grant, S., Wight, E., & Thomson, N. (1975). Imagery and visual working memory. *Attention and performance V*, 205-217.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baddeley, A. D. (1990). Human memory: Theory and practice. Allyn & Bacon.

Baddeley, A. D., & Lieberman, K. (1980). Spatial working memory. In R. S. Nickerson (Ed.), Attention and performance VIII (pp. 521–539). Hillsdale, NJ: Erlbaum.

Baddeley, A., & Logie, R. (1992). Auditory imagery and working memory. In D. Reisberg (Ed.), *Auditory imagery* (p. 179–197). Lawrence Erlbaum Associates, Inc.

Ballard, D. H., Hayhoe, M. M., Li, F., & Whitehead, S. D. (1992). Hand-eye coordination during sequential tasks. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *337*(1281), 331-339.

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). " I show how you feel": Motor mimicry as a communicative act. *Journal of personality and social psychology*, *50*(2), 322.

Bekker, E. M., Kenemans, J. L., & Verbaten, M. N. (2005). Source analysis of the N2 in a cued Go/NoGo task. *Cognitive Brain Research*, *22*(2), 221-231.

Benikos, N., Johnstone, S., & Roodenrys, S. (2013). *Varying task difficulty in the Go/Nogo task: The effects of inhibitory control, arousal, and perceived effort on ERP components.*

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony.

Bertenthal, B. I., Longo, M. R., & Kosobud, A. (2006). Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 210.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, *2*(8), 561.

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clinical Neurophysiology*, *112*(12), 2224-2232.

Brass, M., Bekkering, H., & Prinz, W. (2001). *Movement observation a€ects movement execution in a simple response task.*

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: comparing symbolic, spatial, and imitative cues. *Brain and cognition*, *44*(2), 124-143.

Brooks, L. R. (1967). The suppression of visualization by reading. *The Quarterly journal of experimental psychology*, *19*(4), 289-299.

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*(5712), 1118-1121.

Bruin, K. J., & Wijers, A. A. (2002). Inhibition, response mode, and stimulus probability: a comparative event-related potential study. *Clinical Neurophysiology*, *113*(7), 1172-1182.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., ... & Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *European journal of neuroscience*, *13*(2), 400-404.

Buccino, G., Binkofski, F., & Riggio, L. (2004). The mirror neuron system and action recognition. *Brain and Language* , *89* (2), 370-376.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in cognitive sciences*, *10*(12), 551-557.

Busiello, M., Costantini, M., Galati, G., & Committeri, G. (2011). Sensory-motor interference abolishes repetition priming for observed actions, but not for action-related verbs. *Neuroscience letters*, *492*(2), 89-93.

Byrne, B. (1974). Item concreteness vs spatial organization as predictors of visual imagery. *Memory & Cognition*, *2*(1), 53-59.

Cai, W., Oldenkamp, C. L., & Aron, A. R. (2011). A proactive mechanism for selective suppression of response tendencies. *Journal of Neuroscience*, *31*(16), 5965-5969.

Chang, Y., & Choi, S. (2014). Effects of seductive details evidenced by gaze duration. *Neurobiology of learning and memory*, *109*, 131-138.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893.

Chelazzi, L., Biscaldi, M., Corbetta, M., Peru, A., Tassinari, G., & Berlucchi, G. (1995). Oculomotor activity and visual spatial attention. *Behavioural brain research*, *71*(1-2), 81-88.

Chum, M., Bekkering, H., Dodd, M., & Pratt, J. (2007). Motor and visual codes interact to facilitate visuospatial memory performance. *Psychonomic Bulletin and Review , 14* (6), 1189-1193.

Cochin, S. (1999). Observation and execution of movement: Similarities demonstrated by quantified electroencephalography. *European Journal of Neuroscience , 11* (5), 1839-1842.

Cohen, R. (1983). *The effect of encoding variables on the free recall of words and action events.*

Cohen, R. L., & Bean, G. (1983). Memory in educable mentally retarded adults: Deficit in subject or experimenter?. *Intelligence*, *7*(3), 287-298.

Cohen, R. L. (1989). Memory for action events: The power of enactment. *Educational psychology review*, *1*(1), 57-80.

Cohen, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, *22*(1), 267-281.

Cohen, R. L., Peterson, M., & Mantini-Atkinson, T. (1987). Interevent differences in event memory: Why are some events more recallable than others?. *Memory & Cognition*, *15*(2), 109-118.

Cohen, J., & Walter, W. G. (1966). The interaction of responses in the brain to semantic stimuli. *Psychophysiology*, *2*(3), 187-196.

Coles, R. (1989). *The call of stories: Teaching and the moral imagination*. Houghton Mifflin Harcourt.

Constable, M., Pratt, J., & Welsh, T. (2018). "Two minds don't blink alike": The attentional blink does not occur in a joint context. *Frontiers in Psychology , 9* (SEP).

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., ... & Shulman, G. L. (1998). A common network of functional areas for attention and eye movements. *Neuron*, *21*(4), 761-773.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1993). A PET study of visuospatial attention. *Journal of Neuroscience*, *13*(3), 1202-1226.

Costantini, M., Ambrosini, E., & Sinigaglia, C. (2012). Does how I look at what you're doing depend on what I'm doing?. *Acta Psychologica*, *141*(2), 199-204.

Couperus, J. W., & Mangun, G. R. (2010). Signal enhancement and suppression during visual–spatial selective attention. *Brain research*, *1359*, 155-177.

Craft, J. L., & Simon, J. R. (1970). Processing symbolic information from a visual display: interference from an irrelevant directional cue. *Journal of experimental psychology*, *83*(3p1), 415.

d'Ydewalle, G., & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European psychologist*, *12*(3), 196-205.

De Jong, R., Wierda, M., Mulder, G., & Mulder, L. J. M. (1988). The use of partial information in response preparation. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 682-692.

De Leonibus, E., Lafenetre, P., Oliverio, A., & Mele, A. (2003). Pharmacological evidence of the role of dorsal striatum in spatial memory consolidation in mice. *Behavioral neuroscience*, *117*(4), 685

De Leonibus, E., Oliverio, A., & Mele, A. (2005). A study on the role of the dorsal striatum and the nucleus accumbens in allocentric and egocentric spatial memory consolidation. *Learning and Memory , 12* (5), 491-503.

De Renzi, E., & Nichelli, P. (1975). Verbal and non-verbal short-term memory impairment following hemispheric damage. *Cortex*, *11*(4), 341-354.

Decety, J., & Grèzes, J. (2006). The power of simulation: imagining one's own and other's behavior. *Brain research*, *1079*(1), 4-14.

Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioural brain research*, *34*(1-2), 35-42.

Denis~, M., Engelkamp, J., & Mohr, G. (1991). *Psycho gical Research Memory of imagined actions: Imagining oneself or another person.*

Dick, M. B., Kean, M. L., & Sands, D. (1989). Memory for action events in Alzheimer-type dementia: Further evidence of an encoding failure. *Brain and Cognition*, *9*(1), 71-87.

DiMattia, B. V., & Kesner, R. P. (1988). Spatial cognitive maps: differential role of parietal cortex and hippocampal formation. *Behavioral neuroscience*, *102*(4), 471.

Dodd, M., & Shumborski, S. (2009). Examining the influence of action on spatial working memory: The importance of selection. *Quarterly Journal of Experimental Psychology , 62* (6), 1236-1247.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2011). How “social” is the social Simon effect?. *Frontiers in Psychology*, *2*, 84.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2014). The joint Simon effect: a review and theoretical integration. *Frontiers in Psychology*, *5*, 974.

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) social Simon effect: a referential coding account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1248.

Donchin, E. (1981). Surprise!… surprise?. *Psychophysiology*, *18*(5), 493-513.

Eimer, M. (1993). *Spatial cueing, sensory gating and selective response preparation: an ERP study on visuo-spatial orienting.*

Elekes, F., Bródy, G., Halász, E., & Király, I. (2016). Enhanced encoding of the co-actor's target stimuli during a shared non-motor task. *The Quarterly Journal of Experimental Psychology*, *69*(12), 2376-2389.

Engelkamp, J. (1990). Memory for action events: Some implications for memory theory and for imagery. In C. Cornoldi & M. A. McDaniel (Eds.),*Imagery and cognition* (pp. 183–219). New York, Berlin, Heidelberg: Springer.

Engelkamp, J. (1998). *Memory for actions*. Psychology Press/Taylor & Francis (UK).

Engelkamp, J. (1986). Nouns and verbs in paired-associate learning: Instructional effects. *Psychological Research*, *48*(3), 153-159.

Engelkamp, J., & Cohen, R. (1991). Current issues in memory of action events. *Psychological Research* , *53* (3), 175-182.

Engelkamp, J., & Dehn, D. (2000). Item and Order Information in Subject-Performed Tasks and Experimenter-Performed Tasks. *Journal of Experimental Psychology: Learning Memory and Cognition* , *26* (3), 671-682.

Engelkamp, J., & Krumnacker, H. (1980). Image-and motor-processes in the retention of verbal materials. *Zeitschrift für experimentelle und angewandte Psychologie*.

Engelkamp, J., & Zimmer, H. (1989, 12). Memory for action events: A new field of research. *Psychological Research* , *51* (4), pp. 153-157.

Engelkamp, J., & Zimmer, H. (1984). Motor programme information as a separable memory unit. *Psychological Research* , *46* (3), 283-299.

Engelkamp, J., & Zimmer, H. D. (1985). Motor programs and their relation to semantic memory. *German Journal of psychology*.

Engelkamp, J., & Zimmer, H. D. (1994). Motor similarity in subject-performed tasks. *Psychological Research*, *57*(1), 47-53.

Engelkamp, J., & Zimmer, H. (1997). Sensory factors in memory for subject-performed tasks. *Acta Psychologica* , *96* (1-2), 43-60.

Engelkamp, J., & Zimmer, H. D. (1983). Zum Einflu–S von Wahrnehmen und Tun auf das Behalten von Verb–Objekt-Phrasen. *Sprache & Kognition*.

Enticott, P. G., Johnston, P. J., Herring, S. E., Hoy, K. E., & Fitzgerald, P. B. (2008). Mirror neuron activation is associated with facial emotion processing. *Neuropsychologia*, *46*(11), 2851-2854.

Eskenazi, T., Doerrfeld, A., Logan, G. D., Knoblich, G., & Sebanz, N. (2013). Your words are my words: Effects of acting together on encoding. *Quarterly Journal of Experimental Psychology*, *66*(5), 1026-1034.

Evans, K. M., & Federmeier, K. D. (2007). The memory that's right and the memory that's left: Event-related potentials reveal hemispheric asymmetries in the encoding and retention of verbal information. *Neuropsychologia*, *45*(8), 1777-1790.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta psychologica*, *101*(2-3), 267-291.

Falkenstein, M., Koshlykova, N. A., Kiroj, V. N., Hoormann, J., & Hohnsbein, J. (1995). Late ERP components in visual and auditory Go/Nogo tasks. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *96*(1), 36-43.

Fallgatter, A. J., & Strik, W. K. (1999). The NoGo-anteriorization as a neurophysiological standard-index for cognitive response control. *International Journal of Psychophysiology*, *32*(3), 233-238.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Federmeier, K. D., & Benjamin, A. S. (2005). Hemispheric asymmetries in the time course of recognition memory. *Psychonomic bulletin & review*, *12*(6), 993-998.

Fischer, M. H., & Hoellen, N. (2004). Space-and object-based attention depend on motor intention. *Journal of General Psychology*, *131*(4), 365-378.

Flanagan, J., & Johansson, R. (2003). Action plans used in action observation. *Nature , 424* (6950), 769-771.

Franca, M., Turella, L., Canto, R., Brunelli, N., Allione, L., Andreasi, N. G., ... & Fadiga, L. (2012). Corticospinal facilitation during observation of graspable objects: a transcranial magnetic stimulation study. *PLoS One*, *7*(11), e49025.

Freunberger, R., Höller, Y., Griesmayr, B., Gruber, W., Sauseng, P., & Klimesch, W. (2008). Functional similarities between the P1 component and alpha oscillations. *European Journal of Neuroscience, 27*(9), 2330-2340.

Forrin, N. D., Ralph, B. C., Dhaliwal, N. K., Smilek, D., & MacLeod, C. M. (2019). Wait for it… performance anticipation reduces recognition memory. *Journal of Memory and Language, 109*, 104050.

Gaillard, A. W. K. (1977). The late CNV wave: Preparation versus expectancy. *Psychophysiology*, *14*(6), 563-568.

Galantucci, B., & Sebanz, N. (2009). Joint action: current perspectives. *Topics in Cognitive Science*, *1*(2), 255-259.

Galdo-Álvarez, S., & Carrillo-de-la-Peña, M. T. (2004). ERP evidence of MI activation without motor response execution. *Neuroreport*, *15*(13), 2067-2070.

Gallese, V., Eagle, M. N., & Migone, P. (2007). Intentional attunement: Mirror neurons and the neural underpinnings of interpersonal relations. *Journal of the American psychoanalytic Association*, *55*(1), 131-175.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). *Action recognition in the premotor cortex.*

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

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage*, *35*(4), 1674-1684.

Gazzola, V., Rizzolatti, G., Wicker, B., & Keysers, C. (2007). The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage , 35* (4), 1674-1684.

Gevins, A., Smith, M. E., Le, J., Leong, H., Bennett, J., Martin, N., ... & Whitfield, S. (1996). High resolution evoked potential imaging of the cortical dynamics of human working memory. *Electroencephalography and clinical Neurophysiology*, *98*(4), 327-348.

Georgescu, A. L., Kuzmanovic, B., Schilbach, L., Tepest, R., Kulbida, R., Bente, G., & Vogeley, K. (2013). Neural correlates of "social gaze" processing in high-functioning autism under systematic variation of gaze duration. *NeuroImage: Clinical*, *3*, 340-351.

Gonzalez, C. M. G., Clark, V. P., Fan, S., Luck, S. J., & Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain topography*, *7*(1), 41-51.

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

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human brain mapping*, *12*(1), 1-19.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: an fMRI study. *Neuroimage*, *21*(2), 744-750.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological review*, *77*(2), 73.

Guagnano, D., Rusconi, E., & Umiltà, C. A. (2010). Sharing a task or sharing space? On the effect of the confederate in action coding in a detection task. *Cognition*, *114*(3), 348-355.

Guillot, A., & Collet, C. (2005). Duration of mentally simulated movement: a review. *Journal of motor behavior*, *37*(1), 10-20.

Guillot, A., Lebon, F., Rouffet, D., Champely, S., Doyon, J., & Collet, C. (2007). Muscular responses during motor imagery as a function of muscle contraction types. *International Journal of Psychophysiology*, *66*(1), 18-27.

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings of the National Academy of Sciences , 95* (25), 15061-15065.

Harnishfeger, K. K. (1995). The development of cognitive inhibition: Theories, definitions, and research evidence. In *Interference and inhibition in cognition* (pp. 175-204). Academic Press.

Hayhoe, M., Bensinger, D., & Ballard, D. (1998). Task constraints in visual working memory. *Vision Research* .

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition*, *7*(1-3), 43-64.

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

Hogeveen, J., & Obhi, S. S. (2012). Social interaction enhances motor resonance for observed human actions. *Journal of Neuroscience*, *32*(17), 5984-5989.

Holländer, A., Jung, C., & Prinz, W. (2011). Covert motor activity on NoGo trials in a task sharing paradigm: Evidence from the lateralized readiness potential. *Experimental Brain Research , 211* (3-4), 345-356.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological research*, *59*(3), 176-186.

Hommel, B., Colzato, L., & Van Den Wildenberg, W. (2009). How social are task representations? *Psychological Science , 20* (7), 794-798.

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

Hopf, J. M., Vogel, E., Woodman, G., Heinze, H. J., & Luck, S. J. (2002). Localizing visual discrimination processes in time and space. *Journal of Neurophysiology*, *88*(4), 2088-2095.

Hornstein, S., & Mulligan, N. (2001). *Memory of action events: The role of objects in memory of self-and other-performed tasks.*

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: the Self-Attention Network (SAN). *Cognitive neuroscience*, *7*(1-4), 5-17.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, *20*(5), 497-514.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, *32*(4), 421-445.

Hunt, R. R., & Worthen, J. B. (2006). Distinctiveness and memory. New York, NY, US: Oxford University Press.

Iacoboni, M. (2005). Neural mechanisms of imitation. *Current opinion in neurobiology*, *15*(6), 632-637.

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

Iani, C., Anelli, F., Nicoletti, R., Arcuri, L., & Rubichi, S. (2011). The role of group membership on the modulation of joint action. *Experimental Brain Research*, *211*(3-4), 439.

Ikeda, A., Shibasaki, H., Kaji, R., Terada, K., Nagamine, T., Honda, M., ... & Kimura, J. (1996). Abnormal sensorimotor integration in writer's cramp: study of contingent negative variation. *Movement disorders: official journal of the Movement Disorder Society*, *11*(6), 683-690.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, *17*(3), 259-273.

James, W., Burkhardt, F., Bowers, F., & Skrupskelis, I. K. (1890). *The principles of psychology* (Vol. 1, No. 2). London: Macmillan.

Jansson, E., Wilson, A. D., Williams, J. H., & Mon-Williams, M. (2007). Methodological problems undermine tests of the ideo-motor conjecture. *Experimental Brain Research*, *182*(4), 549-558.

Jeannerod, M. (2006). *Motor cognition: What actions tell the self* (No. 42). Oxford University Press.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, *14*(1), S103-S109.

Jeannerod, M. (1999). The 25th Bartlett Lecture: To act or not to act: Perspectives on the representation of actions. *The Quarterly Journal of Experimental Psychology Section A*, *52*(1), 1-29.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, *17*(2), 187-202.

Jha, A., Nachev, P., Barnes, G., Husain, M., Brown, P., & Litvak, V. (2015). The frontal control of stopping. *Cerebral Cortex*, *25*(11), 4392-4406.

Johannes, S., Münte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, *2*(3), 189-205.

Johansson, R., & Johansson, M. (2014). Look here, eye movements play a functional role in memory retrieval. *Psychological Science*, *25*(1), 236-242.

Johansson, R. S., Westling, G., Bäckström, A., & Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, *21*(17), 6917-6932.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin*, *114*(1), 3.

Johnstone, S. J., Barry, R. J., Anderson, J. W., & Coyle, S. F. (1996). Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *International Journal of Psychophysiology*, *24*(3), 223-238.

Kawashima, R., Satoh, K., Itoh, H., Ono, S., Furumoto, S., Gotoh, R., ... & Yanagisawa, T. (1996). Functional anatomy of GO/NO-GO discrimination and response selection—a PET study in man. *Brain research*, *728*(1), 79-89.

Kenemans, J. L. (2015). Specific proactive and generic reactive inhibition. *Neuroscience & Biobehavioral Reviews*, *56*, 115-126.

Kilner, J., Paulignan, Y., & Blakemore, S. (2003). An Interference Effect of Observed Biological Movement on Action tral premotor cortex. These neurons discharge both when the monkey performs specific hand movements and when it observes another monkey or human per. *Current Biology , 13,* 522-525.

Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M., & Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain research*, *1104*(1), 114-128.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*(3), 99-104.

Kok, A. (1986). Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biological psychology*, *23*(1), 21-38.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557-577.

Kopp, B., Mattler, U., Goertz, R., & Rist, F. (1996). *N2, P3 and the lateralized readiness potential in a nogo task involving selective response priming.*

Koriat, A., Ben-Zur, H., & Druch, A. (1991). The contextualization of input and output events in memory. Psychological Research, 53, 260-270.

Kourtis, D., Sebanz, N., & Knoblich, G. (2010). Favouritism in the motor system: Social interaction modulates action simulation. *Biology Letters , 6* (6), 758-761.

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience , 8* (1), 31-42.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision research*, *35*(13), 1897-1916.

Kuhbandner, C., Pekrun, R., & Maier, M. A. (2010). The role of positive and negative affect in the “mirroring” of other persons' actions. *Cognition and Emotion*, *24*(7), 1182-1190.

Kuhlen, A., & Abdel Rahman, R. (2017). Having a task partner affects lexical retrieval: Spoken word production in shared task settings. *Cognition , 166*, 94-106.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in retinal and eye research*, *25*(3), 296-324.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1358), 1231-1239.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities?. *Vision research*, *41*(25-26), 3559-3565.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature neuroscience*, *3*(12), 1340.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311-1328.

Land, M. F. (2009). Vision, eye movements, and natural behavior. *Visual neuroscience*, *26*(1), 51-62.

Lavric, A., Pizzagalli, D. A., & Forstmeier, S. (2004). When 'go'and 'nogo'are equally frequent: ERP components and cortical tomography. *European Journal of Neuroscience*, *20*(9), 2483-2488.

LaFrance, M. (1982). Posture mirroring and rapport: Analysis by the cross-lag panel technique. Social Psychology Quarterly, 42, 66–70.

LA FRANCE, M. (1979). NONVERBAL SYNCHRONY AND RAPPORT: ANALYSIS BY THE CROSS-LAG PANEL TECHNIQUE.

LaFrance, M., & Broadbent, M. (1976). Group rapport: Posture sharing as a nonverbal indicator. Group and Organizational Studies, 1, 328–333.

Lange, N., Hollins, T., & Bach, P. (2017). Testing the motor simulation account of source errors for actions in recall. *Frontiers in Psychology* , *8* (SEP).

Leuthold, H., Sommer, W., & Ulrich, R. (2004). Preparing for Action: Inferences from CNV and LRP. *Journal of Psychophysiology* , *18*, 77-88.

Leynes, P. A., & Kakadia, B. (2013). Variations in retrieval monitoring during action memory judgments: Evidence from event-related potentials (ERPs). *International Journal of Psychophysiology*, *87*(2), 189-199.

Liepelt, R., Prinz, W., & Brass, M. (2010). When do we simulate non-human agents? Dissociating communicative and non-communicative actions. *Cognition* , *115* (3), 426-434.

Lindner, I., Echterhoff, G., Davidson, P., & Brand, M. (2010). Observation inflation: Your actions become mine. *Psychological Science* , *21* (9), 1291-1299.

Lindner, I., Schain, C., & Echterhoff, G. (2016). Other-self confusions in action memory: The role of motor processes. *Cognition* , *149*, 67-76.

Logie, R. H. (1995). *Essays in cognitive psychology. Visuo-spatial working memory.* Lawrence Erlbaum Associates, Inc.

Logie, R. H. (1991). Visuo-spatial short-term memory: Visual working memory or visual buffer?. In *Imagery and cognition* (pp. 77-102). Springer, New York, NY.

Logie, R. H., Zucco, G. M., & Baddeley, A. D. (1990). Interference with visual short-term memory. *Acta Psychologica*, *75*(1), 55-74.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive?. In *Advances in psychology* (Vol. 80, pp. 105-115). North-Holland.

Lotze, R. H. (2005). Excerpt from Medicinische Psychologie oder Physiologie der Seele. *History of Psychiatry*, *16*(1), 117-127.

Loveless, N. E., & Sanford, A. J. (1974). Slow potential correlates of preparatory set. *Biological psychology*, *1*(4), 303-314.

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Luck, S. J. (1998). Sources of dual-task interference: Evidence from human electrophysiology. *Psychological Science*, *9*(3), 223-227.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, *31*(3), 291-308.

Luck, S. J., & Kappenman, E. S. (Eds.). (2011). *The Oxford handbook of event-related potential components*. Oxford university press.

Ma, H., Zhang, D., Li, X., Ma, H., Wang, N., & Wang, Y. (2019). Long-term exposure to high altitude attenuates verbal and spatial working memory: Evidence from an event-related potential study. *Brain and behavior*, *9*(4), e01256.

Maeno, T., Gjini, K., Iramina, K., Eto, F., & Ueno, S. (2004, August). Event-related potential P2 derived from visual attention to the hemi-space. Source localization with LORETA. In *International Congress Series* (Vol. 1270, pp. 262-265). Elsevier.

Majid, D. A., Cai, W., Corey-Bloom, J., & Aron, A. R. (2013). Proactive selective response suppression is implemented via the basal ganglia. *Journal of Neuroscience*, *33*(33), 13259-13269.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, *32*(1), 4-18.

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human perception and performance*, *17*(4), 1057.

Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1993). IQ electrocortical substrates of visual selective attention. *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, *14*, 219.

Manzone, J., Cole, G. G., Skarratt, P. A., & Welsh, T. N. (2017). Response-specific effects in a joint action task: social inhibition of return effects do not emerge when observed and executed actions are different. *Psychological research*, *81*(5), 1059-1071.

McDaniel, M. A., & Bugg, J. M. (2008). Instability in memory phenomena: A common puzzle and a unifying explanation. *Psychonomic Bulletin & Review*, *15*(2), 237-255.

McEvoy, L. K., Smith, M. E., & Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. *Cerebral cortex (New York, NY: 1991)*, *8*(7), 563-574.

Miller, J., & Hackley, S. A. (1992). Electrophysiological evidence for temporal overlap among contingent mental processes. *Journal of Experimental Psychology: General*, *121*(2), 195.

Minelli, A., Marzi, C. A., & Girelli, M. (2007). Lateralized readiness potential elicited by undetected visual stimuli. *Experimental brain research*, *179*(4), 683-690.

Mohr, G., Engelkamp, J., & Zimmer, H. D. (1989). Recall and recognition of self-performed acts. *Psychological Research*, *51*(4), 181-187.

Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences*, *98*(3), 1273-1276.

MORRIS, R. G. Exploring the visuospatial scratch pad. Quarterly Joumol of Experimen~al Psychology, 1987, 39, 409-430.

Müller, B. C., Brass, M., Kühn, S., Tsai, C. C., Nieuwboer, W., Dijksterhuis, A., & van Baaren, R. B. (2011). When Pinocchio acts like a human, a wooden hand becomes embodied. Action co-representation for non-biological agents. *Neuropsychologia*, *49*(5), 1373-1377.

Mulligan, N. W., & Hornstein, S. L. (2003). Memory for actions: Self-performed tasks and the reenactment effect. *Memory & Cognition*, *31*(3), 412-421.

Mulligan, N. W., Smith, S. A., & Buchin, Z. L. (2018). The generation effect and experimental design. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Nakata, H., Inui, K., Nishihira, Y., Hatta, A., Sakamoto, M., Kida, T., ... & Kakigi, R. (2004). Effects of a go/nogo task on event-related potentials following somatosensory stimulation. *Clinical Neurophysiology*, *115*(2), 361-368.

Nakata, H., Inui, K., Wasaka, T., Tamura, Y., Kida, T., & Kakigi, R. (2005). Effects of ISI and stimulus probability on event-related go/nogo potentials after somatosensory stimulation. *Experimental brain research*, *162*(3), 293-299.

Nakata, H., Inui, K., Wasaka, T., Tamura, Y., Kida, T., & Kakigi, R. (2006). The characteristics of the nogo-N140 component in somatosensory go/nogo tasks. *Neuroscience letters*, *397*(3), 318-322.

Oades, R. D., Dittmann-Balcar, A., Zerbin, D., & Grzella, I. (1997). Impaired attention-dependent augmentation of MMN in nonparanoid vs paranoid schizophrenic patients: a comparison with obsessive-compulsive disorder and healthy subjects. *Biological psychiatry*, *41*(12), 1196-1210.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and clinical neurophysiology*, *60*(5), 423-434.

Press, C., Bird, G., Flach, R., & Heyes, C. (2005). Robotic movement elicits automatic imitation. *Cognitive Brain Research*, *25*(3), 632-640.Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Prinz, W. (1990). A common coding approach to perception and action. In *Relationships between perception and action* (pp. 167-201). Springer, Berlin, Heidelberg.

Prinz, W. (1997). Perception and Action Planning. *European Journal of Cognitive Psychology , 9* (2), 129-154.

Quinn, G. (1991). Encoding and maintenance of information in visual working memory. In *Advances in psychology* (Vol. 80, pp. 95-104). North-Holland.

Quinn, G. (1988). Interference effects in the visuo-spatial sketchpad. In *Cognitive and neuropsychological approaches to mental imagery* (pp. 181-189). Springer, Dordrecht.

Quinn, J. G. (2008). Movement and visual coding: The structure of visuo-spatial working memory. *Cognitive Processing*, *9*(1), 35-43.

Quinn, J. G., & Ralston, G. E. (1986). Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology Section A*, *38*(4), 689-703.

Ramnani, N., & Miall, R. C. (2004). A system in the human brain for predicting the actions of others. *Nature neuroscience*, *7*(1), 85.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology*, *62*(8), 1457-1506.

R.H. Lotze (1852). Medicinische psychologie oder physiologie der seele, Weidmann, Leipzig (1852)

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

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

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). *Premotor cortex and the recognition of motor actions.*

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (1999). Resonance behaviors and mirror neurons. *Archives italiennes de biologie*, *137*(2), 85-100.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental brain research*, *111*(2), 246-252.

Rizzolatti, G., & Matelli, M. (2003). Two different streams form the dorsal visual system: anatomy and functions. *Experimental brain research*, *153*(2), 146-157.

Rossi-Arnaud, C., Spataro, P., & Longobardi, E. (2012). Effects of pointing on the recall of simultaneous and sequential visuospatial arrays: A role for retrieval strategies? *Psychological Research , 76* (6), 699-712.

Rossi-Arnaud, C., Pieroni, L., Spataro, P., & Baddeley, A. (2012). Working memory and individual differences in the encoding of vertical, horizontal and diagonal symmetry. *Acta psychologica*, *141*(1), 122-132.

Rotman, G., Troje, N. F., Johansson, R. S., & Flanagan, J. R. (2006). Eye movements when observing predictable and unpredictable actions. *Journal of neurophysiology*, *96*(3), 1358-1369.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic bulletin & review*, *16*(2), 225-237.

Rubia, K., Smith, A. B., Brammer, M. J., & Taylor, E. (2003). Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection. *Neuroimage*, *20*(1), 351-358.

Ruchkin, D. S., Johnson Jr, R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and clinical Neurophysiology*, *76*(5), 419-439.

SALTZ, E. (1988). The role of motoric enactment (m-processing) in memory for words and sentences. In M. M. Gruneberg, P.E. Morris, & R. N. Sykes (Eds.), Practical aspects ofmemory: Current research and issues (Vol. I, pp. 408-414). Chichester, U.K.: Wiley.

Saltz, E., & Donnenwerth-Nolan, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning and Verbal Behavior*, *20*(3), 322-332.

Sapkota, R., Pardhan, S., & Van der Linde, I. (2013). Manual tapping enhances visual short-term memory performance where visual and motor coordinates correspond. *British Journal of Psychology , 104* (2), 249-264.

Save, E., & Moghaddam, M. (1996). Effects of lesions of the associative parietal cortex on the acquisition and use of spatial memory in egocentric and allocentric navigation tasks in the rat. *Behavioral Neuroscience , 110* (1), 74-85.

Schmajuk, M., Liotti, M., Busse, L., & Woldorff, M. G. (2006). Electrophysiological activity underlying inhibitory control processes in normal adults. *Neuropsychologia*, *44*(3), 384-395.

Schult, J. C., & Steffens, M. C. (2011). On the representation of intentions: Do personally relevant consequences determine activation?. *Memory & cognition*, *39*(8), 1487.

Schult, J. C., & Steffens, M. C. (2013). Tuned for the future: Intentions are only accessible when a retrieval opportunity is near. *Memory & cognition*, *41*(8), 1252-1260.

Sebanz, N., & Knoblich, G. (2009). Prediction in Joint Action: What, When, and Where. *Topics in Cognitive Science , 1* (2), 353-367.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in cognitive sciences , 10* (2), 70-6.

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance , 31* (6), 1234-246.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own?. *Cognition*, *88*(3), B11-B21.

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of cognitive neuroscience*, *18*(5), 859-870.

Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., & Frith, C. (2007). Is it really my turn? An event-related fMRI study of task sharing. *Social Neuroscience , 2* (2), 81-95.

Sebanz, N., Knoblich, G., Stumpf, L., & Prinz, W. (2005). Far from action-blind: Representation of others' actions in individuals with autism. *Cognitive Neuropsychology*, *22*(3-4), 433-454.

Sellaro, R., Treccani, B., Rubichi, S., & Cubelli, R. (2013). When co-action eliminates the simon effect: Disentangling the impact of co-actor's presence and task sharing on joint-task performance. *Frontiers in Psychology , 4* (NOV).

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology*, *38*(3), 475-491.

Simon, J. R. 1990. “The effects of an irrelevant directional cue on human information processing”. In *Stimulus-response compatibility: An integrated perspective*, Edited by: Proctor, R. W. and Reeve, T. G. 31–86. Amsterdam: North-Holland

Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B., Agid, Y., & Pierrot-Deseilligny, C. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport*, *6*(7), 997-1001.

Slagter, H. A., Prinssen, S., Reteig, L. C., & Mazaheri, A. (2016). Facilitation and inhibition in attention: functional dissociation of pre-stimulus alpha activity, P1, and N1 components. *Neuroimage*, *125*, 25-35.

Slamecka, N. J., & Katsaiti, L. T. (1987). The generation effect as an artifact of selective displaced rehearsal. *Journal of Memory and Language*, *26*(6), 589-607.

Smyth, M. M., Pearson, N. A., & Pendleton, L. R. (1988). Movement and working memory: Patterns and positions in space. *The Quarterly Journal of Experimental Psychology*, *40*(3), 497-514.

Smyth, M. M., & Pendleton, L. R. (1989). Working memory for movements. *The quarterly Journal of experimental Psychology*, *41*(2), 235-250.

Smyth, M., & Scholey, K. (1994). *Characteristics of Spatial Memory Span: Is There an Analogy to the Word Length Effect, Based on Movement Time?*

Smyth, M. M., & Pelky, P. L. (1992). Short-term retention of spatial information. *British Journal of Psychology*, *83*(3), 359-374.

Spataro, P., Marques, V., Longobardi, E., & Rossi-Arnaud, C. (2015). Does pointing facilitate the recall of serial positions in visuospatial working memory? *Cognitive Processing , 16*, 377-381.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, *38*(4), 387-401.

Stanley, J., Gowen, E., & Miall, R. C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(4), 915.

Steffens, M. C. (1999). The role of relational processing in memory for actions: A negative enactment effect in free recall. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(4), 877-903.

Steffens, M., Von Stülpnagel, R., & Schult, J. (2015). Memory recall after "learning by doing" and "learning by viewing": Boundary conditions of an enactment benefit. *Frontiers in Psychology , 6* (DEC).

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1746.

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human perception and performance*, *38*(5), 1105.

Tai, Y. F., Scherfler, C., Brooks, D. J., Sawamoto, N., & Castiello, U. (2004). The human premotor cortex is' mirror'only for biological actions. *Current biology*, *14*(2), 117-120.

Tecce, J. J. (1972). Contingent negative variation (CNV) and psychological processes in man. *Psychological bulletin*, *77*(2), 73.

Theeuwes, J., Belopolsky, A., & Olivers, C. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica , 132* (2), 106-114.

Thomas, S. J., Gonsalvez, C. J., & Johnstone, S. J. (2009). Sequence effects in the Go/NoGo task: Inhibition and facilitation. *International Journal of Psychophysiology*, *74*(3), 209-219.

Toni, I., Gentilucci, M., Jeannerod, M., & Decety, J. (1996). Differential influence of the visual framework on end point accuracy and trajectory specification of arm movements. *Experimental Brain Research*, *111*(3), 447-454.

Touzel, M., Snidal, C., Segal, J., Renoult, L., & Debruille, J. B. (2018). A central component of the N1 event-related brain potential could index the early and automatic inhibition of the actions systematically activated by objects. *bioRxiv*, 341057.

Tsai, C. C., & Brass, M. (2007). Does the human motor system simulate Pinocchio's actions? Coacting with a human hand versus a wooden hand in a dyadic interaction. *Psychological Science*, *18*(12), 1058-1062.

Tsai, J. C. C., Sebanz, N., & Knoblich, G. (2011). The GROOP effect: Groups mimic group actions. *Cognition*, *118*(1), 135-140.

Tsai, C., Kuo, W., Jing, J., Hung, D., & Tzeng, O. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research , 175* (2), 353-362.

Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience*, *20*(11), 2015-2024.

Turk, D. J., Cunningham, S. J., & Macrae, C. N. (2008). Self-memory biases in explicit and incidental encoding of trait adjectives. *Consciousness and cognition*, *17*(3), 1040-1045.

Ulrich, R., Leuthold, H., & Sommer, W. (1998). Motor programming of response force and movement direction. *Psychophysiology*, *35*(6), 721-728.

Van Boxtel, G. J. M., & Brunia, C. H. (1994). Motor and non-motor aspects of slow brain potentials. *Biological Psychology*, *38*(1), 37-51.

Van Boxtel, G. J., & Brunia, C. H. M. (1994). Motor and non-motor components of the contingent negative variation. *International Journal of Psychophysiology*, *17*(3), 269-279.

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British journal of psychology*, *95*(1), 57-79.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*(8-9), 998-1003.

Vlainic, E., Liepelt, R., Colzato, L., Prinz, W., & Hommel, B. (2010). The virtual co-actor: The social Simon effect does not rely on online feedback from the other. *Frontiers in Psychology , 1* (DEC).

Vogel, E. K., & Luck, S. J. (2002). Delayed working memory consolidation during the attentional blink. *Psychonomic Bulletin & Review*, *9*(4), 739-743.

Vogel, E. K., Luck, S. J., & Shapiro, K. L. (1998). Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(6), 1656.

von Stülpnagel, R., & Steffens, M. C. (2013). Active route learning in virtual environments: disentangling movement control from intention, instruction specificity, and navigation control. *Psychological research*, *77*(5), 555-574.

Von Stülpnagel, R., & Steffens, M. C. (2012). Can active navigation be as good as driving? A comparison of spatial memory in drivers and backseat drivers. *Journal of experimental psychology: applied*, *18*(2), 162.

Welsh, T. N., Higgins, L., Ray, M., & Weeks, D. J. (2007). Seeing vs. believing: Is believing sufficient to activate the processes of response co-representation?. *Human Movement Science*, *26*(6), 853-866.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *nature*, *203*(4943), 380-384.

Weinberg, A., & Hajcak, G. (2011). The late positive potential predicts subsequent interference with target processing. *Journal of cognitive neuroscience*, *23*(10), 2994-3007.

Welsh, T., Elliott, D., Anson, J., Dhillon, V., Weeks, D., Lyons, J., et al. (2005). Does Joe influence Fred's action? Inhibition of return across different nervous systems. *Neuroscience Letters , 385* (2), 99-104.

Wild-Wall, N., Jo¨, J., Sangals, J., Sommer, W., & Leuthold, H. (2003). *Are fingers special? Evidence about movement preparation from event-related brain potentials.* Blackwell Publishing Inc.

Wronka, E. A., Kaiser, J., & Coenen, A. M. (2012). Neural generators of the auditory evoked potential components P3a and P3b.

Yantis, S., Schwarzbach, J., Serences, J. T., Carlson, R. L., Steinmetz, M. A., Pekar, J. J., & Courtney, S. M. (2002). Transient neural activity in human parietal cortex during spatial attention shifts. *Nature neuroscience*, *5*(10), 995.

Zaehle, T., Jordan, K., Wüstenberg, T., Baudewig, J., Dechent, P., & Mast, F. W. (2007). The neural basis of the egocentric and allocentric spatial frame of reference. *Brain research*, *1137*, 92-103.

Zimmer, H. D., & Engelkamp, J. (1985). An attempt to distinguish between kinematic and motor memory components. *Acta psychologica*, *58*(1), 81-106.

Zimmer, H. D., & Engelkamp, J. (1984). Planungs-und Ausführungsanteile motorischer Gedächtniskomponenten und ihre Wirkung auf das Behalten ihrer verbalen Bezeichnungen. *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie*.

Appendix I

Full Analyses of Experiment 1 and Experiment 2 from the eye tracking study

Full Analysis of Experiment 1:

Recognition accuracy (%). A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 7.97$, $p = 0.01$, $\eta_p^2 = 0.29$], indicating that recognition accuracy was higher for move arrays ($M = 91.14\%$) than for no-move arrays ($M = 76.76\%$), and a significant main effect of Array Order [$F(1, 19) = 20.10$, $p < 0.001$, $\eta_p^2 = 0.51$], indicating that recognition accuracy was higher for arrays presented as second ($M = 89.79\%$) than for arrays presented as first ($M = 78.12\%$). These effects were qualified by a significant three-way interaction among Array Order, Condition and Agent Cue [$F(1, 19) = 15.74$, $p = 0.001$, $\eta_p2 = 0.45$]. However, a follow-up analysis showed that the effect of Condition was significant both when the pointing movements were performed by the participant (P-cued) [$F(1, 19) = 4.42$, $p = 0.049$, $\eta_p2 = 0.19$] and when they were performed by the experimenter (E-cued) [$F(1, 19) = 11.31$, $p = 0.003$, $\eta_p2 = 0.37$].

Relative fixation percentages for the encoded arrays. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA was performed on number of fixations during encoding phase which revealed: a significant main effect of Condition [$F(1, 19) = 16.46$, $p = 0.001$, $\eta_p^2 = 0.46$], indicating that relative fixation percentages were higher for move ($M = 11.69\%$) than for no-move arrays ($M = 10.84\%$); a significant main effect of Array Order [$F(1,$

19) = 4.22, $p = 0.05$, $\eta_p^2 = 0.18$], indicating that relative fixation percentages were higher for arrays presented as second ($M = 11.53\%$) than for arrays presented as first ($M = 11.01\%$); a significant main effect of Array Size [$F(1, 19) = 11.36$, $p = 0.003$, $\eta_p^2 = 0.37$], indicating that relative fixation percentages were higher for 4-item ($M = 11.71\%$) than for 3-item arrays ($M = 10.82\%$). These effects were qualified by a four-way interaction between Array Order, Condition, Array Size and Agent Cue [$F(1, 19) = 5.31$, $p = 0.03$, $\eta_p^2 = 0.21$]. A follow-up analysis focusing on the effects of Condition indicated that the advantage of move arrays was significant in all cases, except for 3-item arrays presented as first – both when the move arrays were pointed by the participant, $F(1, 19) = 2.91$, $p = 0.10$, and when they were pointed by the experimenter, $F(1, 19) = 1.13$, $p = 0.30$.

Relative fixation percentages for the blank screen. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 10.63$, $p = 0.004$, $\eta_p^2 = 0.36$], indicating that relative fixation percentages were higher for move ($M = 7.67\%$) than for no-move arrays ($M = 6.47\%$); a significant main effect of Array Order [$F(1, 19) = 5.84$, $p = 0.02$, $\eta_p^2 = 0.23$], indicating that relative fixation percentages were higher for arrays presented as first ($M = 7.54\%$) than for arrays presented as second ($M = 6.61\%$); a significant main effect of Array Size [$F(1, 19) = 96.55$, $p < 0.001$, $\eta_p^2 = 0.83$], indicating that relative fixation percentages were higher for 3-item ($M = 8.63\%$) than for 4-item arrays ($M = 5.51\%$). These effects were qualified by a significant interaction between Array Size and Condition [$F(1, 19) = 4.84$, $p = 0.04$, $\eta_p2= 0.20$]: a follow-up analysis indicated the advantage of move arrays was significant for 3-item arrays [$F(1, 19) = 13.65$, $p = 0.002$, $\eta_p2 = 0.42$], but not for 4-item arrays [$F(1, 19) = 2.49$, $p = 0.13$, $\eta_p2 = 0.12$].

Relative gaze duration for the encoded arrays. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA revealed a significant main effect of Condition [$F(1, 19) = 22.88$, $p < 0.001$, $\eta_p^2 = 0.55$], indicating that relative gaze duration was higher for move ($M = 12.18\%$) than for no-move arrays ($M = 9.92\%$); a significant main effect of Array Size [$F(1, 19) = 16.20$, $p = 0.001$, $\eta_p^2 = 0.46$], indicating that relative gaze duration was higher for 4-item ($M = 11.66\%$) than for 3-item arrays ($M = 10.44\%$); a significant main effect of Agent Cue [$F(1, 19) = 8.33$, $p = 0.009$, $\eta_p^2 = 0.30$], indicating that relative gaze duration was higher for participant ($M = 11.54\%$) than for experimenter-pointed arrays ($M = 10.56\%$). These effects were qualified by a significant two-way interaction between Condition and Array Size [$F(1, 19) = 7.04$, $p = 0.01$, $\eta_p 2 = 0.27$] and by a three-way interaction between Array Order, Condition and Agent Cue [$F(1, 19) = 11.41$, $p = 0.003$, $\eta_p 2 = 0.37$]. However, for both interactions, follow-up analyses focusing on the effects of Condition showed that the advantage of move arrays over no-move arrays was significant in all conditions [$F(1, 19) > 12.42$, $p < 0.002$, $\eta_p 2 > 0.39$ and $F(1, 19) > 14.68$, $p < 0.001$, $\eta_p 2 > 0.43$].

Relative gaze duration for the blank screen. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) × 2 (Agent Cue: P-cued vs. E-cued) repeated measures ANOVA was performed on average gaze duration during retention phase which revealed a significant main effect of Condition [$F(1, 19) = 13.81$, $p = 0.001$, $\eta_p^2 = 0.42$], indicating that relative gaze duration was higher for move ($M = 8.46\%$) than for no-move arrays ($M = 6.72\%$); a significant main effect of Array Order [$F(1, 19) = 12.21$, $p = 0.002$, $\eta_p^2 = 0.39$], indicating that relative gaze duration was higher for arrays presented as first ($M = 8.36\%$) than for arrays presented as second ($M = 6.82\%$); a significant main effect of Array

Size [$F(1, 19) = 60.73$, $p < 0.001$, $\eta_p^2 = 0.76$], indicating that average gaze duration was higher for the 3-item ($M = 9.17\%$) than for 4-item arrays ($M = 6.01\%$). These effects were qualified by a marginal two-way interaction between Condition and Array Size [$F(1, 19) = 3.78$, $p = 0.06$, $\eta_p2 = 0.16$]. However, a follow-up analysis focusing on the effects of condition showed that the advantage of move arrays over no-move arrays was significant for both 3- and 4-item arrays [$F(1, 19) = 16.60$, $p = 0.001$, $\eta_p2 = 0.46$ and $F(1, 19) = 4.98$, $p = 0.038$, $\eta_p2 = 0.21$].

Full Analysis of Experiment 2:

Recognition accuracy (%). A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) repeated measures ANOVA revealed a significant main effect of Array Order [$F(1, 19) = 9.50$, $p = 0.006$, $\eta_p^2 = 0.33$], indicating that recognition performance was higher for arrays presented as second ($M = 83.04\ \%$) than for arrays presented as first ($M = 69.66\ \%$), and a significant main effect of Array Size [$F(1, 19) = 8.16$, $p = 0.01$, $\eta_p^2 = 0.30$], indicating that recognition performance was higher for 3-item ($M = 79.33\%$) than for 4-item arrays ($M = 73.37\%$). Crucially, the main effect of Condition was not significant [$F(1, 19) = 0.25$, $p = 0.62$, $\eta_p^2 = 0.01$], suggesting that move and no-move arrays were recognized with equal accuracy [M = 76.91% and M = 75.78% for move and no-move arrays respectively]. In addition, the analysis revealed a significant interaction between Array Order and Array Size [$F(1, 19) = 4.31$, $p = 0.05$, $\eta_p^2 = 0.18$]: a follow-up analysis of simple effects showed that 3-item arrays were recognized better than 4-item arrays when they were presented as first [M = 74.99% vs. M = 64.33%: $F(1, 19) = 12.05$, $p = 0.003$, $\eta_p^2 = 0.38$], but not when they were presented as second [M = 83.66% vs. M = 82.41%: $F(1, 19) = 0.16$, $p = 0.69$, $\eta_p^2 = 0.009$]. No other effects or interactions were significant [$F(1,19) < 0.75$, $p > 0.396$].

Relative fixation percentages for the encoded arrays. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) repeated measures ANOVA revealed a significant main effect of Array Size [$F(1, 19) = 9.84$, $p = 0.005$, $\eta_p^2 = 0.34$], indicating that fixation percentages were higher for 4-item ($M = 10.82\%$) than for 3-item arrays ($M = 9.68\%$), and a marginal main effect of Array Order [$F(1, 19) = 3.68$, $p = 0.07$, $\eta_p^2 = 0.16$], indicating that fixation percentages tended to be higher for arrays presented as first ($M = 10.60\%$) than for arrays presented as second ($M = 9.91\%$). The main effect of Condition was not significant [$F(1, 19) = 0.42$, $p = 0.52$, $\eta_p^2 = 0.02$]. Finally, a significant interaction between Array Order and Array Size was also found [$F(1, 19) = 87.54$, $p < 0.001$, $\eta_p^2 = 0.82$]: a follow-up analysis of simple effects showed that fixation percentages were higher for 4-item than for 3-item arrays, when they were presented as first [$F(1, 19) = 41.05$, $p < 0.001$, $\eta_p^2 = 0.68$], but not when they were presented as second [$F(1, 19) = 1.51$, $p = 0.234$, $\eta_p^2 = 0.07$]. No other effects or interactions reached the significance level [$F(1,19) < 0.42$, $p > 0.523$].

Relative fixation percentages for the blank screen. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) repeated measures ANOVA revealed a significant main effect of Array Order [$F(1, 19) = 18.07$, $p < 0.001$, $\eta_p^2 = 0.48$], indicating that fixation percentages were higher for arrays presented as first ($M = 8.28\%$) than for arrays presented as second ($M = 6.55\%$), and a significant main effect of Array Size [$F(1, 19) = 81.84$, $p < 0.001$, $\eta_p^2 = 0.81$], indicating that fixation percentages were higher for 3-item ($M = 9.01\%$) than for 4-item arrays ($M = 5.82\%$). The main effect of Condition was not significant [$F(1, 19) = 0.06$, $p = 0.80$, $\eta_p^2 = 0.003$], as they were all other effects and interactions [$F(1, 19) < 0.43$, $p > 0.516$].

Relative gaze duration for the encoded arrays. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) repeated measures ANOVA revealed a significant main effect of Array Size [$F(1, 19) = 13.71$, $p = 0.002$, $\eta_p^2 = 0.41$], indicating that relative gaze duration was longer for 4-item ($M = 11.33\%$) than for 3-item arrays ($M = 9.79\%$), and a marginal main effect of Array Order [$F(1, 19) = 3.42$, $p = 0.08$, $\eta_p^2 = 0.15$], indicating that average gaze duration tended to be longer for arrays presented as first ($M = 10.98\%$) than for arrays presented as second ($M = 10.14\%$). The main effect of Condition was not significant [$F(1, 19) = 1.11$, $p = 0.30$, $\eta_p^2 = 0.05$], confirming that relative gaze duration was the same for move and no-move arrays. In addition, a significant interaction between Array Order and Array Size was also found [$F(1, 19) = 92.82$, $p < 0.001$, $\eta_p^2 = 0.83$]: a follow-up analysis of simple effects showed that relative gaze duration was longer for 4-item than for 3-item arrays when they were presented as first [$F(1, 19) = 57.27$, $p < 0.001$, $\eta_p^2 = 0.75$], but not when they were presented as second [$F(1, 19) = 1.40$, $p = 0.251$, $\eta_p^2 = 0.06$]. No other effects or interactions were found significant [$F(1, 19) < 1.11$, $p > 0.305$].

Relative gaze durations for the blank screen. A 2 (Array Order: first vs. second array) × 2 (Array Size: 3- vs. 4-item array) × 2 (Condition: move vs. no-move array) repeated measures ANOVA revealed a significant main effect of Array Order [$F(1, 19) = 20.95$, $p < 0.001$, $\eta_p^2 = 0.52$], indicating that relative gaze duration was longer for arrays presented as first ($M = 9.04\%$) than for arrays presented as second ($M = 6.85\%$), and a significant main effect of Array Size [$F(1, 19) = 57.19$, $p < 0.001$, $\eta_p^2 = 0.75$], indicating that relative gaze duration was longer for 3-item ($M = 9.62\%$) than for 4-item arrays ($M = 6.27\%$). The main effect of Condition was not significant [$F(1, 19) = 0.01$, $p = 0.91$, $\eta_p^2 = 0.001$], as they were all other effects or interactions [$F(1, 19) < 0.94$, $p > 0.344$].

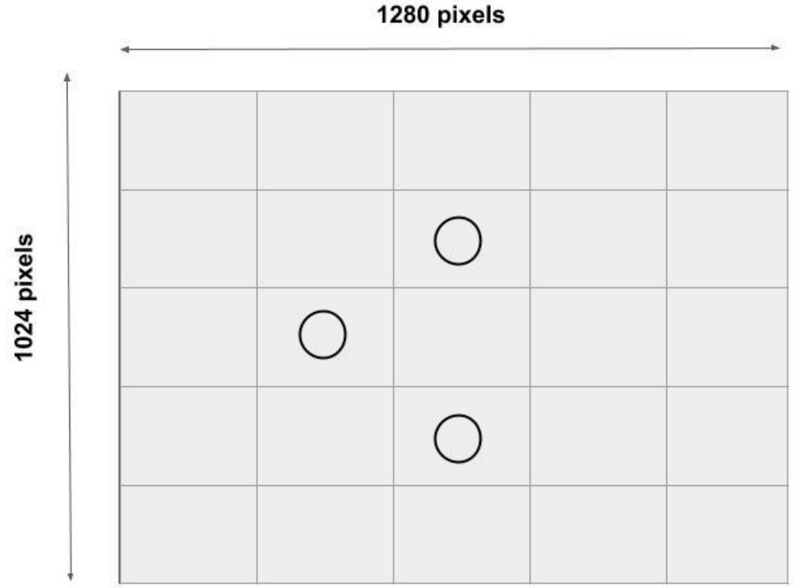


Figure 1: A schematic figure of the regions of interest for the analysis of fixations and gaze-duration.

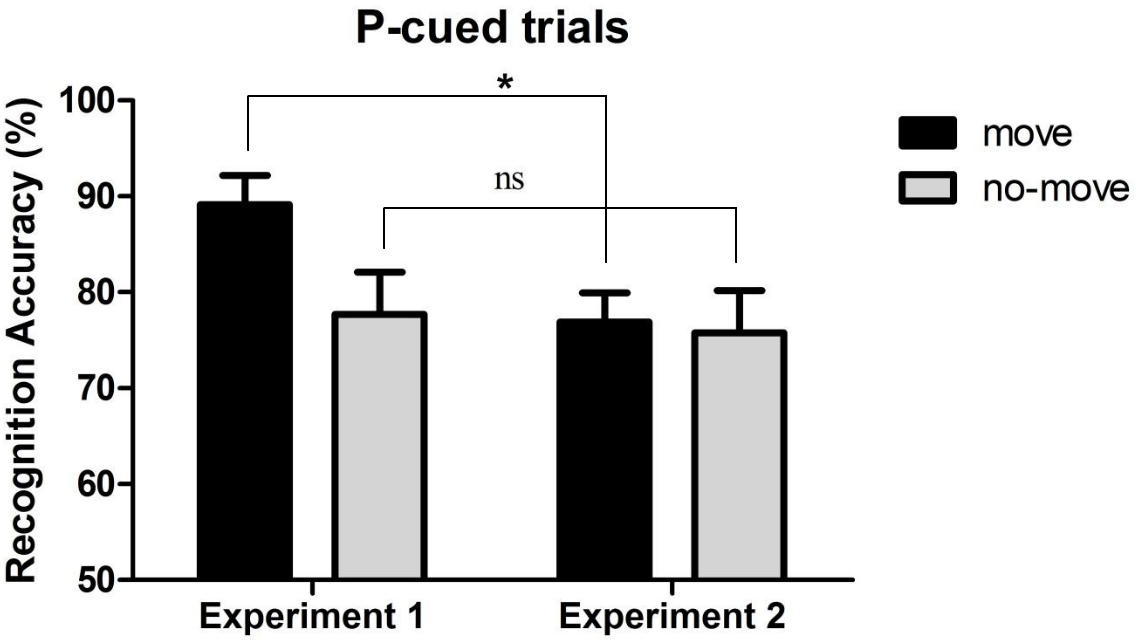
Appendix II

Cross experimental analyses of Experiment 1 and Experiment 2 from the eye-tracking study

To determine whether the advantage in the recognition of the arrays pointed by both the participants and the experimenter in the joint-action condition was due to the increase in the recognition of move arrays or to the inhibition in the recognition of no-move arrays, in study 1 we performed a cross-experiment comparison with the performance observed in the passive condition. The assumption was that, given the absence of a significant memory difference between move and no-move arrays, the performance in the passive condition could be considered as an appropriate baseline against which the decrements produced by motor inhibition and/or the improvements produced by self-pointing movements in the joint-action condition could be verified. In study 1, we found that the recognition of move arrays did not differ between the two conditions (suggesting that self-pointing movements did not increase spatial memory), whereas the recognition of no-move arrays was significantly lower in the joint action than in the passive condition (suggesting that the advantage in the recognition of move arrays could be entirely ascribed to the negative effect of motor inhibition on the encoding of no-move arrays). To ascertain whether the same pattern emerged in the eye-tracking study, we performed a cross-experiment comparison between the recognition data obtained in Experiments 1 and 2. Separate analyses were run for P- and E-cued trials of Experiment 1. For P-cued trials, a 2 (Condition: move vs. no-move array) × 2 (Experiment: Exp.1 vs. Exp.2) mixed ANOVA found a significant main effect of Condition [F(1, 38) = 4.56, *p* = 0.039, ηp2 = 0.11] and a marginally significant two-way interaction between Condition and Experiment [F(1, 38) = 3.07, *p* = 0.088, ηp2 = 0.07]. As illustrated in the top panel of **Figure 2**, the recognition of no-move arrays was not lower (instead was numerically higher) in the joint-action condition of Experiment 1 than in the passive condition of Experiment 2 [M = 77.70% vs. M = 75.78%: F(1, 38) = 0.09, p = 0.76, ηp2 = 0.002], whereas the recognition of move arrays showed the opposite pattern – it was significantly higher in the joint-action condition of Experiment 1 than in the passive condition of Experiment 2 [M = 89.16% vs. M = 76.91%: F(1, 38) = 8.17, p = 0.007, ηp2 = 0.17].

Regarding E-cued trials, we found that the main effect of Condition and the two-way interaction between Condition and Experiment were significant [F(1, 38) = 10.77, p = 0.002, ηp2 = 0.22 and F(1, 38) = 8.30, p = 0.006, ηp2 = 0.18, respectively]: as illustrated in the bottom panel of **Figure 2**, the recognition of no-move arrays was similar in the joint-action condition of

Experiment 1 and in the passive condition of Experiment 2 [M = 75.83% vs. M = 75.79%: F(1, 38) = .00, $p = 0.99$, ηp2 = 0.00], whereas the recognition of move arrays was significantly higher in the joint-action condition of Experiment 1 than in the passive condition of Experiment 2 [M = 93.12% vs. M = 76.91%: F(1, 38) = 14.69, $p < 0.001$, ηp2 = 0.28].


P-cued trials
Recognition Accuracy (%)
100
90
80
70
60
50
*
ns
Experiment 1
Experiment 2
move
no-move

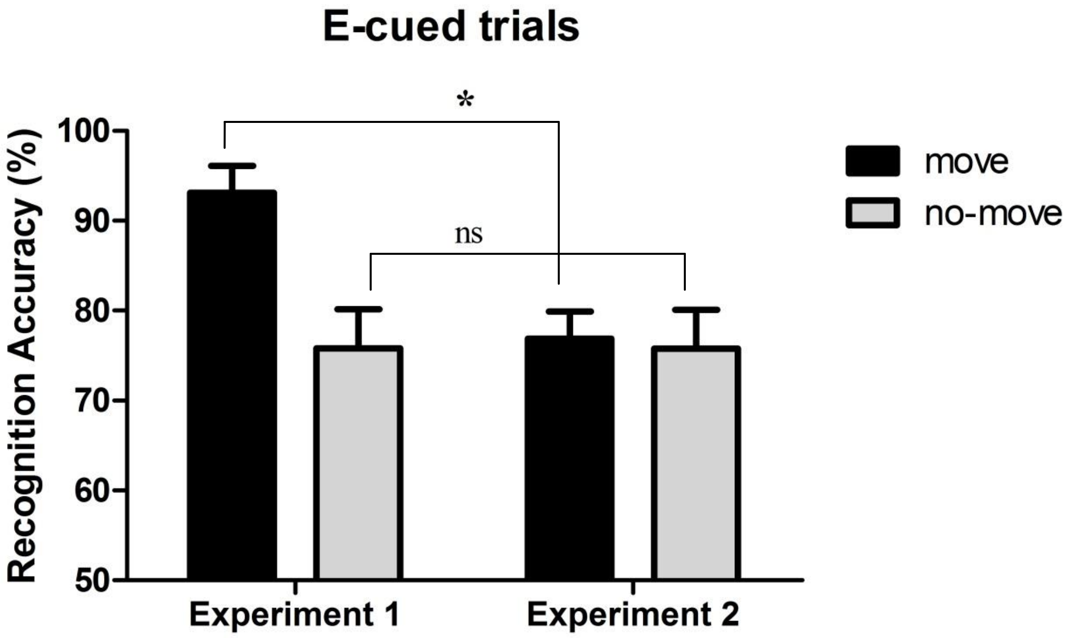

E-cued trials
Recognition Accuracy (%)
100
90
80
70
60
50
*
ns
Experiment 1
Experiment 2
move
no-move

Figure 2: Cross experimental analyses of memory performance for move and no-move arrays between Experiment 1 and Experiment 2. * represents the significant difference between the two conditions, independently of the significance level. ns represents not significant difference between the two conditions

Appendix III

Higher order interactions from the event related potential study

A significant three-way interaction between Order, Condition and Pointing hand [$F(1, 21) = 11.14$, $p < 0.01$, $\eta_p^2 = 0.34$] was found. A follow-up analysis focusing on the effect of Condition showed that move arrays were recognized better than the no-move arrays in both hand conditions (for the Right hand pointing: [$F(1, 21) = 15.18$, $p = 0.001$, $\eta_p^2 = 0.34$] and for the Left hand pointing: [$F(1, 21) = 10.14$, $p < 0.01$, $\eta_p^2 = 0.32$]) but only when the items were presented at order first.

A significant higher order interaction between Order, Condition, Pointing agent and Pointing hand was observed [$F(1, 21) = 14.19$, $p = 0.001$, $\eta_p^2 = 0.40$]. The analysis of simple effects focusing on the Condition showed:

(a) At Order first, a three way interaction between Condition, Pointing agent and Pointing hand was found [$F(1, 21) = 5.59$, $p < 0.05$, $\eta_p^2 = 0.21$]. Follow-up analysis showed that in the self-pointing trials, the move arrays were remembered significantly better than the no-move arrays irrespective of the pointing hand. For the right hand pointing condition accuracy for move (M = 84.77%) was higher than no-move (M = 60.40%) [$F(1, 21) = 26.30$, $p < 0.001$, $\eta_p^2 = 0.55$], so was true for the left hand pointing condition (move M = 78.09%, no-move M = 67.95%) [$F(1, 21) = 6.17$, $p < 0.05$, $\eta_p^2 = 0.22$]. In the confederate-pointing trials, for the right hand pointing condition, accuracy for move arrays (M = 67.90%) was numerically higher than no-move arrays (M = 60.45%) but could not reach significance [$F(1, 21) = 1.87$, $p > 0.05$, $\eta_p^2 = 0.08$]. For the left hand pointing, move arrays were recognized significantly higher (M = 69.54%) than no-move arrays (M = 56.50%) [$F(1, 21) = 6.17$, $p < 0.05$, $\eta_p^2 = 0.22$].

(b) At Order second, an interaction between Condition and Pointing hand was found [$F(1, 21) = 9.97$, $p < 0.01$, $\eta_p^2 = 0.32$]. Follow-up analysis showed that move arrays (M = 84.56%) were recognized significantly better than no-move arrays (M = 76.90%) when pointing was performed by left hand [$F(1, 21) = 4.21$, $p = 0.05$, $\eta_p^2 = 0.16$]. However, an opposite pattern was found when the pointing was performed by right hand, no-move arrays (M = 86.29%) were recognized better than move arrays (M= 79.54%) and the difference was marginally significant [$F(1, 21) = 3.97$, $p = 0.06$, $\eta_p^2 = 0.16$]. An interaction between Pointing agent and hand was found [$F(1, 21) = 6.32$, $p < 0.05$, $\eta_p^2 = 0.23$].

(c) For the right hand pointing condition, a significant three way interaction was found between Order, Condition and Pointing agent [$F(1, 21) = 13.99$, $p = 0.001$, $\eta_p^2 = 0.40$]. Follow-up analysis showed that at order first, move arrays were recognized significantly better (M = 84.77%) than no-move arrays (M = 60.40%) in the self-pointing trials, $F(1, 21) = 26.30$, $p < 0.001$, $\eta_p^2 = 0.55$]. No significant interaction was found in the left hand pointing condition.

Supplementary Material

Page No.	Type of work:	Reprinted work	Source of work	Copyright holder	permission granted	permission date	note
7	figure	Figure 1.2: The results shown by Chum et al. (2007, p. 1191) for Experiment 1...	Motor and visual codes interact to facilitate visuospatial memory performance," by Chum, M., Bekkering, H., Dodd, M., & Pratt, J. 2007, *Psychonomic Bulletin and Review*, 14 (6), p. 1191	© 2007 Springer Nature	yes	21/12/2019	Written permission
9	figure	Figure 1.3: Results shown by Dodd and Shumborski (2009, p. 1240) in Experiment 1 and 2...	Examining the influence of action on spatial working memory: The importance of selection. By Dodd, M., & Shumborski, S. 2009, *Quarterly Journal of Experimental Psychology , 62* (6), p. 1240	© Taylor & Francis	yes	06/01/2020	Written permission
17	figure	Figure 1.4: Sitting arrangements in joint and individual go-nogo task conditions and the results reported in the study by Sebanz, Knoblich..	Representing others' actions: just like one's own?," by Sebanz, N., Knoblich, G., & Prinz, W. , 2003, *Cognition*, *88*(3), p. 14, 16	© Elsevier	yes	21/12/2019	Written permission:
20	figure	Figure 1.5: LRP results shown by Tsai et al. (2008, p. 2021)…	Action co-representation is tuned to other humans," by	© Massachusetts Institute of Technology	yes	23/12/2019	Written permission

			Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J., 2008, *Journal of Cognitive Neuroscience*, *20* (11), p. 2021				
22	figure	Figure 1.6: The effect of social context between the task-partners illustrated	Favouritism in the motor system: Social interaction modulates action simulation," by Kourtis, D., Sebanz, N., & Knoblich, G., 2010, *Biology Letters* , *6* (6), p. 759	© The Royal Society	yes	21/12/2019	Written permission
25	figure	Figure 1.7: Attention-attracting object used by Dolk et al. (2013)…	The (not so) social Simon effect: a referential coding account," by Dolk, T., Hommel, B., Prinz, W., & Liepelt, R., 2013, *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), p. 1251	© American Psychological Association	yes	21/12/2019	Written permission