

**DIFFERENCES IN VISUAL CAPTURE OF CONNECTED BODY PARTS IN
THE MIRROR-BOX ILLUSION REFLECT LOCAL VARIATIONS IN
CROSS-MODAL CONGRUENCE AND UNIMODAL VARIANCE**

by

Pushpita Bhattacharyya

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Psychological and Brain Sciences

Summer 2023

© 2023 Pushpita Bhattacharyya
All Rights Reserved

**DIFFERENCES IN VISUAL CAPTURE OF CONNECTED BODY PARTS IN
THE MIRROR-BOX ILLUSION REFLECT LOCAL VARIATIONS IN
CROSS-MODAL CONGRUENCE AND UNIMODAL VARIANCE**

by

Pushpita Bhattacharyya

Approved: _____
Jared Medina, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Tania Roth, Ph.D.
Chair of the Department of Psychological and Brain Sciences

Approved: _____
John A. Pelesko, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

DEDICATION

In memory of my grandmother Chitra Bhattacharya – my very first guide into the worlds of critical thinking and storytelling, who never missed an opportunity to nurture and celebrate my mind.

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude toward my advisor Dr. Jared Medina, for providing essential hands-on guidance at every stage of this project, from conception to final write-up. Thank you for supporting me throughout this journey, and pushing me to discover new heights of my potential as a scientist. I would also like to thank my committee members Dr. Tim Vickery and Dr. Keith Schneider for valuable feedback on this thesis.

I would also like to thank my lab members Anu Nair, Luisa Raigosa-Posada, Riwa Safa, Elisabetta Ambron, and Carli Fine for providing valuable insight at various stages of the project, which helped refine the design of these experiments and develop a nuanced interpretation of their results. I would also like to thank my undergraduate assistants Audrey Berry, Kallie Sweetman, Christopher Rochester and Rebecca Metellus for their dedicated time, effort and attention to detail in collecting the data reported in this study.

Next, I would like to thank the constant support of my ‘found family’ here in Newark: Banjit Singh, Alyssa Levy, Josiane Mukahirwa, Lynn Baker, Anton and Olga Lebed, and the rest of my friends in the PBS, as well as my dear friends from the Queer and Trans Graduate Student Union – Cameron, Willow, Casey, Ash, Shar, Maximilien, Celine, and the rest – who have provided me with a strong sense of belonging in the greater graduate student community here at the University of Delaware. Special thanks to Anthony Sigman-Lowery, who drove me home on cold, winter nights after long hours spent analyzing data in the office.

I am eternally grateful to my family for their unconditional support. A special shout-out to my uncle Amartya Samadder and my aunt Sarbani Samadder, for providing a warm home and cooked meals throughout the winter while I worked day and night on this thesis, and my cousins Angan and Arin for keeping my spirits up through the holidays despite my intense work schedule. I would also like to thank my little brother Shovit, for his steady supply of humor and patience during my initial attempts at verbalizing my research ideas in clear, comprehensible terms. Last but not the least – none of this would have been possible without the relentless support from my mother Nayana Bhattacharya, and my father Debdatta Bhattacharyya, whose unshakeable belief in me kept me going through the toughest challenges throughout my journey as an aspiring scientist.

TABLE OF CONTENTS

| | |
|---|----|
| LIST OF FIGURES | ix |
| ABSTRACT | xi |
| Chapter | |
| 1 INTRODUCTION | 1 |
| 1.1 How does the brain solve the unique challenges in multisensory integration? | 2 |
| 1.2 How do these multisensory integration principles inform one’s perception of their body? | 4 |
| 1.3 How are locally integrated position estimates for different body parts integrated into a unitary ‘whole-body’ percept? | 7 |
| 1.4 How do bottom-up factors like cross-modal congruence and unimodal variance interact with top-down information from a stored body representation to generate integrated estimates of the body position? | 13 |
| 1.4.1 Experiment 1 – the location of synchronous multisensory stimulation | 14 |
| 1.4.2 Experiment 2 – the location of the visual fixation point | 16 |
| 1.5 How do local and global effects of visual capture vary over the time-course of the mirror-box illusion? | 18 |
| 2 EXPERIMENT 1 – THE EFFECT OF THE LOCATION OF SYNCHRONOUS MULTISENSORY STIMULATION ON HAND AND FOREARM VISUAL CAPTURE IN THE MBI | 22 |
| 2.1 Materials and Methods | 22 |
| 2.1.1 Participants | 22 |
| 2.1.2 Apparatus | 23 |
| 2.1.3 Experimental Paradigm | 23 |
| 2.1.4 Procedure | 25 |
| 2.1.5 Data Analysis | 26 |
| 2.2 Results | 32 |

| | | |
|---------|---|----|
| 2.2.1 | Within-question analyses – The effect of stroking location on visual capture of the hand and the forearm, respectively | 33 |
| 2.2.1.1 | ‘Initial phase’ of stimulation – 0-10 seconds | 33 |
| 2.2.1.2 | ‘Subsequent phase’ of stimulation – 10-60 seconds..... | 35 |
| 2.2.2 | Between-question analyses – comparing visual capture of the hand versus the forearm under each stroking condition | 36 |
| 2.2.2.1 | The difference between hand and forearm visual capture in each stroking condition, after 60 seconds of stroking..... | 36 |
| 2.2.2.2 | How the location of synchronous stroking drives ‘fragmented visual capture’ of a connected hand and forearm in the mirror | 38 |
| 2.2.3 | Top-down effects of a stored ‘global’ body representation on visual capture | 40 |
| 2.2.3.1 | Likelihood of visual capture of the unstimulated forearm, given visual capture of the stimulated hand... 40 | |
| 2.2.3.2 | Likelihood of visual capture of the unstimulated hand, given visual capture of the stimulated forearm... 41 | |
| 2.3 | Discussion..... | 42 |
| 3 | THE EFFECT OF THE LOCATION OF VISUAL FIXATION ON HAND AND FOREARM VISUAL CAPTURE IN THE MBI..... | 47 |
| 3.1 | Materials and Methods | 47 |
| 3.1.1 | Participants | 47 |
| 3.1.2 | Apparatus..... | 48 |
| 3.1.3 | Experimental Paradigm | 48 |
| 3.1.4 | Procedure..... | 49 |
| 3.1.5 | Data Analysis: | 50 |
| 3.2 | Results | 53 |
| 3.2.1 | Within-question analyses – The effect of the location of fixation on visual capture of the hand and the forearm, respectively | 54 |
| 3.2.1.1 | ‘Initial phase’ of stimulation – 0-10 seconds | 54 |

| | | |
|----------|--|----|
| 3.2.1.2 | ‘Subsequent phase’ of stimulation – 10-60 seconds..... | 55 |
| 3.2.2 | Between-question analyses – comparing visual capture of the hand versus the forearm under each fixation condition..... | 56 |
| 3.2.2.1 | The difference between hand and forearm visual capture in each fixation condition, at t = 60 seconds. .. | 57 |
| 3.2.2.2 | How the location of visual fixation drives ‘fragmented visual capture’ of a connected hand and forearm in the mirror | 58 |
| 3.2.3 | Top-down effects of a stored ‘global’ body representation on visual capture..... | 60 |
| 3.2.3.1 | Likelihood of visual capture of the un-foveated forearm, given visual capture of the foveated hand, during hand-fixation | 60 |
| 3.2.3.2 | Likelihood of visual capture of the un-foveated hand, given visual capture of the foveated forearm, during forearm-fixation..... | 62 |
| 3.3 | Discussion..... | 63 |
| 4 | GENERAL DISCUSSION | 67 |
| | REFERENCES | 74 |
| Appendix | | |
| | IRB/HUMAN SUBJECTS APPROVAL | 77 |

LIST OF FIGURES

| | | |
|-------------|--|----|
| Figure 2.1: | Mean percentage of ‘mirror’ responses for hand and forearm positions at each timepoint during synchronous stroking of the hands (orange) and the forearms (teal). Values at time=0 indicate responses recorded before stroking began. Error bars represent 95% CIs. | 32 |
| Figure 2.2: | Mean percentage of ‘mirror’ responses for the hand and forearm positions at t=60s of synchronous stroking of the hands (orange) and the forearms (teal). Error bars represent 95% CIs. | 38 |
| Figure 2.3: | Mean percentage of trials with ‘fragmented visual capture’ of the viewed hand (orange) and forearm (teal), during synchronous stroking of the hands and the forearms. Error bars represent 95% CIs. | 39 |
| Figure 2.4: | Percentage of mirror responses for the forearm, given a mirror response for the hand on the same trial versus no mirror response for the hand on the same trial - when the hands were stroked. Error bars represent 95% confidence intervals. | 41 |
| Figure 2.5: | Percentage of mirror responses for the hand, given a mirror response for the forearm on the same trial v. no mirror response for the forearm on the same trial - when the forearms were stroked. Error bars represent 95% confidence intervals. | 42 |
| Figure 3.1: | Mean percentage of ‘mirror’ responses for hand and forearm positions at each timepoint during synchronous stroking of the full lower arm, as participants fixated on the hand (orange) or the forearm (teal) in the mirror. Values at time=0 indicate responses recorded before stroking began. Error bars represent 95% CIs. | 53 |
| Figure 3.2: | Mean percentage of ‘mirror’ responses for the hand (orange) and forearm positions (teal) at t=60s of synchronous stroking, during hand-fixation and the forearm-fixation. Error bars represent 95% CIs... | 58 |
| Figure 3.3: | Mean percentage of trials with ‘fragmented visual capture’ of the viewed hand (orange) and forearm (teal), hand-fixation and forearm-fixation. Error bars represent 95% CIs. | 59 |

Figure 3.4: Percentage of mirror responses for the forearm, given a mirror response for the hand on the same trial v. no mirror response for the hand on the same trial, during hand-fixation. Error bars represent 95% confidence intervals..... 61

Figure 3.5: Percentage of mirror responses for the hand, given a mirror response for the forearm on the same trial v. no mirror response for the forearm on the same trial – during forearm-fixation. Error bars represent 95% confidence intervals..... 62

ABSTRACT

In the mirror box illusion, synchronous multisensory stimulation of a hidden hand alongside a reflected hand in the mirror can lead to 'visual capture' – where participants experience their hidden hand in the location of the reflected 'mirror hand'. This effect is attributed to two main factors influencing multisensory integration: (1) cross-modal congruence between visual, tactile and/or motor signals, and (2) lower unisensory variance of vision compared to proprioception, weighting the integrated outcome toward the 'visual' input ('mirror hand'). Previous studies have primarily explored how these factors influence visual capture of a single, stimulated body part (e.g., visual capture of a viewed 'mirror hand' during synchronous bimanual movements), with limited investigations into how adjacent, unstimulated body parts are represented. Given the tendency to experience the body as a unitary whole composed of connected body parts, what happens when optimal integration of immediate multisensory information leads to 'disconnected' positions for connected body parts? We explored this question in two experiments, where participants' hands were placed at unequal distances on either side of a mirror to create a spatial disparity between the 'visual' and 'proprioceptive' estimates of the hand and forearm behind the mirror. In Experiment 1, we manipulated the degree of cross-modal congruence between the two parts by synchronously stroking either the hands or the forearms, in isolation. In Experiment 2, we manipulated visual input variance by having participants fixate on either their hand or forearm while they were stroked synchronously across both parts. Analyses of participants' perceived positions of their

hand and forearm revealed stronger ‘local’ effects of visual capture at the site of stimulation (Experiment 1) or foveation (Experiment 2), as participants were significantly more likely to perceive either part of their hidden arm ‘in the mirror’ when it was the location of stimulation/foveation. Moreover, stimulation/foveation of a given part significantly raised the likelihood of ‘fragmented visual capture’ of that part in the mirror (e.g., simultaneously perceiving the hand but NOT the forearm in the ‘mirror’ location, when the hands were stroked). Meanwhile, we found distinct ‘global’ effects in both experiments, as visual capture of the stimulated (Experiment 1) or foveated (Experiment 2) part of the arm significantly raised the likelihood of visual capture for the adjacent unstimulated/un-foveated part. Taken together, these results indicate that multisensory integration for the body is driven by a weighted integration of bottom-up properties of ‘online’ sensory information (e.g., cross-modal congruence; unimodal variance), with top-down information from a stored, ‘offline’ body representation of how the different body parts are connected.

Chapter 1

INTRODUCTION

To provide a coherent percept of the body in space, the brain must integrate information from different sensory modalities such as vision, touch and proprioception (Maravita et al., 2013). Since individual sensory signals are inherently noisy, this cross-modal integration provides an improved estimate of the body by combining information from multiple sources. However, the process of integrating information from different sensory inputs involves unique challenges.

First, this multisensory integration is only beneficial when the brain correctly identifies which sensory inputs to integrate, as integrating inputs that arise from different sources would be counterproductive to this goal. Given the complex multisensory environments encountered in daily life, determining which stimuli to integrate, and which to segregate, is not a trivial task. Secondly, there is the problem of unimodal sensory noise - any sensory signal transmitted from a receptor to its respective sensory cortex is inherently noisy due to spontaneous neuronal firing as a result of stochastic transduction channel activity (Baylor et al. 1980, Laughlin & Lillywhite 1982; Rieke & Baylor 2000), which in turn varies across modalities due to differences in specialized neuronal properties. Thus, the brain must also find a way to deal with the complexity introduced by different noise levels for each sensory input.

1.1 How does the brain solve the unique challenges in multisensory integration?

Previous studies have explored how the brain tackles these challenges when integrating multisensory information to perceive one's environment. One theory for how the brain determines which sensory inputs to integrate is the 'causal inference model' (Kording et al., 2007), which predicts how different sensory inputs will be integrated based on the likelihood of those inputs arising from a 'common cause'. This likelihood is computed by a Bayesian estimator that takes into account the different sensory signals, their variances, and a 'prior' probability based on previous experiences involving those stimuli. This computed likelihood then informs whether the different stimuli will be integrated, and if so, how the integration will affect how each stimulus is perceived.

Evidence for this model was provided using an experiment where participants were presented with either visual, auditory or simultaneous audiovisual stimuli at different locations along the azimuth, and participants were instructed to report where the stimulus was located in each trial. In addition, participants were asked to report their subjective perception of 'unity' between the auditory and visual stimuli on trials where both were present (i.e., whether they shared a 'common cause'). Taking the response distributions from the auditory- and visual-only trials as measures of their unimodal variances, they found that the proposed causal inference model performed significantly better than other models at predicting a perceived common cause between auditory and visual stimuli. They also observed a strong positive bias in the perceived location of the auditory stimulus toward the visual stimulus when a common cause was perceived, whereas a strong negative bias away from the visual stimulus was observed when separate causes were perceived, supporting the idea that the

perception of a common cause between different sensory inputs leads to those inputs being integrated into a single percept at one location.

Next, we have the question of how the brain handles the different levels of unimodal sensory noise (i.e., ‘variance’) when integrating stimuli from different modalities. In the above study, response distributions for stimulus location in the auditory-only trials were more variable than in the visual-only trials. Meanwhile, when a common cause was perceived between the visual and auditory stimuli on trials where both were presented together, it was always the auditory location response that was biased toward the visual, rather than the other way round. This suggests that when different sensory inputs are integrated into a single percept, the resulting perception of each input is closer to input with the lowest variance. This is in line with the ‘optimal integration’ model proposed by Ernst and Banks (2002), where different sensory signals are optimally integrated to minimize variance in the final integrated estimate. This is done using a maximum-likelihood estimator (MLE) that computes a weighted average of the sensory inputs where each input is multiplied by the reciprocal of its unimodal variance. To test this model, Ernst and Banks (2002) conducted a series of experiments where participants judged whether two heights (of varying height difference) were equal or different, using 1) haptic information, 2) visual information and 3) simultaneous haptic and visual information, while the unimodal noise of the visual input was manipulated to be higher, lower or equal to that of the haptic input. Results from the visual-haptic combined task showed that participants’ responses were strongly predicted by the MLE formula, where the sensory modality with the highest manipulated variance was consistently weighted the lowest in the final estimate.

Overall, this model supports the idea that multisensory integration serves to improve the final estimate by minimizing the effects of inherent sensory noise.

Taken together, these studies provide a framework that addresses how the brain deals with the two key challenges in multisensory integration. While the causal inference model explains how the brain determines which sensory inputs to integrate based on the computed likelihood of a ‘common cause’ between the inputs, the maximum-likelihood estimator explains how sensory variance is minimized in the final integrated estimate by weighting the different inputs based on their respective unimodal variances.

1.2 How do these multisensory integration principles inform one’s perception of their body?

Although the multisensory integration principles demonstrated above were primarily in the context of perceiving one’s environment, the same principles have been observed in the integration of multisensory information regarding the body. Evidence for these mechanisms can be found in numerous studies that have used body illusions such as the ‘rubber hand illusion’ (Botvinick and Cohen, 1998) and the ‘mirror-box illusion (Holmes & Spence, 2005; Medina et al., 2015) to study how the brain represents the body.

In the rubber hand illusion (RHI), synchronous stroking of a viewed rubber hand along with one’s own hidden hand leads to experiences of ‘ownership’ over the rubber hand, as well as the feeling that the touch they are experiencing originates from the brush stroking the rubber hand. When asked to localize their real, hidden hand, participants report locations biased in the direction of the rubber hand – an effect

called ‘proprioceptive drift’. These effects disappear when the rubber hand and real hand are stroked asynchronously, indicating that visuotactile congruence plays a role in the integration of vision, touch and proprioception in this illusion (Botvinick & Cohen, 1998).

A similar effect is seen in the mirror box illusion (MBI), where a participant’s hands are placed at unequal distances on either side of mirror placed at the mid-sagittal plane. This creates a spatial disparity between the proprioceptive estimate of the real hand hidden behind the mirror and the visual estimate of the reflected hand in the mirror. When the participant taps both index fingers synchronously while looking at the mirror in this setup, a strong effect of ‘visual capture’ is seen, where real hand behind the mirror is perceived to share the same location as the reflected hand in the mirror (Holmes and Spence, 2005; Medina et al., 2015). As in the RHI, this effect is accompanied by reports of ‘ownership’ over the viewed hand in the mirror. Interestingly, these effects of the MBI not only decrease during asynchronous tapping, but also when both hands are held still (Medina et al., 2015). This indicates that the temporally dynamic nature of cross-modal congruence from synchronous tapping likely contributes to multisensory integration.

These illusions provide considerable evidence supporting both causal inference and optimal integration models of multisensory integration. In both illusions, multisensory integration is enhanced under conditions of cross-modal congruence (i.e., synchronous tapping/stroking). According to the causal inference model, multisensory integration is driven by the Bayesian estimation of a ‘common cause’ that takes into account a ‘prior’, based on previous experiences with that set of stimuli. In the case of

the MBI, previous experiences of concurrent visual, tactile and motor sensations from moving one's finger provide evidence for a 'common cause' between synchronous visual, tactile and motor inputs. Thus, prolonged synchronous tapping in the mirror box leads to an accumulation of evidence that strengthens the 'prior' for a common cause, resulting in increased multisensory integration. Evidence for the role of causal inference in the RHI has been explicitly demonstrated by Chancel et al. (2022), where the RHI was induced under different conditions of visual noise. The results showed that the causal inference model outperformed a non-Bayesian model in predicting the emergence of the RHI, with a strong, direct correlation between prior probability of inferring a common cause between the visual and tactile inputs, and the emergence of the RHI. Meanwhile, the effect of visual capture in the MBI provides strong evidence for the optimal integration model – given the significantly lower variance of the visual input of the reflected hand compared to the proprioceptive input from the real hand (Welch et al., 1979), the strong visual bias observed in the final, integrated estimate is consistent with the MLE principles underlying this model.

A key point worth noting is the role of temporal congruence in this process. Although Kording et al. (2007) focused on studying causal inference under conditions of different spatial disparities between the simultaneously presented stimuli, several other studies have shown a clear effect of temporal congruence on multisensory integration on both a behavioral (Lewald et al., 2001) and neural level (Royal et al., 2009). This could explain why a stronger MBI was found in the synchronous tapping condition compared to the 'no movement' condition in Medina et al. (2015). Although the 'no movement' condition provides static multisensory information from the

identically stationary ‘mirror’ and real hands, the lack of change in these stimuli limits opportunities for the brain to assess temporal relationships between these inputs. In contrast, the dynamic nature of synchronous tapping provides multiple opportunities to assess temporal congruence between the stimuli, which provides additional evidence that they come from the same source, yielding increased binding compared to the no movement condition.

1.3 How are locally integrated position estimates for different body parts integrated into a unitary ‘whole-body’ percept?

Previous studies on body illusions have primarily focused on investigating multisensory integration for isolated body parts. For instance, the classic MBI paradigm uses congruent multisensory stimulation of the hand to induce a perceived shift in the location of that hand toward the visual location (i.e., the ‘mirror’ hand). However, since these experiments only ever asked the participant about the stimulated hand, they have not been able to provide any insights on how the brain represents the rest of the body when the shift occurs. Given that humans experience their body as a unitary whole, this information gap leaves open a fundamental question about body perception: If the perceived position of each body part is constructed by optimally integrating local multisensory information from that part, how are these optimal estimates for different parts stitched together to produce a contiguous, ‘global’ whole-body percept? Despite some theoretical discussions comparing ideas of ‘local’ and ‘global’ body schemata (de Vignemont, 2021), there has been limited empirical research directly addressing this question.

One theory is that in addition to an ‘online’ body representation that informs the current position of the body based on optimal integration of immediate multisensory information, there is also a stored, ‘offline’ representation of the body that holds information regarding the possible positions the body can assume at any point in time, given the body’s biomechanical constraints (Tsakiris, 2010). The stored, offline representation of the body is more temporally stable than the ‘online’ representation, as it does not change with immediate multisensory input. Such representations are evidenced by phenomena such as phantom limb perception in amputees where a ‘phantom’ of the amputated limb is perceived despite the lack of multisensory information suggesting its existence (Ramachandran & Hirstein, 1998), or ‘body integrity identity disorder’ where patients experience a lack of ownership toward an intact limb despite the presence of multisensory information (Brang et al., 2008). Thus, a contiguous, whole-body percept may arise from the integration of ‘online’ and ‘offline’ body representations, where the former estimates the current state of the body by integrating immediate multisensory information while the latter ensures congruency of the integrated estimates with a global body representation (Tsakiris, 2010).

Some studies have shown evidence for the influence of a stored, internal body representation on multisensory integration for the body. Liu and Medina (2017) demonstrated how the knowledge of biomechanical constraints of the body can influence multisensory integration using a novel variation of the MBI, where they induced illusory hand rotation by having participants perform synchronous bimanual movements with their hands in opposite orientations (palm-up/palm-down) in a mirror

box. They found that the perceived rotation was less likely when wrist rotation was more biomechanically difficult (palm-down to palm up) compared to when it was easier (palm-up to palm-down), even though no actual wrist movement was involved. In another study, Tsakiris and Haggard (2005) showed that the RHI was less likely to occur when the rubber hand was in an ‘incongruent posture’ (Experiment 1; rubber hand rotated 90°) or ‘incongruent identity’ (Experiment 2; opposite handedness). Taken together, these results indicate that multisensory integration is driven not only by bottom-up properties of multisensory information such as cross-modal congruence, but also by top-down information about the time-invariant characteristics of the body from a stored, internal representation.

Despite evidence indicating an interaction between top-down and bottom-up information in multisensory integration, the nature of this interaction remains unclear. More specifically, the current literature lacks a clear understanding of how the brain handles situations where optimal integration of multisensory information conflicts with a stored, internal representation of the body. In such scenarios, how do these two kinds of information integrate to construct a coherent percept of the body? For instance, if an optimal integrated estimate is prioritized when integrating multisensory information for each part of the body, what happens when the ‘optimal’ estimates for two adjacent parts of the body are incongruent, as dictated by a stored internal representation of the body?

A few alternative possibilities come to mind. First, it may be the case that in situations of conflict, top-down information from the stored, internal body representation dominates multisensory integration for different parts of the body. This

would constrain ‘optimal integration’ for each adjoining body part to solutions where connections between adjoining body parts are maintained, to preserve the integrity of a whole-body percept. However, there is substantial evidence contradicting this account. In Tsakiris and Haggard (2005)’s third experiment, they synchronously stroked either the index or little finger of the real and rubber hand, but measured proprioceptive shift for both fingers. If the entire hand was represented as a whole representation, one would predict that a shift in the position estimate of the stroked finger would result in a shift in the position estimate of the unstroked finger. However, they only found significant proprioceptive drift toward the rubber hand for the finger that was stroked, not for the unstroked finger. This was replicated in a fourth experiment where one finger was stroked synchronously, while the other was stroked asynchronously with respect to the rubber hand; significant drift was only found for the synchronously stroked finger. Another study looked at passive visuomotor stimulation of different fingers, and once again found local, ‘fragmented’ effects where the only stimulated finger showed significant proprioceptive drift toward a projected image of the hand (Tsakiris et al., 2006). These results do not support an account of multisensory integration where a stored, ‘global’ representation of the body completely constrains optimal integration of immediate multisensory information, which would predict any perceived shifts in body position to be global across adjacent parts of the stimulated part.

Another possibility is that multisensory integration for any given body part is entirely driven by bottom-up properties of the immediate multisensory stimuli being integrated. In this case, the perceived position of each body part would be the optimal

integrated estimate given the specific sensory inputs local to that part, regardless of how these parts are connected to each other within a global body representation. However, while the studies demonstrating local, ‘fragmented’ body illusions (Tsakiris & Haggard, 2005 – Experiments 3 & 4; Tsakiris et al., 2006) support this account, other evidence pointing to clear effects of the influence of a stored, offline body representation on multisensory integration (Tsakiris & Haggard, 2005 – Experiments 1 & 2; Liu & Medina, 2017) contradicts this account.

A third possibility in the event of a conflict between top-down (e.g., stored, global body representation) and bottom-up factors (e.g., cross-modal congruence, unimodal variances) driving multisensory integration, is that both kinds of information are weighted to varying degrees in the final percept of the body. Findings from a recent experiment done in our lab provide evidence for this case. This experiment involved a variation of the MBI where both the hand and forearm were reflected in the mirror. Participants were asked to report where they perceived their hand and forearm to be positioned with respect to the reflected ‘mirror limb’ as they tapped both index fingers synchronously, asynchronously, or held them stationary. Participants were asked about the perceived position of their hand or forearm, choosing one out of four categorical response options – ‘mirror’, indicating they felt like it shared the location of the reflected limb in the mirror; ‘hidden’, indicating that it was in a hidden location behind the mirror that did not match that of the reflected limb; ‘both’, indicating that it was located in both the ‘mirror’ and ‘hidden’ positions simultaneously, and ‘neither’ for when they could not determine the position of their limb.

Based on the possibilities outlined above, if multisensory integration for individual body parts is purely based on the optimal integration of immediate sensory information, we would expect the differences in multisensory information between the hand and the forearm to result in significant differences in visual capture for each, with frequent occurrences of ‘fragmented’ visual capture of one part of the viewed arm without the other. If instead, optimal integration of immediate multisensory information for each part is completely constrained by top-down influences of a stored ‘global’ representation of the body, we would expect visual capture of either part of the viewed arm to always occur alongside visual capture of the adjacent part. Finally, if optimal integration of immediate multisensory information is only partially influenced via top-down information from a stored, ‘global’ body representation, then we would expect to see both ‘local’ and ‘global’ effects of visual capture simultaneously, such that there will be significant differences in visual capture for each part of the arm based on local multisensory information, but each part will also raise the likelihood of visual capture of the other, adjacent part. The results of this experiment were consistent with this third hypothesis. On one hand, participants reported visual capture of the reflection of the hand significantly more often than that of the adjacent forearm (and often reported feeling like their hand was in the mirror but their forearm was not, on the same trial), indicating a fragmented illusion. On the other hand, movement synchrony was still shown to have an effect on forearm visual capture, as the forearm was significantly more likely to be localized in the ‘mirror’ location during synchronous compared to asynchronous tapping. Moreover, it was found that visual capture of the reflected hand significantly raised the likelihood of

visual capture of the forearm on the same trial. Thus, we found that while there were significant local differences in visual capture between the two parts, there were also significant, albeit weaker global effects of visual capture across both adjacent body parts, indicating weighted effects of both bottom-up and top-down processes in the final integrated estimates of body position.

1.4 How do bottom-up factors like cross-modal congruence and unimodal variance interact with top-down information from a stored body representation to generate integrated estimates of the body position?

Although these incidental findings have interesting implications for how bottom-up and top-down processes may interact to different degrees to inform body position estimates, it is not immediately obvious what may have contributed to observed ‘fragmentation’ of visual capture in this experiment. Indeed, given that participants saw a visibly connected hand and forearm in the mirror throughout the manipulation, one would expect that full visual capture of one segment of the viewed arm would not occur in the absence of the other. This is notably more dramatic than the results seen in the work of Tsakiris et al. (2005; 2006) described above, where the difference between stimulated and unstimulated fingers were at the centimeter-level for proprioceptive drift, and never involved full visual capture of the rubber hand.

The aim of this study was to dig deeper into how a stored, offline body representation may interact with specific bottom-up factors that influence optimal integration and causal inference mechanisms, and how these interactions may give rise to distinct local and global effects of visual capture in the MBI. To this end, we identified two specific variables that may have contributed to the local differences in visual capture between the hand and the forearm in our previous findings – i. the

location of synchronous multisensory stimulation, corresponding to varied cross-modal congruence across the two parts, and ii. the location of visual fixation, corresponding to varied unimodal variance across the two parts. We performed two separate MBI experiments to study the effects of these two potential factors – one where we manipulated the location of synchronous visuotactile stimulation (brush-stroking) on the arm, and another where we manipulated the location of the visual fixation point on the reflected image of the arm. Studying each factor in a separate experiment allowed us to explore how two different factors involved in bottom-up processing of multisensory information – cross-modal congruence and unimodal variance – individually interact with top-down influences of a stored, internal representation of the body, to construct how we perceive our bodies.

1.4.1 Experiment 1 – the location of synchronous multisensory stimulation

One potential explanation for the significantly higher rate of visual capture for the hand compared to the visibly adjacent forearm in our previous experiment may be the variation in information to be integrated across the two parts of the viewed arm. Although both the hand and the forearm were clearly visible in the mirror, the origin of synchronous movement (tapping index fingers) was at the hand, while the forearm remained stationary. As established previously, the MBI is significantly stronger when multisensory congruence is dynamic (e.g., synchronous tapping) compared to when it is static (no movement), possibly due to increased opportunities for assessing temporal congruence between the different sensory modalities during dynamic stimulation (Medina et al., 2015). Thus, increased visual capture of the reflected hand may be because there is more information providing evidence that information from each

modality should be combined in comparison to the static forearm. This is supported by the results of previous RHI studies that showed increased proprioceptive drift for stimulated fingers compared to unstimulated fingers (Tsakiris & Haggard, 2005; Tsakiris et al., 2006).

However, it is also important to account for the global visual capture effects we observed in our previous experiment, as we found that there were significant effects of movement synchrony on both parts of the reflected arm despite the movement being localized at the hand. This may be explained by a top-down influence of a stored, global body schema. As dynamic cross-modal congruence directly promotes visuoproprioceptive integration, and thus visual capture of the ‘mirror’ hand, stored knowledge of the anatomical connection between the hand and forearm causes this primary, local effect to influence the perceived location of connected forearm. However, this secondary, global effect of forearm visual capture is weaker than the primary effect for the hand, which is directly mediated by dynamic cross-modal congruence.

To test this theory, we performed an experiment using a similar mirror-box setup to our original experiment, where the participants’ arms were positioned at unequal distances from a mirror large enough to reflect both hand and forearm. Participants were then stroked synchronously with two identical brushes on either their hands or their forearms in isolation, and were asked to report where they perceived their hidden hand and forearm to be located with respect to the mirror. To control for potential effects of foveation on visuoproprioceptive integration (see section 1.4.2) in this experiment, we instructed participants to fixate on their wrist throughout the

experiment such that both parts of the arm were in peripheral vision, to eliminate relative differences in visual input variance across the two parts.

Based on our original findings, we predicted strong local effects of visual capture, such that a ‘mirror’ response for the perceived location of either part of the arm would occur significantly more often when it was the location of stroking, compared to when the other part was stroked. We also predicted a significant dissociation between hand and forearm visual capture in both conditions, such that the part being stroked would experience visual capture significantly more often than the unstimulated part. Moreover, we predicted that this dissociation would lead to more frequent experiences of ‘fragmented’ visual capture of the stimulated part of the arm, where participants would report visual capture of the stimulated part in the absence of visual capture for the adjacent, unstimulated part on the same trial. Finally, we predicted a global, ‘spreading’ effect of visual capture from the stimulated part to the connected, unstimulated part, such that visual capture of the stimulated part would significantly raise the likelihood of visual capture of the unstimulated part in that trial.

1.4.2 Experiment 2 – the location of the visual fixation point

Another potential factor that may explain a higher rate of visual capture for the hand compared to the forearm during synchronous tapping, may be related to where participants foveated while viewing the ‘mirror limb’ during the manipulation. Given the instruction to tap their fingers synchronously or asynchronously in different conditions, participants may be likely to foveate on their hand to maintain the correct finger movements. We know that visual capture in the MBI is largely explained by the fact that optimal integration favors the sensory input with the lowest unimodal

variance, which in this case is the visual input of the reflected limb in the mirror. Meanwhile, it has long been established that foveal vision has significantly lower variance than peripheral vision, especially in the judgment of spatial phase relationships (Rentschler & Treutwein, 1985) and relative position estimates (Levi & Klein, 1986). Given that proprioceptive variance remains constant, the lower variance of foveal vision may result in an even stronger weighting of the visual input during optimal integration for the foveated ‘mirror hand’, and thus stronger visual capture, compared to the unfoveated ‘mirror forearm’. Curiously, the role of foveation in multisensory integration remains unexplored at the time of this writing. Thus, evidence for this would also provide novel insight into a new factor to consider in multisensory integration overall.

To test this hypothesis, we used a similar paradigm as Experiment 1, where participants’ arms were positioned at unequal distances on either side of a mirror, such that both the hand and forearm were clearly visible in the mirror. To isolate the effect of visual fixation on multisensory integration for each part of the arm, participants were instructed to focus on one of two fixation points – either on the hand or the forearm – as the full length of both lower arms were stroked synchronously, to control for the potential effect of the location of synchronous stroking. Our predictions for this experiment were analogous to our predictions for Experiment 1 – we predicted a significantly higher rate of visual capture for a given part of the arm when it was foveated location compared to when it was not, and that there would be significantly more visual capture for the foveated part of the arm, compared to the unfoveated part in each fixation condition. We also predicted foveating on either part would raise the

likelihood of experiencing ‘fragmented’ visual capture of that part in the mirror, in the absence the adjacent, unfoveated part on the same trial. Finally, we predicted global effects of integration to spread from foveated to unfoveated parts of the viewed arm, such that visual capture of the foveated part would raise the likelihood of visual capture of the unfoveated part.

1.5 How do local and global effects of visual capture vary over the time-course of the mirror-box illusion?

A secondary question of interest relates to the temporal progression of the MBI across the duration of sustained synchronous multisensory stimulation. Previous work using the RHI has shown that body illusions take time to occur (Ehrsson et al., 2004; Lloyd, 2007), and that illusion onset time is significantly faster for synchronous compared to asynchronous conditions (Lane et al., 2017). This is consistent with the hypothesis that cross-modal congruence enhances multisensory integration (and thus embodiment) in body illusions by providing an accumulation of evidence for a ‘common cause’ over time, which increases the prior in the causal inference model that predicts multisensory integration. Thus, the likelihood of experiencing the illusion should grow with increased duration of congruent multisensory stimulation.

However, to our knowledge, only one study has explicitly explored how the degree of illusion varies over the duration of stimulation. Rohde et al. (2011) studied how proprioceptive drift varied over time by measuring the participant’s perceived location of their hidden hand on a centimeter scale at different timepoints during synchronous stroking. When position estimates were measured at 10-second intervals over a 120-second period of stimulation, they found a steady increase in

proprioceptive drift over the first 40 seconds, followed by a plateau. This makes intuitive sense, as the rate increase in illusion strength is likely to diminish as the illusion approaches its maximum effect for a given condition.

Although this study provides a solid starting point for understanding the time-course of multisensory integration in body illusions, several aspects of the problem remain unexplored. First, the specific time-course of the MBI remains unexplored, and inherent differences between the RHI and the MBI are such that results from one illusion may not generalize to the other. For instance, the relative lack of realism of a rubber hand compared to the reflection of one's own hand makes for a substantially weaker illusion compared to the MBI. This is especially obvious when comparing perceived positional changes in the two illusions – while the MBI frequently induces full visual capture of the viewed 'mirror hand', this is uncommon in the RHI, where perceived positional changes in the RHI are typically observed as partial shifts toward the viewed rubber hand.

Furthermore, there are qualitative differences between the two illusions that suggest different underlying mechanisms for seemingly shared effects. For instance, Rohde et al. (2011) found that despite the lack of a distinct effect of synchronous stroking on proprioceptive drift, it still resulted in significantly higher ratings of 'ownership' compared to the 'vision only' condition. This dissociation between drift and ownership directly contradicts the results of Medina et al. (2015), which showed not only a strong correlation between ownership ratings and drift, but also a clear, positive effect of synchronous movement on both drift as well as ownership relative to no movement (i.e., 'vision only'). This suggests potentially different mechanisms for

perceived positional changes in the two illusions. Given that these differences are directly related to the explanations provided by Rohde et al. (2011) for their time-course results, it is likely that the MBI's time-course has different attributes worth exploring separately.

Finally, there is the question of how the time-course of the MBI may differ for local versus global effects of the illusion. Does the time-course of multisensory integration look different based on the relative weightings of bottom-up (local) and top-down (global) processes for a given body part? To explore this question, we included a time variable in the designs of both Experiments 1 and 2 described above. In both experiments, participants were asked to report where they felt their hand and forearm were located (with respect to the mirror) at the start of each condition block, before stroking began. Synchronous stroking was then administered in 60 s blocks (according to condition parameters), during which participants were asked to report their perceived locations of their hand and forearm at 10-second intervals. This provided us with a timeline of how visual capture progressed for each part of the arm under each condition, which allowed us to compare the local and global effects of visual capture for each part of the arm, as well as study the main effect of visual capture over time.

To summarize, our main objectives for this secondary question of 'time-course' were to characterize the temporal dynamics of visual capture in the MBI, and compare the time-courses of local and global visual capture under different conditions of local cross-modal congruence and unimodal variance. Due to limited overall information on the time-course of these illusions, our goals were more exploratory

than confirmatory. However, a few possibilities may be considered based on the limited information available. First, given the theory that cross-modal congruence enhances multisensory integration via a gradual accumulation of evidence that raises the inferred likelihood of a ‘common cause’ between the stimuli, we may expect to see a gradual increase in the rate of visual capture for a given part of the arm as it is stroked synchronously. Given the results seen in Rohde et al. (2011), we may also expect to see a decay and eventual plateau in the slope of this increase, as the effect reaches its maximum for a given condition. Next, we may see a similar gradual increase in global visual capture alongside local visual capture, as our previous results indicate that the latter significantly raises the likelihood of the former at a given point in time. Finally, given the stronger effects of cross-modal congruence or unimodal variance on local visual capture compared to global, we may see a steeper initial slope for the former, as larger increases at each consecutive time point will lead to a faster rate of increase over time.

Chapter 2

EXPERIMENT 1 – THE EFFECT OF THE LOCATION OF SYNCHRONOUS MULTISENSORY STIMULATION ON HAND AND FOREARM VISUAL CAPTURE IN THE MBI

The broad theoretical question for this experiment concerns how local multisensory congruence interacts with top-down information from a stored, ‘global’ body representation during multisensory integration, to determine the final position estimates of different, connected body parts. To explore this, we designed an experiment to address the more targeted question of how the location of synchronous multisensory stimulation may differentially affect visual capture of two segments of a viewed ‘mirror arm’ – the hand and forearm – in the MBI.

2.1 Materials and Methods

2.1.1 Participants

Forty-three participants (15 male, ages 18-21 years) were recruited from the SONA participant pool at the University of Delaware, consisting of undergraduate students taking an introductory psychology course. Participants were compensated with 0.5 SONA credits which went toward their grade in the course. Voluntary informed consent was obtained from each participant. All studies were approved by the University of Delaware IRB.

2.1.2 Apparatus

A mirror box setup was used, which consisted of a black wooden base (92 cm wide × 42 cm deep) with a vertical mirror (30.5 cm tall × 37.3 cm deep) mounted at the center of the base, aligned with the participant's midsagittal plane. The participant's left arm was positioned facing the mirror, with the middle finger of their left hand positioned 9 cm from the mirror on the distal edge of the wooden base from the participant. The right arm was placed behind the mirror, with the middle finger positioned 27 cm behind the mirror on the distal edge of the wooden base, creating an 18 cm difference between the mirror reflection of the hand (what we will refer to as the 'mirror' hand) and their real location of the hand hidden behind the mirror. A black cloth was draped over the right shoulder and arm to conceal any visual information of the real arm's position. The arms were placed such that they were parallel to each other and the central mirror, and the participant could easily see both their hand and forearm in the mirror. Two identical 1-inch-wide foam brushes were used to stroke both arms.

2.1.3 Experimental Paradigm

We used a modified version of the MBI paradigm in which visuotactile stimuli were administered by using two identical brushes to stroke both limbs synchronously. Using brush-stroking to administer multisensory stimuli is a well-established method of inducing embodiment in the RHI (Botvinick & Cohen, 1998; Tsakiris & Haggard, 2005; Rohde et. al., 2011), an illusion that shares similar multisensory integration mechanisms as the MBI. The effectiveness of synchronous stroking in inducing the MBI was verified in a pilot study. Since we were primarily interested in the effect of

the location of multisensory stimuli on visual capture and were not concerned with effects of synchronicity, we only used synchronous stroking in this experiment. For our dependent variable of interest, we operationalized ‘visual capture’ as a function of how frequently participants felt their actual limb and the mirror limb shared the same location.

To observe the effect of synchronous stroking on visual capture of the location of the arm being stroked, as well as the adjacent, unstroked part of the viewed arm, we set our independent variable as ‘stroking location’ – a categorical variable with two levels – (1) hand-stroking, where both hands were stroked synchronously, and (2) forearm-stroking, where both forearms were stroked. We studied the effect of this variable on two dependent variables – (1) the perceived position of the hand, and (2) the perceived position of the forearm. On each trial, participants were asked to report where they perceived each part of their hidden arm to be located categorically with respect to the mirror. For each question, participants were asked to respond by choosing one out of four response choices:

1. “Mirror” - if it shared the same location as the reflection of the hand/forearm seen in the mirror
2. “Hidden” - if it was in a separate ‘hidden’ location behind the mirror; i.e., where the hand actually was located
3. “Both” - if they felt like it simultaneously inhabited both the ‘mirror’ and ‘hidden’ locations
4. “Neither” – if they could not determine the location of their hand/forearm

These responses were then used to operationalize ‘visual capture’ as the binary response variable ‘mirror or both’, with ‘1’ representing a ‘mirror’ or ‘both’ response

and ‘0’ representing and a ‘hidden’ or ‘neither’ response – indicating the presence or absence of any illusion on that trial.

2.1.4 Procedure

There were ten blocks in total – five blocks of hand stroking and five blocks of forearm stroking. The blocks were counterbalanced for condition in ABBAABBAAB order (A = hand-stroking; B = forearm-stroking). To control for any potential effects of visual fixation, participants were instructed to fixate on their wrist in the mirror throughout each block. The experimenter paid close attention to make sure the participant followed the fixation instruction while stroking, redirecting them if their gaze strayed.

On each trial, participants were asked to report where they felt their hand and their forearm were located at regular, 10-second intervals during each 60-second block of stroking. Participants responded with one out of the four choices described in the previous section (‘mirror’, ‘hidden’, ‘both’, ‘neither’). Participants were reminded at each trial to keep their gaze fixed on their wrist even while answering the questions.

The following protocol was followed for each block:

1. The participant was instructed to look at their wrist in the mirror, and not look anywhere else until otherwise instructed, with the experimenter closely monitoring the participant to ensure their gaze did not stray from the instructed location.
2. Before beginning stroking, the participant was asked:
 - Where do you feel your is hand located? (options: mirror, hidden, both, neither)
 - Where do you feel your forearm is located?

3. The participant received synchronous brushstrokes on both [hands/forearms] continuously for 10 seconds
4. The two questions from step 2 were repeated, reminding the participant to keep their gaze focused on their wrist in the mirror while answering. The experimenter paused stroking to record answers, and immediately resumed stroking again after, once again reminding the participant not to shift their gaze.
5. Steps 1-4 was cycled through 5 more times, switching question order for each new trial – for a total of 6 trials per block.

Between each block and the next, participants were instructed to remove their arms from the mirror box and shake them five times to eliminate carryover illusion effects between blocks. After shaking, the experimenter placed their middle fingers back on the marked points (9 cm from the mirror on the left; 27 cm behind the mirror on the right).

2.1.5 Data Analysis

Based on the established relationship between dynamic cross-modal congruence and multisensory integration (Botvinick & Cohen, 1998; Holmes & Spence, 2005; Medina et. al., 2015) and the effects of ‘fragmented visual capture’ we observed in our previous MBI experiment, we predicted the following main effects for our primary question of interest concerning the role of the location of synchronous multisensory stimulation on hand and forearm visual capture:

- Prediction 1: A significant main effects of ‘stroking location’ on visual capture of each individual segment of the viewed arm, such that visual capture of a given part (e.g., ‘mirror hand’) will be significantly more likely when it is the location of synchronous stroking (e.g., hand-stroking), compared to when the other part is stroked (e.g., forearm-stroking).

- Prediction 2: A significant difference between hand and forearm visual capture under each stroking condition, such that the stroked part of the arm will experience significantly higher rates of visual capture compared to the adjacent, un-stroked part.

In addition to confirming these predictions, we were also interested in investigating the minimally explored time-course of the MBI. For this question, we were interested in characterizing the overall time-course of visual capture in the MBI, as well as comparing how this time-course may vary based on whether the visual capture is driven by ‘local’ or ‘global’ integration processes.

A series of binomial generalized linear mixed models (GLMMs) were used to collectively address these questions. For the first prediction, the data were first divided into two sets based on ‘question’ (i.e., queried part of the arm), such that one set contained all the responses for ‘hand position’ and the other contained all the responses for ‘forearm position’, enabling separate analyses of the fixed effect of ‘stroking location’ on hand and forearm visual capture.

In addition to testing Prediction 1, these same models were simultaneously used to explore the time-course of visual capture, and how it may vary for ‘stimulated’ and ‘unstimulated’ parts of the viewed arm (representing local and global integration, respectively). When visualizing the results, we noticed that responses could be divided into two phases – a sharp increase in reports of visual capture in the first ten seconds, followed by a more gradual increase after ten seconds. Based on this, we separated the data into two further subsets based on time, with an ‘initial phase’ subset containing trials from $t=0-10$ s, and a ‘subsequent phase’ subset containing trials from $t=10-60$ s. This yielded a total of four datasets – one for each combination of ‘question’ and time-period. We performed separate binomial GLMM on each of these four datasets, with

each model including fixed effects of ‘stroking location’, ‘time’ and a ‘time × stroking location’ interaction term, along with a random intercept for ‘participant’ and a random slope for ‘stroking location’. Each of these models used the binary response variable ‘mirror or both’, coded for the presence (1) or absence (0) of a ‘mirror’ or ‘both’ response for a given question on a given trial.

Since the ‘initial phase’ only contained two trials – one before the manipulation began, and one after, we expected the effect of the first trial after 10 seconds of stroking to be cancelled out by the trial before stroking began. Thus, we predicted that we would not see a significant effect of ‘stroking location’ in the binomial GLMMs for this phase. Meanwhile, since each trial in the ‘subsequent phase’ was preceded by isolated synchronous stroking of one part of the arm, we predicted that we would see a significant main effect of ‘stroking location’ on both hand and forearm visual capture, such that visual capture for the stroked part of the arm would be stronger than for the unstroked part.

For effects of time, due to increasing accumulation of evidence for a ‘common cause’ with increased duration of synchronous multisensory stimulation, we expected to see increases in visual capture for the stimulated location over time. Meanwhile, we expected increases visual capture of the stimulated location to result in subsequent increases in visual capture of the unstimulated location, due to top-down influences of a stored, internal representation of the body containing information on the anatomical connection between the two parts. Thus, we expected both the hand and forearm to show increased visual capture over time in both stroking conditions, leading to a main effect of ‘time’ on visual capture in our binomial GLMMs for both the initial and

subsequent phases of stimulation. Since our visualization of the data showed the slope of increase to be steeper in the first 10 seconds compared to subsequent trials, we expected this main effect of time to be stronger in the ‘initial phase’ models compared to the ‘subsequent phase’ models. This would be consistent with the findings in Rohde et. al. (2011), where they found the slope of increase in proprioceptive drift to decrease and eventually plateau over time as the illusion approached its maximum limit for a given condition.

Finally, we anticipated a significant interaction effect between ‘time’ and ‘stroking location’ in both the ‘initial’ and ‘subsequent’ phase models, as multisensory congruence would provide stronger evidence for a ‘common cause’ at the location of stimulation compared to the adjacent, unstimulated location, resulting in larger increases at each timepoint and stronger visual capture over time for the stimulated location. We also anticipated that this interaction effect would be stronger in the ‘initial phase’ of stimulation compared to the ‘subsequent phase’, as differences in the rate of increase in ‘global’ and ‘local’ visual capture are likely to level out as both approach their respective plateaus at maximum illusion.

A similar approach was used to test Prediction 2, with a few differences. Since we already used the previous four models to explore the question of time-course, we simplified the analyses for Prediction 2 by removing ‘time’ as a variable, and only looking at the final trials of each block ($t=60$ s). These trials were then divided into two subsets based on ‘stroking location’, such that one set contained all the ‘hand-stroking’ trials and the other contained all ‘forearm-stroking’ trials. A separate binomial GLMM was then performed on each dataset with ‘question’ as a fixed effect

variable to look at the significance and direction of the ‘dissociation’ between hand and forearm visual capture for each stroking location. Both models also included a random intercept for participant and a random slope for ‘question’. All of the above analyses were performed using the lme4 package in R.

To delve deeper into the dissociation between hand and forearm visual capture observed in our previous findings, we investigated whether isolated synchronous stroking of a given part of the arm would increase the likelihood of ‘fragmented’ visual capture of that specific part in the mirror. Specifically, we predicted that the percentage of trials in which visual capture is experienced for the hand, but NOT the forearm, would be significantly higher when the hand is stroked compared to when the forearm is stroked, and vice versa. To test this, we looked at occurrences of visual capture of one part of the viewed arm in the absence of the other using two new binary variables – ‘only mirror hand’, which coded for instances where there was a ‘mirror’ response for the hand and a ‘hidden’ or ‘neither’ response for the forearm on the same trial, and ‘only mirror forearm’, which coded for instances where there was a ‘mirror’ response for the forearm and a ‘hidden’ or ‘neither’ response for the hand. We then performed two separate binomial GLMMs looking at the fixed effect of stroking location on ‘fragmented visual capture’ of the hand and the forearm respectively, to see if ‘fragmented visual capture’ of a given part was significantly more likely to occur when it was the location of synchronous stroking, compared to when it was not.

Finally, we predicted that alongside strong ‘local’ effects of synchronous stroking on the visual capture of a stimulated part of the arm, we would see simultaneous evidence for the top-down influence of a stored ‘offline’ body

representation in the form of ‘global’ effects of synchronous stroking on the visual capture of unstimulated parts of viewed ‘mirror arm’. Specifically, we predicted that we would find the likelihood of visual capture for the unstimulated part of the arm to be significantly higher given visual capture of the stimulated part, on the same trial.

To test this, we computed the following probabilities:

1. For hand-stroking: A ‘mirror’ or ‘both’ response for the forearm, given a ‘mirror’ or ‘both’ response for the hand on the same trial
2. For forearm-stroking: A ‘mirror’ or ‘both’ response for the hand, given a ‘mirror’ or ‘both’ response for the forearm on the same trial

To calculate the probability of a ‘mirror’ response for the forearm given a ‘mirror’ response for the hand on the same trial, we coded another binary variable where ‘1’ was assigned to all trials where the response for hand was ‘mirror’, and ‘0’ assigned to all trials where the response for hand was not ‘mirror’. The percentage of ‘mirror’ responses for the forearm was then calculated across the trials for each of these two groups, and compared to each other to assess the likelihood of forearm embodiment in relation to whether or not the hand was embodied. Finally, a binomial GLMM was performed with ‘hand in mirror’ as a fixed variable on the likelihood of a ‘mirror’ response for the forearm on the same trial. A random intercept for each participant and a random slope for ‘hand in mirror’ were also included, to account for baseline variation as well as random variation of hand embodiment across participants.

The same process was then used to calculate the probability of a ‘mirror’ response for the hand, given a mirror response for the forearm on the same trial, in the forearm-stroking condition, as well as analyze the effect of forearm embodiment on hand embodiment on a given trial using a binomial GLMM.

All confidence intervals for mean visual capture in each group were calculated using normal approximation method for binomial confidence intervals, using the formula: $M \pm 1.96 \frac{\sqrt{M(1-M)}}{n}$ for 95% CIs.

2.2 Results

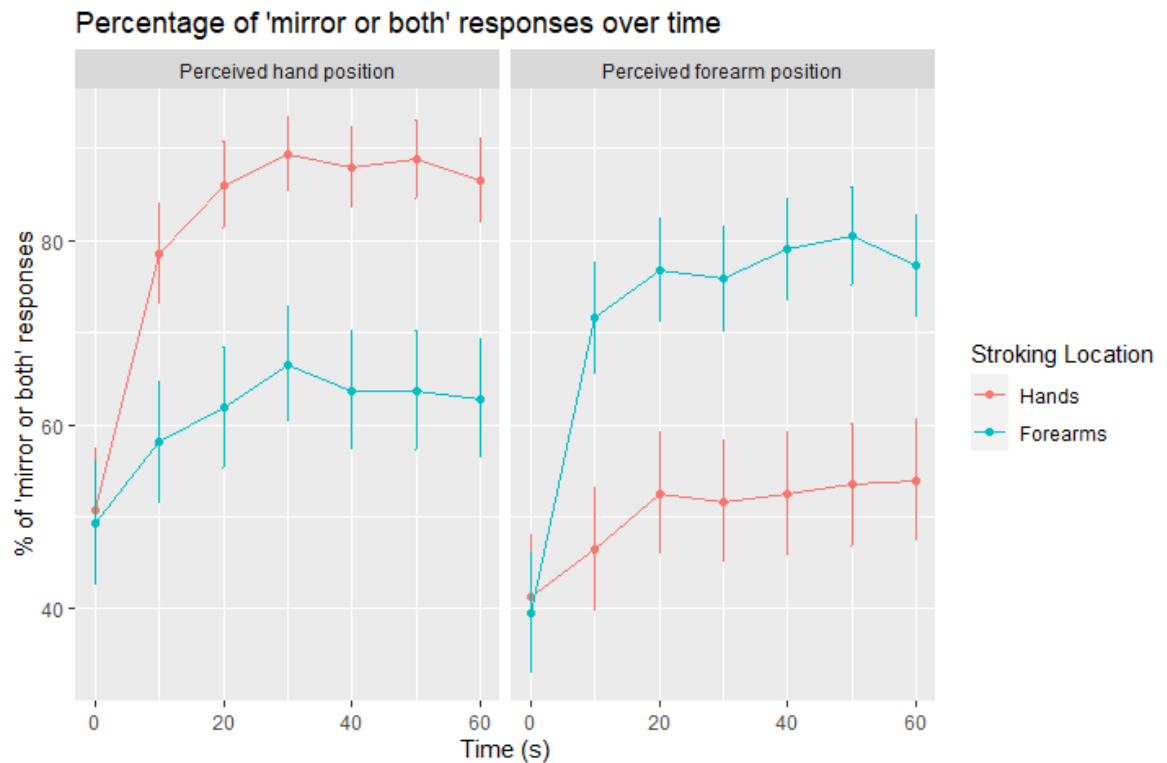


Figure 2.1: Mean percentage of 'mirror or both' responses for perceived hand and forearm positions at each time point, during synchronous stroking of the hands (orange) and the forearms (teal). Values at time=0 indicate responses recorded before stroking began. Error bars represent 95% CIs.

As predicted, a visualization of the data (Figure 2.1) shows larger percentages of ‘mirror or both’ responses (i.e., ‘visual capture’) for the hand when the hands were stroked, and larger percentages for the forearm when the forearms were stroked. We also found that the gradient of increasing visual capture over time is steepest toward the beginning of the block, which then plateaus over the course of the block. Moreover, the sharp initial increase in visual capture for either part of the arm is steeper when it is the location of stimulation, compared to when it is not. We analyzed each of these effects using a series of binomial GLMMs, as reported below.

2.2.1 Within-question analyses – The effect of stroking location on visual capture of the hand and the forearm, respectively

In our first set of analyses, we looked at the effect of ‘stroking location’ on visual capture of the hand and the forearm separately, in the initial (0-10 seconds) and subsequent (10-60 seconds) phases of the illusion. The same models were also used to explore the time-course of visual capture for each of the two parts, across these two time phases.

2.2.1.1 ‘Initial phase’ of stimulation – 0-10 seconds

Consistent with our predictions, we found a significant ‘time × stroking location’ interaction effect on hand visual capture ($z = 3.468$, $p < .001$), indicating a significantly greater increase in ‘mirror or both’ responses for the hand over the first 10 seconds when the hands were stroked ($M = +27.90\%$, 95% CI = [+17.62%, +38.19]) compared to when the forearms were stroked ($M = +8.84\%$, 95% CI = [-0.59%, +18.27%]). We also found a main effect of ‘time’ on visual capture of the

‘mirror hand’ ($z = 6.846, p < .001$) in our binomial mixed model looking at ‘perceived hand location’ over the first 10 seconds of synchronous stroking, indicating that participants significantly more likely to respond with ‘mirror’ or ‘both’ for their perceived hand location after 10 seconds of synchronous stroking ($M = 68.37\%$, 95% CI = [63.98%, 72.77%]) compared to before stroking began ($M = 50.00\%$, 95% CI = [45.27%, 54.73%]), regardless of the location of stroking. As expected, there was no significant main effect of ‘stroking location’ over this time period ($z = .273, p = .784$) in this model.

Similarly, our binomial mixed model looking at the responses for ‘perceived forearm location’ also found a significant ‘time \times stroking location’ interaction effect on forearm visual capture ($z = 4.672, p < .001$), indicating that the increase in ‘mirror or both’ responses for the forearm over the first 10 seconds of stimulation was significantly greater when the forearms were stroked ($M = +32.10\%$, 95% CI = [+21.01%, +43.17%]) compared to when the hands were stroked ($M = +5.11\%$, 95% CI = [-3.64, +13.88]). However, we found no main effect of ‘time’ ($z = 1.494, p = .135$), indicating that the overall increase in forearm visual capture over the first 10 seconds across both stroking conditions was non-significant. As with the ‘hand response’ model, we found no main effect of ‘stroking location’ on forearm visual capture.

Overall, these results indicate sharper initial increases in visual capture for the stimulated part of the arm (i.e., ‘local’ visual capture) compared to the unstimulated part (i.e., ‘global’ visual capture), for both parts of the viewed arm.

2.2.1.2 ‘Subsequent phase’ of stimulation – 10-60 seconds

Next we examined performance from 10-60 seconds. For hand visual capture, we found a significant main effect of ‘stroking location’ ($z = 3.584$, $p < .001$), as participants were far more likely to respond with ‘mirror’ or ‘both’ when the hands were stroked ($M = 86.2\%$, 95% CI = [84.31%, 88.08%]) compared to when the forearms were stroked ($M = 62.79\%$, 95% CI = [60.15%, 65.43%]). We also found a significant main effect of time ($z = 3.308$, $p = .002$), indicating a gradual increase in ‘mirror’ or ‘both’ responses over time from $t = 10$ -60 seconds. However, we did not find a significant ‘time \times stroking location’ interaction ($z = .917$, $p = .359$), suggesting that there was no difference in the rate of increase in hand visual capture based on whether or not it was at the location of stroking.

We also found a similar main effect of ‘stroking location’ on forearm visual capture ($z = 4.043$, $p < .001$), such that participants were much more likely to respond with ‘mirror’ or ‘both’ for their forearm when their forearms were stroked ($M = 76.82\%$, 95% CI = [74.52%, 79.12%]) than when their hands were stroked ($M = 51.78\%$, 95% CI = [49.07%, 54.51%]). There was also a significant main effect of time ($z = 2.256$, $p = .024$), indicating a gradual increase in forearm visual capture over time from $t = 10$ -60 seconds. However, similar to responses for the hand, we did not find a significant ‘time \times stroking location’ interaction on mirror responses for the forearm ($z = .311$, $p = .756$) from 10-60 seconds.

Taken together, the results of our ‘within-question analyses’ are consistent with our prediction that visual capture of a given part of the viewed arm would be greater when it is the part being stroked, compared to when the adjacent part is stroked, as supported by the significant main effects of ‘stroking location’ in the

binomial mixed models for both the hand and forearm responses from $t = 10-60$ seconds. We also found support for our predictions regarding the time-course of visual capture, as visual capture gradually increased over the duration of the block for both the stimulated and unstimulated adjacent parts of the viewed ‘mirror arm’ over the course of the block, with a sharper initial slope of increase that decays as the illusion approaches maximum effect. Finally, we found support for our prediction that the sharper initial slope would be steeper for ‘local’ visual capture compared to ‘global’ visual capture, as there were larger increases for the stimulated part than the unstimulated part over the first 10 seconds and that these differences would level out with continued synchronous stimulation as the illusion approaches maximum effect for both locations. However, we had also anticipated a significant, albeit weaker interaction of time by stroking location in the $t = 10-60$ second phase, which we did not find.

2.2.2 Between-question analyses – comparing visual capture of the hand versus the forearm under each stroking condition

Next, we performed ‘between-question’ analyses, where we compared visual capture between the two parts of the viewed arm to explore the phenomenon of ‘fragmented visual capture’ of one part of the arm in the absence of the other, under each stroking condition.

2.2.2.1 The difference between hand and forearm visual capture in each stroking condition, after 60 seconds of stroking

In our previous work with the MBI, we had observed the interesting phenomenon of ‘fragmented’ visual capture, where participants often reported

experiencing their hand in the ‘mirror’ location and their forearm in the ‘hidden’ location on the same trial, as they synchronously tapped both index fingers. Given that synchronous tapping only provides congruent multisensory stimulation at the location of the hand (and not the forearm), we asked whether this effect of fragmented visual capture was driven by the location of synchronous multisensory stimulation. To investigate this, we first looked at the differences in visual capture for the hand compared to the forearm, within each stroking condition. We found that participants were significantly more likely to respond with ‘mirror’ or ‘both’ for their hand ($M = 86.51\%$, $95\% \text{ CI} = [81.94\%, 91.08\%]$) compared to their forearm ($M = 53.95\%$, $95\% \text{ CI} = [47.29\%, 60.61\%]$) when their hands were stroked ($z = 3.572$, $p < .001$), and significantly more likely to respond with ‘mirror’ or ‘both’ for their forearm ($M = 77.21\%$, $95\% \text{ CI} = [71.60\%, 82.82\%]$) compared to their hand ($M = 62.79\%$, $95\% \text{ CI} = [56.34\%, 69.25\%]$) when their forearms were stroked ($z = 2.451$, $p = .014$). These results provide further evidence that synchronous multisensory stimulation results in stronger ‘local’ effects of visual capture at the site of stimulation, compared to adjacent body parts. Mean percentages of ‘mirror or both’ responses for each condition at $t = 60$ seconds are plotted in Figure 2.2.

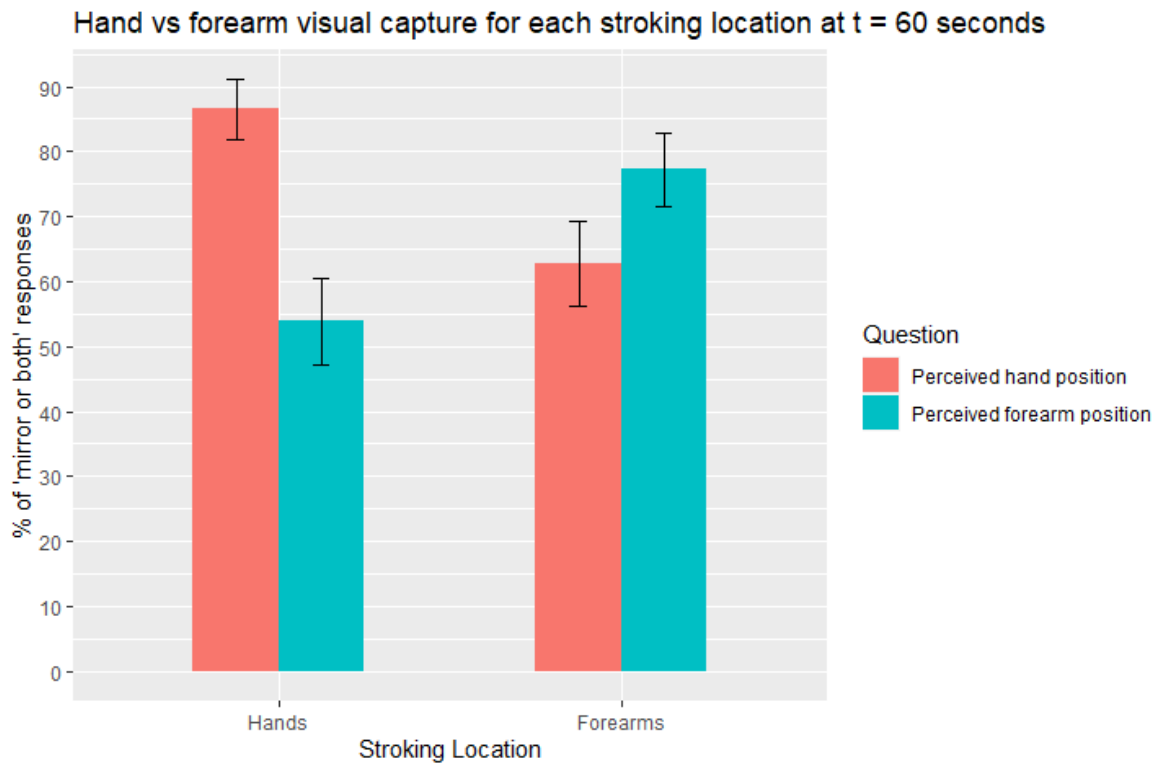


Figure 2.2: Mean percentage of ‘mirror or both’ responses for perceived hand (orange) and forearm (teal) locations at t = 60 seconds of synchronous hand-stroking and forearm-stroking. Error bars represent 95% CIs.

2.2.2.2 How the location of synchronous stroking drives ‘fragmented visual capture’ of a connected hand and forearm in the mirror

Next, we explicitly looked at instances of ‘fragmented visual capture’ of each part of the viewed arm under the two stroking conditions, by calculating percentages of trials where participants reported feeling their hand in the ‘mirror’ location and their forearm in the ‘hidden’ location on the same (i.e., ‘fragmented visual capture’ of the hand), and trials where they reported feeling their forearm in the ‘mirror’ location and

their hand in the ‘hidden’ location on the same trial (i.e., ‘fragmented visual capture’ of the forearm). The results are plotted in Figure 2.3.

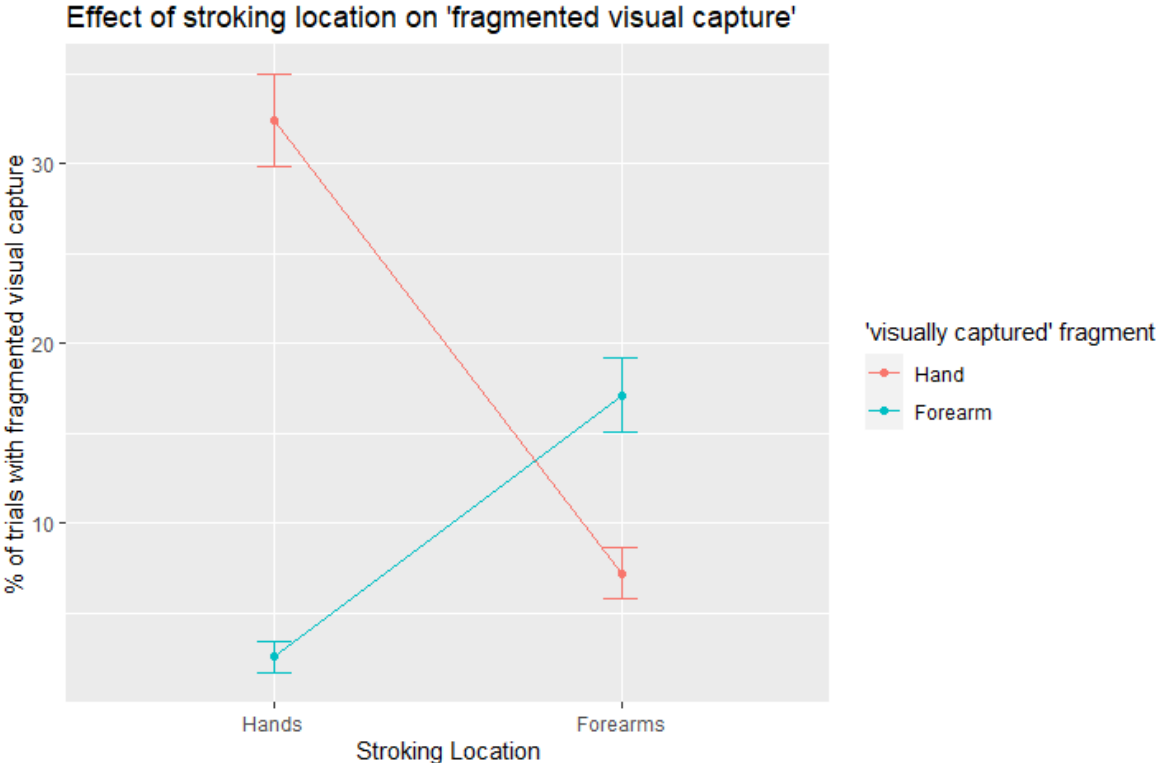


Figure 2.3: Mean percentage of trials with ‘fragmented visual capture’ of the hand (orange) and the forearm (teal), during synchronous stroking of the hands and the forearms. Error bars represent 95% CIs.

We then performed two binomial mixed models to look at the effect of ‘stroking location’ on the likelihood of ‘fragmented visual capture’ of the hand, and the forearm, respectively. As expected, participants were significantly more likely to experience ‘fragmented visual capture’ of the ‘mirror hand’ when the hands were

stroked ($M = 32.40\%$, 95% CI = [29.85%, 34.96%]) than when the forearms were stroked ($M = 7.21\%$, 95% CI = [5.80%, 8.62%]), $z = 4.148$, $p < .001$. Similarly, they were significantly more likely to experience ‘fragmented visual capture’ of the forearm when the forearms were stroked ($M = 17.13\%$, 95% CI = [15.08%, 19.19%]) than when the hands were stroked ($M = 2.56\%$, 95% CI = [1.70%, 3.42%]), $z = 2.563$, $p = .010$.

2.2.3 Top-down effects of a stored ‘global’ body representation on visual capture

Finally, to test our predictions regarding how stored knowledge of a connected hand and forearm in an internal representation of the body may influence ‘global’ effects on multisensory integration for unstimulated parts of the viewed arm, we looked at the conditional probabilities of visual capture of the unstimulated part of the viewed arm in the mirror, given visual capture of the stimulated part on the same trial.

2.2.3.1 Likelihood of visual capture of the unstimulated forearm, given visual capture of the stimulated hand

As predicted, when the hands were stroked, participants were significantly more likely to respond with ‘mirror’ or ‘both’ for the unstimulated forearm on trials where they also responded with ‘mirror’ or ‘both’ for the hand ($M = 58.14\%$, 95% CI = [55.38%, 60.02%]), compared to when they responded with ‘hidden’ or ‘neither’ for the hand ($M = 16.55\%$, 95% CI = [12.23%, 20.87%]), $z = 3.747$, $p < .001$. Summary statistics are plotted in Figure 2.4.

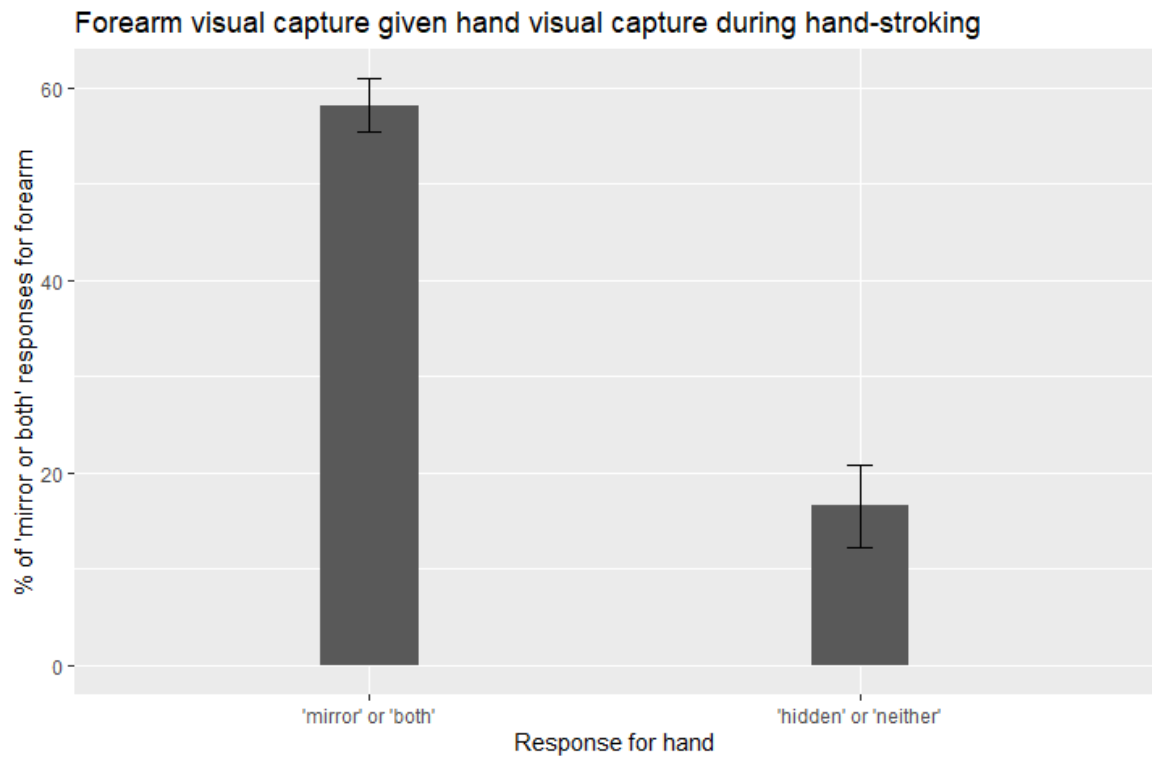


Figure 2.4: Percentage of ‘mirror or both’ responses for the forearm, given the presence or absence of a ‘mirror’ or ‘both’ response for the hand – when the hands were stroked. Error bars represent 95% confidence intervals.

2.2.3.2 Likelihood of visual capture of the unstimulated hand, given visual capture of the stimulated forearm

Similar effects were found when the forearms were stroked, where participants were significantly more likely to respond with ‘mirror’ or ‘both’ for the unstimulated hand on trials where they also responded as such for the forearm ($M = 93.79\%$, $95\% \text{ CI} = [92.07\%, 96.51\%]$), compared to when they did not ($M = 68.31\%$, $95\% \text{ CI} = [64.98\%, 71.65\%]$), $z = 4.137$, $p < .001$. See Figure 2.5 for plotted summary statistics.

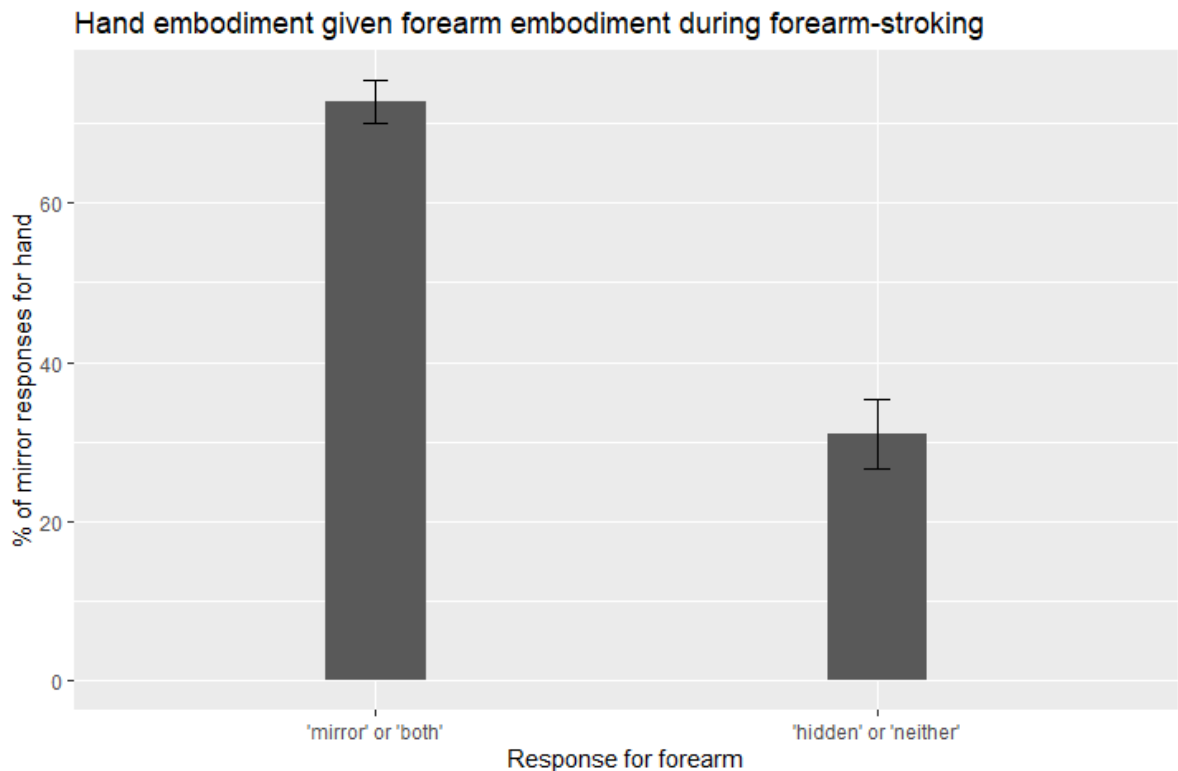


Figure 2.5: Percentage of ‘mirror or both’ responses for the hand, given the presence or absence of a ‘mirror’ or ‘both’ response for the forearm – when the forearms were stroked. Error bars represent 95% confidence intervals.

Taken together, these results indicate clear effects of a top-down influence of a stored information about the ‘global’ relationships between adjoining body parts, on multisensory integration for body position.

2.3 Discussion

The primary goals of this experiment were to look at the role of the location of synchronous multisensory stimulation on the visual capture of two adjacent parts of a viewed limb in the mirror, and to explore how this factor may interact with top-down

information from a stored body representation during multisensory integration. Given the known effects of dynamic cross-modal congruence on multisensory integration, we predicted strong ‘local’ effects of cross-modal congruence on visual capture of the stimulated segment of the viewed arm, alongside weaker ‘global’ effects of a stored internal body representation on visual capture of the unstimulated segment. Our results were largely consistent with these predictions – we found that individuals were significantly more likely to respond with ‘mirror’ or ‘both’ for their hand position when their hands were stroked synchronously than when their forearms were stroked, and were more likely to respond with ‘mirror’ or ‘both’ for their forearm position when their forearms were stroked than when their hands were stroked. These results are consistent with the findings of Tsakiris and Haggard (2005), where significantly more proprioceptive drift toward the rubber hand was seen for a synchronously stimulated finger compared to an unstimulated or asynchronously stimulated finger on the same hand.

Based on the incidental finding of a ‘dissociation’ between hand and forearm visual capture during synchronous tapping in our previous work, we also theorized that the location of synchronous stroking may be a driving factor behind this particular dissociation. To this end, we predicted that isolated synchronous stroking of either part of the arm would result in a significantly higher rate of visual capture for that part of the arm compared to the unstroked, adjacent part. We found substantial evidence for this prediction – on average, participants were significantly more likely to experience visual capture of the hand compared to the forearm after 60 seconds of hand-stroking, and were also significantly more likely to experience visual capture of the forearm

compared to the hand after 60 seconds of forearm-stroking. Thus, we found evidence that local differences in visual capture between different parts of a viewed ‘mirror arm’ can be systematically predicted by the location of dynamic, congruent multisensory stimulation on the viewed and hidden arms. To dig deeper into this dissociation, we also looked specifically at trials where participants experienced ‘fragmented visual capture’, i.e., where participants reported feeling one part of their arm in the ‘mirror’ location, but not the other (i.e. a ‘mirror’ response for one part alongside a ‘hidden’ or ‘neither’ for the other part on the same trial). Here, we found that isolated synchronous stroking of either part of the arm significantly raised the likelihood of ‘fragmented visual capture’ of that part of arm in the mirror. Taken together, these results suggest that the location of synchronous multisensory stimuli may be partially, but not entirely responsible for the previously observed dissociation between hand and forearm visual capture in our original experiment with synchronous tapping. Given that the dissociation between hand and forearm visual capture seen in this experiment is not equally prominent in under both stroking location conditions, there may additional factors responsible for the higher visual capture of the hand compared to the forearm seen in the original experiment.

Next, our exploratory analyses of the time-course of visual capture in the MBI revealed a significant main effect of time, such that visual capture for both parts of the arm increased gradually over the duration of stimulation for both stroking locations. Furthermore, we found that in all cases, there was a steeper initial rise in visual capture over the first 10 seconds, followed by a gradual flattening of the slope of increase as the illusion progressed. We also found distinct differences in the time-

courses of ‘local’ versus ‘global’ effects of visual capture, where the increase in visual capture in the ‘initial phase’ of stimulation ($t=0-10$ s) was significantly steeper for the stimulated part of the arm, compared to the unstimulated part. However, these differences leveled out to similar gradients with continued stimulation ($t=10-60$ s), as ceiling levels of visual capture were approached for a given condition. Overall, these results support the idea that continuous congruent multisensory stimulation ‘feeds’ the prior in the Bayesian inference of a ‘common cause’ between the congruent sensory inputs over time, resulting in increasing multisensory integration as the duration of stimulation increases. They also indicate that the accumulation of evidence is faster at the location of stimulation compared to an adjacent, unstimulated part during the initial phase ($t=0-10$ s) of stimulation.

Finally, we found evidence for ‘global’ effects of synchronous visuotactile stimulation on the adjacent, unstimulated part of the viewed arm in the mirror. In both stroking conditions, we found that on any given trial, visual capture of the stimulated part of the viewed arm significantly raised the likelihood of visual capture of the unstimulated part. This result appears to contradict Tsakiris and Haggard (2005)’s results where the effects of synchronous stroking on the RHI were purely local. One possible reason for this may be the inherent differences between the MBI and the RHI – given that the enhanced realism of the ‘mirror limb’ compared to a rubber limb, we were able to study stronger effects of full visual capture of the viewed limb, compared to the comparatively smaller effects of proprioceptive drift typically seen in the RHI. It may be the case that the larger overall effect of the MBI allows for the detection of

more fine-grained effects of the interacting top-down and bottom-up factors that influence multisensory integration.

Chapter 3

THE EFFECT OF THE LOCATION OF VISUAL FIXATION ON HAND AND FOREARM VISUAL CAPTURE IN THE MBI

In this experiment, we examined the role of the location of the visual fixation point in the embodiment of the hand and the forearm in the mirror, as well as its role in modulating the effect of disjointed hand and forearm location estimates (i.e., ‘fragmented visual capture’), even when both are visually present in the mirror. We also aimed to look at how the differences in unimodal variance of visual input between foveated and un-foveated vision may interact with top-down influences of a stored body representation during multisensory integration. Finally, we aimed to further explore the time-course of visual capture in the MBI, and how this time-course may vary for the different ‘local’ and ‘global’ effects of visual capture induced by varying levels of unimodal visual variance at each location.

3.1 Materials and Methods

3.1.1 Participants

The experiment included 41 participants (14 male, ages 17-21), taken from the SONA participant pool at the University of Delaware, consisting of undergraduate students taking an introductory psychology course. Participants were compensated with 0.5 SONA credits which went toward their grade in the course. Voluntary

informed consent was obtained from each participant. All studies were approved by the University of Delaware IRB.

3.1.2 Apparatus

A mirror box setup identical to the one used in Experiment 1 was used for this experiment. As in the previous experiment, identical 1-inch wide foam brushes were used to stroke both arms.

3.1.3 Experimental Paradigm

To study the role of foveation (i.e., the participant's visual fixation point in the mirror), on hand and forearm embodiment, we used the same synchronous brush-stroking method as in Experiment 1. However, instead of stroking only the hand or the forearm in isolation, the participants had both of their entire upper arms - from elbows to fingertips - stroked synchronously by the two brushes. The idea here was to provide an equal amount of synchronous multisensory stimuli across the entire upper arm to induce the illusion, while varying the participant's fixation point in the mirror between two conditions - i. the first knuckle of the middle finger of the reflected hand in the mirror ('hand location'), ii. the midpoint of the visible portion of their forearm in the mirror ('forearm location').

To ensure fixation throughout each condition, the experimenter monitored the gaze of the participant closely while stroking, and promptly redirected the participant if their gaze shifted from the instructed location. Participants were also reminded to keep their gaze fixated on the correct location when answering questions at each 10

second interval, as well as a general reminder not to shift their gaze for the rest of the block until instructed to do so.

As in Experiment 1, a discrepancy between the location of the ‘mirror arm’ and the participant’s actual hidden arm was created by placing the middle finger of the left visible hand at a point 9 centimeters away from the mirror, and the middle finger of the hidden right hand at a point 27 centimeters behind the mirror, creating a difference of 18 centimeters between the middle finger of the ‘mirror hand’, and the middle finger of the hidden hand behind the mirror. Both middle fingers were placed such that they were at the same distance away from the participant’s coronal plane. Maintaining the same hand positioning across experiments 1 and 2 allowed us to compare the effects of each variable on the embodiment of the hand and the mirror in the MBI.

3.1.4 Procedure

There were 10 blocks in total - 5 blocks for the ‘hand’ fixation point, and 5 for ‘forearm’ fixation point. The blocks were counterbalanced for condition in ABBAABBAAB order (A = ‘hand fixation’; B = ‘forearm fixation’). To control for any potential effects of stroking location, participants were stroked continuously from their elbows to their fingertips, and back, taking care to equally cover the entirety of their upper arm visible in the mirror.

To study the changes in embodiment over time and across conditions, participants were once again asked to respond on the locations of their hand and forearm at regular, 10-second intervals during each 60-second block of stroking. The following protocol was followed for each block:

3. The participant is instructed to look at the [first knuckle of their middle finger/midpoint of their forearm] in the mirror, and not look anywhere else until otherwise instructed.
4. Before stroking, participant is asked:
 - Where do you feel your hand located? (options: mirror, hidden, both, neither - as defined in Experiment 1)
 - Where do you feel your forearm is located?
5. The participant receives synchronous brushstrokes on both upper arms from elbow to fingertips, continuously for 10 seconds
6. Stroking is paused while the participant responds to the same 2 questions from step 2.
7. Synchronous stroking is resumed immediately after responses are noted. Steps 3 and 4 are repeated 5 more times, for a total of 6 trials per block.

3.1.5 Data Analysis:

The analyses we ran for Experiment 2 were largely analogous to those we ran for Experiment 1. Based on the known effects of unimodal variance on multisensory integration, and the observed effects of ‘fragmented visual capture’ in our previous results, we predicted the following main effects for our primary question of interest concerning the role of the location of visual fixation on hand and forearm visual capture:

- Prediction 1: A significant main effects of ‘fixation point’ on visual capture of each individual segment of the viewed arm, such that visual capture of a given part (e.g., ‘mirror hand’) will be significantly more likely when it is fixated on (e.g., hand-fixation), compared to when the other part is stroked (e.g., forearm-fixation).

- Prediction 2: A significant difference between hand and forearm visual capture under each fixation condition, such that the foveated part of the arm will experience significantly higher rates of visual capture compared to the adjacent, un-foveated part.

In addition, we were interested in the same questions regarding the time-course of visual capture as those explored in Experiment 1.

To address these questions, we used the same approach as in Experiment 1. The full dataset was first split into two subsets based on ‘question’, such that all ‘hand’ responses were in one subset and all ‘forearm’ responses were in the other, each of which was divided into two further subsets based on time period – one containing trials from $t=0-10$ s (‘initial phase’), and the other containing trials from $t=10-60$ s (subsequent phase’). A separate binomial GLMM was then run on each of these four datasets, each of which included ‘fixation point’, ‘time’ and a ‘time \times fixation point’ interaction term as fixed effects. In addition, both ‘subsequent phase’ models included a random intercept for ‘participant’ and random slope for ‘fixation point’, while both ‘initial phase’ models included a random intercept for ‘participant’ (a ‘fixation point’ random slope was excluded due to yielding a ‘singular’ fit). The same binary response variable of ‘mirror or both’ (‘1’ for a ‘mirror’ or ‘both’ response; ‘0’ for ‘hidden’ or ‘neither’) was used as the dependent variable in all models. Together, these models were used to address Prediction 1 as well as our questions regarding time-course.

For Prediction 2, we looked exclusively at final trials of each block ($t=60$ s). We split these trials into two sets based on ‘fixation point’, with all ‘hand-fixation’ trials in one set and all ‘forearm-fixation’ trials in the other. Two separate binomial GLMMs were then performed on each of these sets using ‘question’ as a fixed effect

variable, to look at the significance and direction of the ‘dissociation’ between hand and forearm visual capture under each fixation condition. Both models also included a random intercept for ‘participant’ and a random slope for ‘question’.

To dig deeper into this ‘dissociation’ we also investigated whether the relative decrease in unimodal visual variance of the foveated part of the arm would result in an increased likelihood of ‘fragmented visual capture’ for that part. To do this, we used the same two, new binary variables used in Experiment 1 – ‘only mirror hand’, which coded for the co-occurrence of a ‘mirror’ response for the hand and a ‘hidden’ or ‘neither’ response for the forearm on the same trial, and ‘only mirror forearm’, which coded for the co-occurrence of a ‘mirror’ response for the forearm and a ‘hidden’ or ‘neither’ response for the hand on the same trial. Two separate binomial GLMMs were performed to look at the fixed effect of fixation point on each of these two binary variables, to see if the likelihood of ‘fragmented visual capture’ of a given part of the arm was increased by fixating on that part during synchronous stroking.

Finally, we also computed the conditional probabilities of visual capture of the forearm, given visual capture of the hand on the same trial during hand-fixation, and visual capture of the hand given the visual capture of forearm on the same trial during forearm-fixation. We also analyzed the effect of visual capture of the stimulated part, on visual capture of the unstimulated part in both fixation conditions, using the same binomial GLMMs used in Experiment 1.

3.2 Results

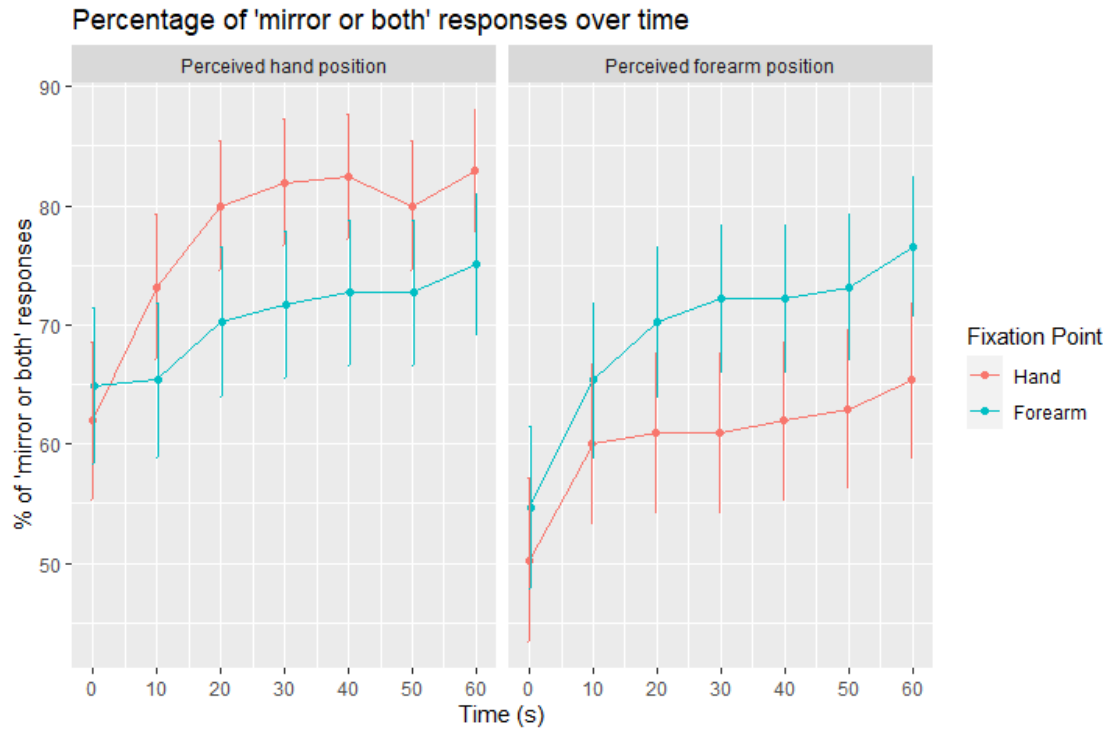


Figure 3.1: Mean percentage of ‘mirror or both’ responses for hand and forearm positions across time during synchronous stroking of the full lower arm, as participants fixated on the hand (orange) or the forearm (teal) in the mirror. Values at time=0 indicate responses recorded before stroking began. Error bars represent 95% CIs.

A visualization of the data (Figure 3.1) suggests an overall effect of the location of visual fixation, with higher mean percentages of ‘mirror or both’ responses for the hand when fixation is on the hand compared to the forearm, and higher mean percentages of ‘mirror or both’ responses for the forearm when fixation is on the

forearm, compared to the hand. Similar to the results from Experiment 1, there appears to be an overall effect of time in both conditions, as visual capture gradually increases over the duration of the block. In most cases, there appears to be a sharp initial increase in ‘mirror or both’ responses over the first 10 seconds, followed by smaller increases on subsequent trials that lead to an eventual plateau. However, the effect of fixation location on the time course of is less clear than the effect of stroking location seen in Experiment 1. In this graph, we see a sharp initial increase visual capture for the hand during hand fixation, but no significant increase in the forearm fixation. Meanwhile, there appears to be no difference in the initial increase in forearm visual capture between the two fixation conditions. We explored these effects using a series of binomial GLMMs, the results of which are described below.

3.2.1 Within-question analyses – The effect of the location of fixation on visual capture of the hand and the forearm, respectively

In our first set of analyses, we looked at the effect of the location of fixation on visual capture of the hand and the forearm separately, in the initial (0-10 seconds) and subsequent (10-60 seconds) phases of stimulation. We also used these models to explore the time-course of visual capture for each of the two parts, across these two phases of stimulation.

3.2.1.1 ‘Initial phase’ of stimulation – 0-10 seconds

We found a significant main effect of time on hand visual capture in the initial 10 seconds of the stimulation ($z = 2.829$, $p = .005$), indicating that participants were significantly more likely to respond with ‘mirror’ or ‘both’ for their perceived hand location after the first 10 seconds of stimulation ($M = 69.27\%$, $95\% \text{ CI} = [64.80\%$,

73.73%]) compared to before stroking began ($M = 63.41\%$, 95% CI = [58.75%, 68.08%]), regardless of the location of fixation. However, there was no significant time by fixation point interaction ($z = .230$, $p = .818$), indicating that the initial increase in ‘mirror or both’ responses for the hand over the first 10 seconds in the ‘hand fixation’ condition ($M = +11.22\%$, 95% CI = [+2.45%, 19.99%]) was not significantly greater than in the ‘forearm fixation’ condition ($M = +0.49\%$, 95% CI = [-8.22%, +9.20%]).

Similarly, we found a significant main effect of time on forearm visual capture over the first 10 seconds ($z = 2.558$, $p < .011$), as participants were significantly more likely to respond with ‘mirror’ or ‘both’ for their forearm location after 10 seconds of stimulation ($M = 62.68\%$, 95% CI = [58.00%, 67.34%]), compared to before stroking began ($M = 52.44\%$, 95% CI = [47.61%, 57.27%]). However, we once again found no significant time by fixation point interaction, indicating that the in the initial increase in forearm visual capture during ‘forearm fixation’ ($M = +10.73\%$, 95% CI = [+0.95%, 20.52%]) was not significantly greater than during ‘hand fixation’ ($M = +9.76\%$, 95% CI = [+0.45%, +19.06%]). As expected, there was no significant main effect of ‘fixation point’ in either of these models.

3.2.1.2 ‘Subsequent phase’ of stimulation – 10-60 seconds

Next, we examined the effect of foveation on visual capture from $t = 10-60$ seconds. We found a significant main effect of fixation point on visual capture for the hand ($z = 2.324$, $p = .020$), such that participants were significantly more likely to respond with ‘mirror’ or ‘both’ for their hand location when they were fixating on the hand in the mirror ($M = 80.08\%$, 95% CI = [77.85%, 82.31%]), compared to when

they were fixating on the forearm ($M = 71.30\%$, $95\% \text{ CI} = [68.77\%, 73.83\%]$). We also found a significant main effect of time ($z = 3.163$, $p = .002$) on hand visual capture in this phase, as ‘mirror or both’ responses gradually increased over the duration of synchronous stroking from $t = 10$ to 60 seconds. Once again, there was no time by fixation point interaction, indicating that the rate of increase in hand visual capture was not significantly different between the two fixation conditions during this period.

Our binomial mixed model looking at the fixed effects of fixation point, time and a time \times fixation point interaction effect on forearm visual capture during $t=10-60$ seconds did not yield significant results for any of these effects (fixation point: $z = 1.576$, $p = .115$; time: $z = 1.687$, $p = .092$, time \times fixation point: $z = 1.307$, $p = .191$).

Participants also appeared to respond with ‘mirror’ or ‘both’ for the forearm more frequently during ‘forearm fixation’ ($M = 71.63\%$, $95\% \text{ CI} = [69.12\%, 74.15\%]$) compared to ‘hand fixation’ ($M = 62.02\%$, $95\% \text{ CI} = [59.32\%, 65.75\%]$). However, our binomial mixed model looking at the fixed effects of fixation point, time and a time \times fixation point interaction effect on forearm visual capture during $t=10-60$ seconds did not yield significant results for any of these effects (fixation point: $z = 1.576$, $p = .115$; time: $z = 1.687$, $p = .092$, time \times fixation point: $z = 1.307$, $p = .191$).

3.2.2 Between-question analyses – comparing visual capture of the hand versus the forearm under each fixation condition

Next, we performed ‘between-question’ analyses, where we compared visual capture between the two parts, and studied the phenomenon of ‘fragmented visual

capture' of one part of the arm in the absence of the other, during each fixation condition.

3.2.2.1 The difference between hand and forearm visual capture in each fixation condition, at t = 60 seconds.

To investigate the potential role of fixation location in the previously observed effects 'fragmented' visual capture in our past experiment, we compared how often participants responded with 'mirror' for their perceived hand and forearm positions at t = 60 seconds of synchronous stroking, under each fixation condition. We found that participants were significantly more likely to respond with 'mirror' or 'both' for their hand (M = 82.93%, 95% CI = [76.65%, 89.21%]) compared to their forearm (M = 65.37%, 95% CI = [58.03%, 72.70%]) when they fixated on the hand in the mirror ($z = 3.122$, $p = .002$). However, there was no significant difference between hand visual capture (M=75.12%, 95% CI = [68.24%, 82.01%]) and forearm visual capture (M = 76.59%, 95% CI = [69.74%, 83.43%]) when they fixated on the forearm in the mirror ($z = .551$, $p = .582$). These results indicate that although the location of visual fixation may influence differences in 'local' and 'global' visual capture for different body parts in the MBI, the nature of these effects are not uniform across body parts. Figure 3.2 shows the different percentages of 'mirror' responses for hand and forearm positions at t=60 s synchronous stroking, for each fixation condition.

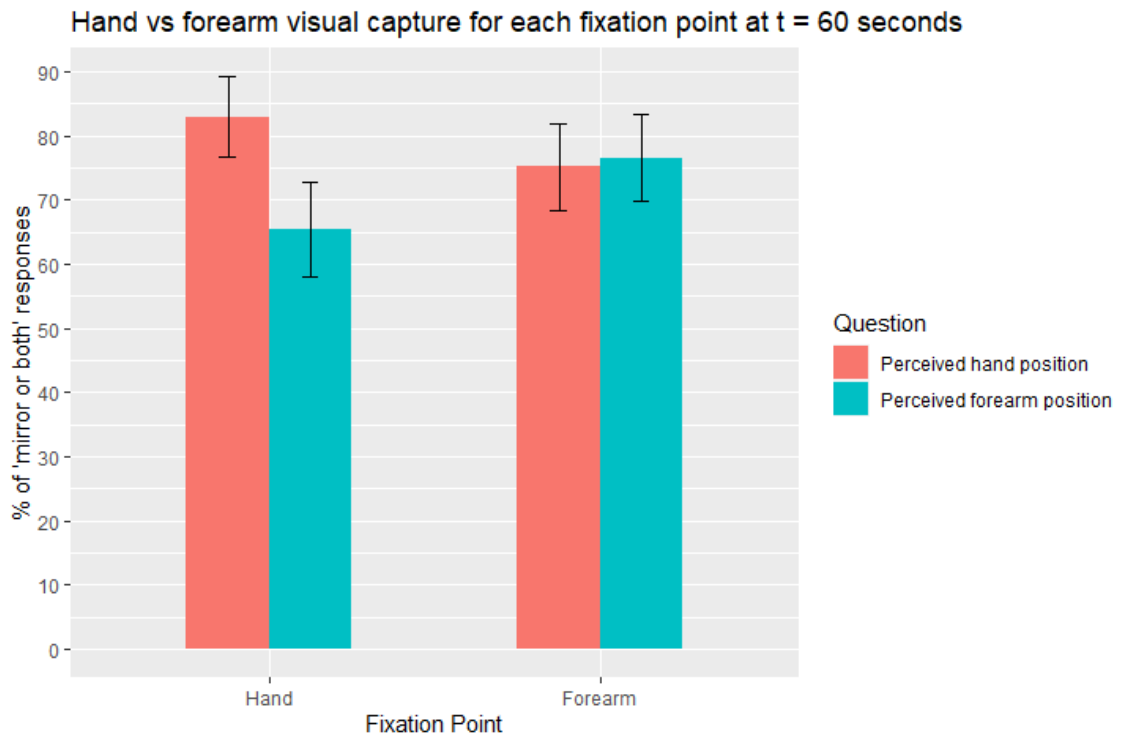


Figure 3.2: Mean percentage of ‘mirror or both’ responses for the hand (teal) and forearm positions (orange) at t = 60 seconds of synchronous stroking, during hand-fixation and the forearm-fixation. Error bars represent 95% CIs.

3.2.2.2 How the location of visual fixation drives ‘fragmented visual capture’ of a connected hand and forearm in the mirror

Next, we looked explicitly at instances of ‘fragmented visual capture’ of each part of the viewed arm during the two fixation conditions by calculating percentages of trials where participants reported ‘mirror’ for their hand position alongside ‘hidden’ or ‘neither’ for their forearm position on the same trial (i.e., ‘fragmented visual capture’ of the hand), and vice versa (i.e., ‘fragmented visual capture’ of the forearm). The results are plotted in Figure 3.3.

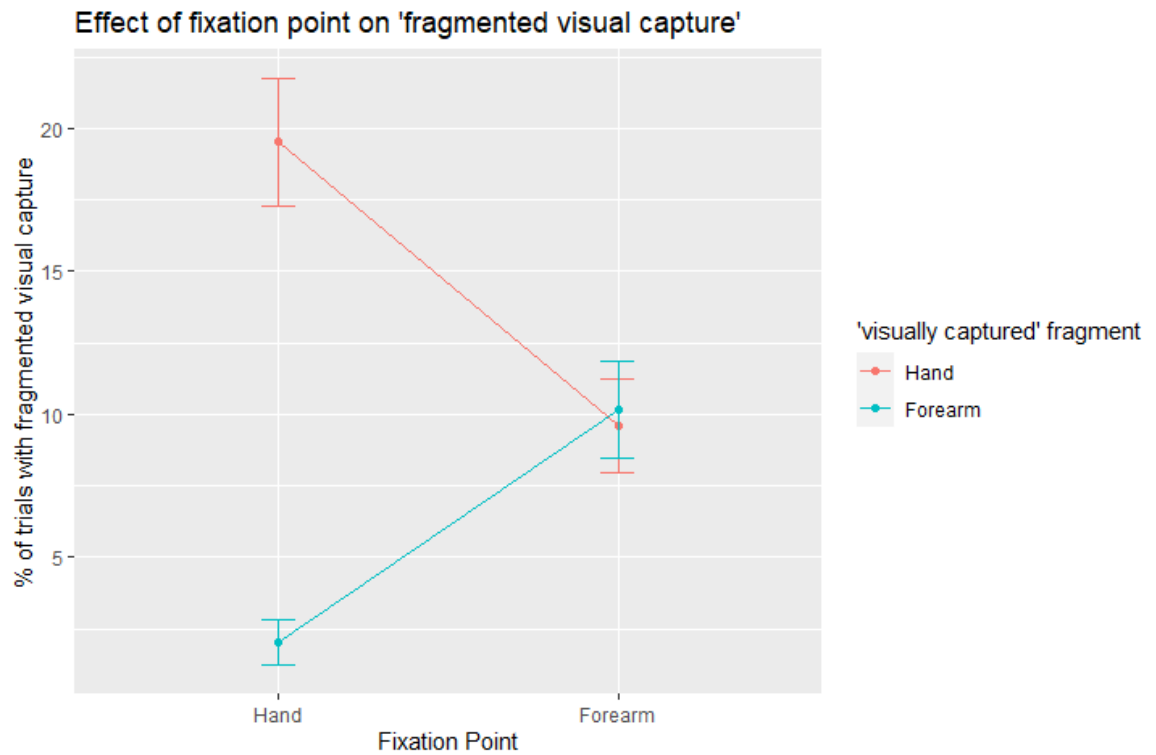


Figure 3.3: Mean percentage of trials with ‘fragmented visual capture’ of the viewed hand (orange) and forearm (teal) in the mirror, during hand-fixation and forearm-fixation, respectively. Error bars represent 95% CIs.

As predicted, participants were significantly more likely to experience ‘fragmented visual capture’ of the viewed hand when they fixated on the hand ($M = 19.51\%$, $95\% \text{ CI} = [17.30\%, 21.73\%]$), compared to when they fixated on the forearm ($M = 9.59\%$, $95\% \text{ CI} = [7.95\%, 11.24\%]$), $z = 3.016$, $p = .003$. Similarly, they were significantly more likely to experience ‘fragmented visual capture’ of the forearm when they fixated on the forearm ($M = 10.16\%$, $95\% \text{ CI} = [8.47\%, 11.85\%]$),

compared to when they fixated on their hand ($M = 2.03\%$, $95\% \text{ CI} = [1.24\%, 2.82\%]$), $z = 2.474$, $p = .013$.

Taken together, these results show that selective foveation on either part of the viewed arm significantly raises the likelihood of the foveated part of the arm being ‘visually captured’ in the mirror in the absence of the adjacent, un-foveated part. This indicates that although our binomial mixed models only found a significant effect of fixation point on visual capture of the hand (but not the forearm), the location of fixation may still play a role in driving ‘fragmented visual capture’ effects for both parts of the arm in the – albeit to a smaller degree than the location of stimulation – and may have been partially responsible for the dissociated effects of hand and forearm visual capture that we observed in our past experiment.

3.2.3 Top-down effects of a stored ‘global’ body representation on visual capture

Finally, to test our predictions regarding how stored knowledge of a connected hand and forearm in an internal representation of the body may influence ‘global’ effects on multisensory integration for un-foveated parts of the viewed arm, we looked at the conditional probabilities of visual capture of the un-foveated part of the viewed arm given visual capture of the foveated part on the same trial.

3.2.3.1 Likelihood of visual capture of the un-foveated forearm, given visual capture of the foveated hand, during hand-fixation

As predicted, when participants were instructed to fixate on their hand, they were significantly more likely to respond with ‘mirror’ or ‘both’ for the unfoveated forearm on trials where they also responded with ‘mirror’ or ‘both’ for the hand ($M =$

74.37%, 95% CI = [71.80%, 76.93%]), compared to trials where they responded with 'hidden' or 'neither' for the hand (M = 16.55%, 95% CI = [12.23%, 20.87%]), $z = 3.747$, $p < .001$. Summary statistics are plotted in Figure 3.4.

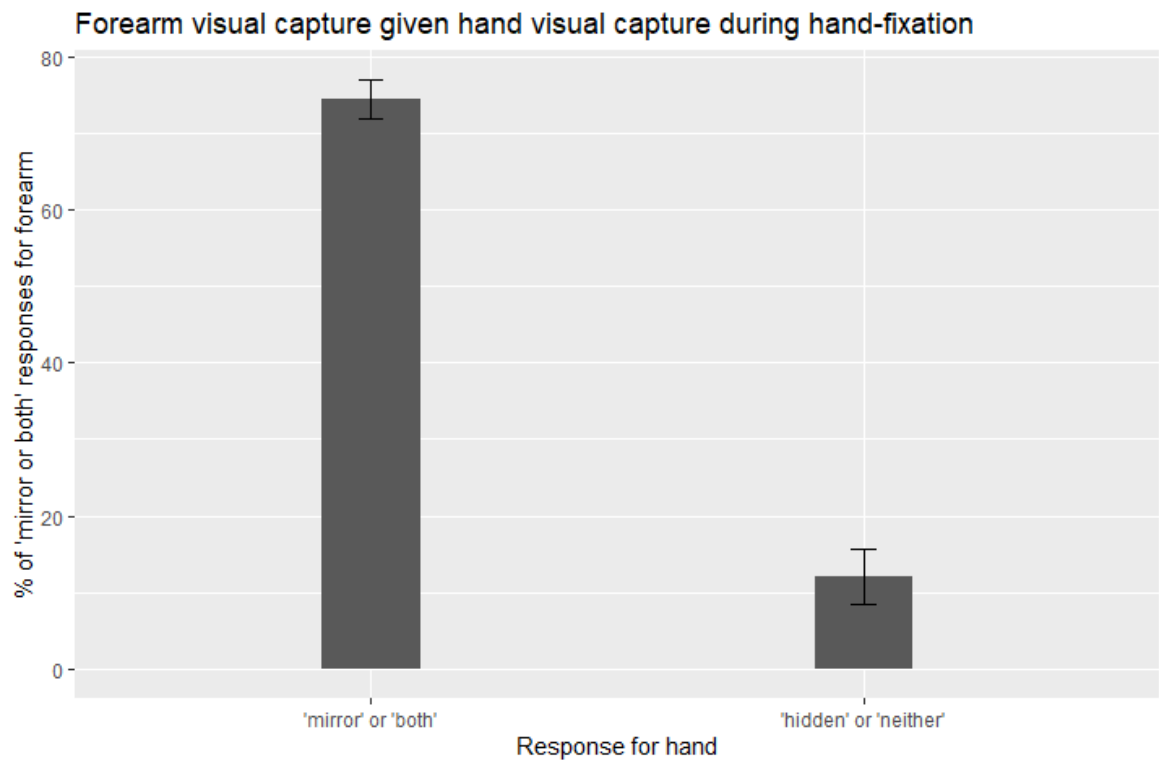


Figure 3.4: Percentage of 'mirror or both' responses for the forearm, given the presence or absence of a 'mirror' or 'both' response for the hand – when participants fixated on the hand. Error bars represent 95% confidence intervals.

3.2.3.2 Likelihood of visual capture of the un-foveated hand, given visual capture of the foveated forearm, during forearm-fixation

Similar effects were found when participants were instructed to fixate on their forearm, where participants were significantly more likely to respond with ‘mirror’ or ‘both’ for the unfoveated hand ($M = 86.20\%$, $95\% \text{ CI} = [84.06\%, 88.35\%]$) on trials where they also responded with ‘mirror’ or ‘both’ for the hand, compared to when they responded with ‘hidden’ or ‘neither’ for the hand ($M = 34.84\%$, $95\% \text{ CI} = [30.40\%, 39.28\%]$), $z = 4.827$, $p < .001$. See Figure 3.5 for plotted summary statistics.

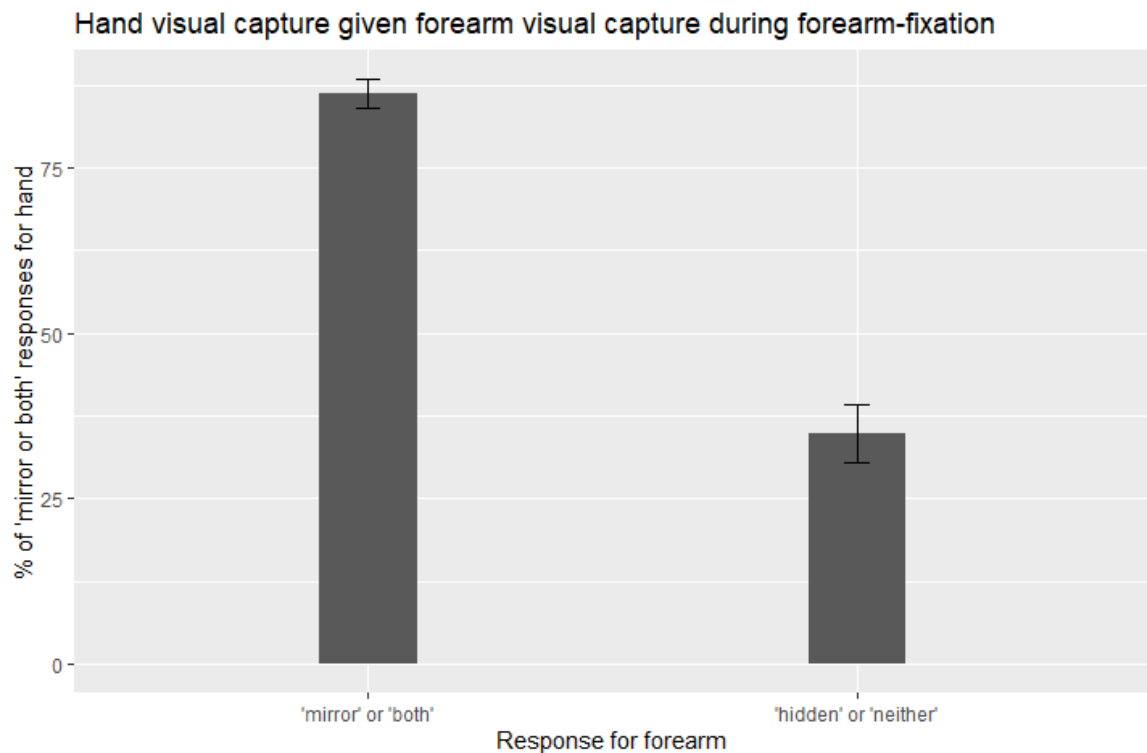


Figure 3.5: Percentage of ‘mirror or both’ responses for the hand, given the presence or absence of a ‘mirror’ or ‘both’ response for the forearm – when participants fixated on the forearm. Error bars represent 95% confidence intervals.

Taken together, these results show that when both the hand and the forearm are visible in the mirror, visual capture of either part significantly raises the likelihood of visual capture of the other adjacent part, regardless of local differences in visual acuity between the two due to selective foveation – providing further evidence for top-down influences of stored knowledge about the body (e.g. relationships between connected body parts) on multisensory integration for body perception.

3.3 Discussion

The primary goals of this experiment were to investigate the role of foveation on visual capture in the MBI – specifically, to see if the location of fixation may be a driving factor for previously observed effects of ‘fragmented’ or ‘dissociated’ visual capture of a connected hand and forearm in the mirror, and how the local, bottom-up effects of relative unimodal variance may interact with top-down effects of a stored ‘global’ body map during multisensory integration. According to optimal integration principles (Ernst and Banks, 2005), the final estimate of multisensory integration is weighted most strongly toward the unimodal input with the lowest variance. Given the higher acuity (and thus lower variance) of foveal vision compared to peripheral vision, we theorized that there would be a stronger shift toward the visual estimate for the ‘foveated’ part of the viewed arm in the mirror, compared to the unfoveated part. Thus, we predicted that when participants were synchronously stroked across their full lower arm as it is visible in the mirror, they would be more likely to experience visual capture of the part of the ‘mirror arm’ that they fixated on, compared the adjacent part in the peripheral visual field. However, we found only partial support for this theory in our study. The location of fixation did seem to have a significant effect on visual

capture of the hand, such that participants were significantly more likely to report feeling their hand ‘in the mirror’ when they fixated on the hand, compared to when they fixated on the forearm. Although there was a higher percentage of ‘mirror or both’ responses for the forearm during forearm-fixation compared to during hand-fixation, this difference was not found to be statistically significant in our binomial mixed model. These results indicate that while local differences in the unimodal variance of vision may influence local differences in multisensory integration across adjacent body parts, this relationship is not universal across different body parts.

Looking at the differences in hand and forearm visual capture for each fixation point, we found that similar to our past results, there was significantly greater hand embodiment than forearm embodiment during hand fixation. Surprisingly, we found visual capture of the hand and the forearm to be roughly equal when participants fixated on their forearm. This was in contrast to our prediction that visual capture of the forearm would be greater than that of the hand, during forearm fixation. Meanwhile, when looking specifically at instances of ‘fragmented visual capture’ of the hand or the forearm (where one was reported in the ‘mirror’ location in the absence of the other on the same trial), both hand fixation and forearm fixation significantly raised the likelihood of fragmented visual capture for the fixated part of the viewed arm. These results indicate that the location of fixation does have some effect on the observed ‘dissociation’ between visual capture of two adjacent parts of a viewed limb, even when both are stimulated equally. However, once again, the effect is not equal across body parts – while hand-fixation results in a clear dissociation between hand and forearm visual capture, forearm-fixation extinguishes this

dissociation. This may be explained by the fact that visual capture for the hand seems to be stronger overall compared to visual capture of the forearm, regardless of the location of fixation. This suggests that although visual fixation may have played a role in the observed dissociation between hand and forearm embodiment in our original experiment, it was likely not the only factor at play.

Next, we also found that during hand fixation, hand visual capture was highly predictive of forearm visual capture, as participants were far more likely to experience the forearm in the mirror on trials where they also experienced the hand in the mirror. Meanwhile, the likelihood of embodying the forearm without embodying the hand during hand fixation was less than 10%. This indicates that when fixating on the hand, the majority of instances for forearm visual capture were likely ‘carryover’ effects from hand visual capture, due to the influence of a stored, global representation of the body. This is interesting, as it suggests that the visual information of the forearm received from peripheral vision during hand fixation had very little contribution to the embodiment of the forearm - further highlighting the importance of foveal vision in embodiment. Moreover, these results provide further support to the theory that multisensory integration for body perception is modulated by top-down influences from stored knowledge about the body, such as the biomechanical relationships between adjacent body parts.

During forearm fixation, we also found forearm embodiment to be highly predictive of hand embodiment. This once again highlights the influence of the global body schema in the embodiment of adjacent parts of the mirror arm, as we see that even in a condition with very similar rates of embodiment for the hand and the

forearm, they are still more likely to occur simultaneously, than in the absence of the other.

Finally, we used the same analyses from Experiment 1 to explore the time-course of the MBI. We found similar main effects of ‘time’ in this experiment, such that visual capture was found to gradually increase over the duration of the manipulation. We also found steeper initial increases in visual capture toward the beginning of the block, compared to subsequent trials, as the main effect of time was stronger in the ‘0-10 seconds’ phase of stimulation. However, unlike in Experiment 1, we did not find any interaction effects between fixation point and time in the ‘initial’ or ‘subsequent’ time periods, indicating that the rate at which visual capture of a given part of the arm increased over time was unaffected by whether or not it was the location of fixation.

Chapter 4

GENERAL DISCUSSION

The overarching aim of this study was to explore how two specific, well-known, bottom-up factors involved in multisensory integration – cross-modal congruence and unimodal variance – may interact with top-down influences of a stored, internal representation of the body, in the construction of a coherent percept of the body in space. Although previous literature has shown clear effects of both bottom-up (Tsakiris & Haggard, 2005 – Experiments 3 & 4; Tsakiris et al., 2006) and top-down (Tsakiris & Haggard, 2005 – Experiments 1 & 2; Liu & Medina, 2017) factors on multisensory integration in body illusions, there has been limited work exploring how these two kinds of information interact to give rise to the final, integrated estimates of bodily percepts.

A previous experiment in our lab involving the MBI found that when participants were asked to report the locations of both their hand and forearm while tapping their index fingers synchronously in a mirror-box that provided a clear visual image of their connected hand and forearm in the mirror, they were significantly more likely to report visual capture of the hand, compared to the forearm. Moreover, participants often reported experiencing their hand in the ‘mirror’ location while simultaneously reporting their forearm in the ‘hidden’ location. Meanwhile, the results also hinted at the ‘global’ effects of visual capture, as the visual capture of either part

of the viewed arm in the mirror significantly predicted visual capture of the adjacent part on the same trial. However, the specific mechanism behind these effects was difficult to ascertain from this original experiment's data. Thus, this experiment provided an intriguing, but incomplete account of how 'local' multisensory integration driven by bottom-up sensory attributes may interact with 'global' integration effects driven by top-down influences of stored body knowledge.

To explore these interactions in further depth, we designed the two experiments reported in this paper to explore two potential factors that may have been responsible for the pattern of results we saw in our old experiment – (1) the location of synchronous multisensory stimulation, and the location, and (2) the location of the visual fixation point during the induction of the illusion. These two variables were uniquely suited to study the 'local' and 'global' effects of cross-modal congruence and unimodal visual variance on visual capture in the MBI, respectively. Furthermore, each of these bottom-up factors may be weighted alongside information from stored body representations, in the final integrated estimates of body position.

Our first experiment provided considerable evidence that the effects of cross-modal congruence on visual capture in the MBI are stronger locally, than globally, as participants were significantly more likely to experience visual capture for the part of the arm that received synchronous stroking, compared to the visibly adjacent, but unstimulated part. The results also showed that during isolated synchronous stroking of either part of the viewed arm, participants were significantly more likely to report visual capture of the stimulated part compared to the unstimulated part. Finally, we found evidence supporting the theory that the location of isolated synchronous

stroking on a viewed limb in the mirror may drive effects of ‘fragmented’ visual capture of the stimulated part of the viewed limb, as synchronous stroking of either part of the arm significantly raised the likelihood of participants experiencing ‘fragmented’ visual capture of the stimulated part in the absence of visual capture of the adjacent, unstimulated part of the viewed arm.

In our second experiment, we found similar results for the effect of location of visual fixation on hand and forearm visual capture – participants were significantly more likely to report visual capture for the hand when they were fixated on the hand in the mirror, than when they fixated on the forearm. However, the increased frequency of forearm visual capture during forearm-fixation compared to hand-fixation was found to be non-significant. A similar pattern was found in our next analysis, where we investigated how the location of fixation may influence the ‘dissociation’ between hand and forearm visual capture seen in our past MBI work. When comparing visual capture of the hand compared to the forearm under each fixation condition, we found that participants were significantly more likely to report visual capture of the hand compared to visual capture of the forearm when they fixated on the hand. However, when they fixated on the forearm, there was no significant difference between hand and forearm visual capture. In fact, rates of forearm and hand visual capture seem to be roughly equal during forearm fixation. Notably, when we next looked at how the location of fixation may predict the incidence of ‘fragmented’ visual capture for each part of the viewed arm, we found similar results as seen in Experiment 1, as fixation of either part of the arm significantly raised the likelihood of participants experiencing ‘fragmented’ visual capture of that part in the absence of the other part.

Taken together, these results indicate that while foveation appears to play a role in modulating local multisensory integration of different body parts in the MBI, its effect is not uniform across different parts of a viewed limb. Specifically, the effect of foveation on visual capture appears to be much stronger for the hand compared to the forearm. Nevertheless, this study is the first to show any evidence for the specific role of foveation in multisensory integration, showing that simply fixating on a part of a viewed limb can significantly raise the likelihood of visual capture for that part, as well as raise the likelihood that an individual will experience ‘fragmented’ visual capture of that part of their body in isolation from adjacent parts.

The above results from both experiments show strong evidence for the independent construction of local body schemata for the hand and the forearm, based on the differences in multisensory information being integrated for each. However, we were also curious as to whether we would see any ‘global’ effects of a stored ‘offline’ body representation on visual capture of each part of the mirror arm. Interestingly, we found that in each condition of both experiments, ‘local’ visual capture of one part of the viewed arm strongly predicted visual capture of the other part on the same trial. For instance, during hand-stroking in Experiment 1, participants were significantly more likely to report visual capture of the unstimulated forearm on trials where they also reported visual capture of the hand, compared to trials where they did not. Similar effects were seen for hand visual capture during forearm-stroking, as well as for visual capture of the unfoveated parts of the arm in Experiment 2. This shows that although the local body schemata for the hand and the forearm are constructed independently to some extent, they are still influenced by each other, as visual capture of the

stimulated/foveated of the arm appears to increase the likelihood of visual for adjacent, unstimulated/unfoveated part. These effects suggest a mechanism that combines the independent integration of multisensory information for a given body part's location, with top-down information from a global body representation that stores the relationships between each body part.

Finally, we used both experiments to explore a secondary question that has seen limited exploration in the literature – that of the ‘time-course’ of visual capture in the MBI. Given the limited information available regarding this question, we were interested in characterizing the overall nature of this time-course, as well as look at how the time-course may vary for ‘local’ and ‘global’ effects of visual capture. Our exploratory analyses in both experiments revealed that visual capture in the MBI shows a gradual increase over time as the duration of stimulation increases, and that this increase is characterized by a sharp initial increase in the first 10 seconds of stimulation, followed by smaller increases over time as stimulation continues and the illusion approaches maximum effect for a condition. This finding was consistent with previous work done on the time-course of proprioceptive drift in the RHI (Rohde et al., 2011), as well as the existing theory that cross-modal congruence in body illusions enhances multisensory integration via an accumulation of evidence over time, as continuous multisensory stimulation is applied. Interestingly, we also found a significant interaction between time and ‘stroking location’ in Experiment 1, such that the initial rate of increase was greater for the stimulated part of the arm compared to the unstimulated part. This suggests that ‘local’ effects of visual capture may accumulate evidence faster than ‘global’ effects in the initial phases of synchronous

multisensory stimulation. However, this effect was not replicated in Experiment 2. It is difficult to determine why we did not see the same interaction effects in both experiments, though it may have something to do with the inherent differences in the effects of foveation and isolated synchronous stroking observed in our other analyses.

In conclusion, these two experiments provide evidence for the roles of two specific factors - the location of dynamic, congruent multisensory stimulation, and the location of visual fixation - in the embodiment of individual parts of a mirror limb in the MBI. It is important to note, however, that neither of the factors studied in the two above experiments were able to completely explain the strong dissociation between hand and forearm embodiment that we saw in the original experiment. Instead, we saw that there was significantly more hand embodiment than forearm embodiment for both hand-stroking and hand-fixation - as seen in the original experiment - whereas the differences between hand and forearm embodiment for both forearm-stroking and forearm-fixation were non-significant. One possibility is that despite the modulating effects of these two factors, there may be an additional, inherent tendency to embody an illusory hand compared to an illusory forearm, regardless of external factors. This may be the case given that we use our hands to interact with our environment much more than we use our forearms. We know that the integration of sensory signals from different modalities is influenced by a 'prior' probability for those signals to share a common cause, which is informed by the learned statistical relationships between stimuli that are likely to co-occur based on past sensory experiences. It may be that due to a larger number of sensory events experienced by the hand, compared to the forearm, the brain develops stronger priors for integrating congruent multisensory

information received by the hand, compared to the forearm - and this, stronger embodiment of an illusory hand, compared to an illusory forearm.

In summary, these experiments also provide strong evidence for the idea that the location estimates of different body parts in space are computed independently, to an extent, through the integration of the specific sets of multisensory information available for each part. However, there is also evidence to show that these estimates are influenced by a global body representation in a top-down manner, as the embodiment of part of an illusory limb raises the likelihood of the adjacent part also being embodied. Together, these experiments point to a framework of multisensory integration for body perception that involves an interaction of top-down influences of a stored representation of the body, with different bottom-up sensory attributes of the signals being integrated, such as cross-modal congruence and unimodal variance – each of which are weighted to different degrees.

REFERENCES

- Baylor, D., Matthews, G., & Yau, K. W. (1980). Two components of electrical dark noise in toad retinal rod outer segments. *The Journal of physiology*, 309(1), 591-621.
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, 391(6669), 756-756.
- Brang, D., McGeech, P. D., & Ramachandran, V. S. (2008). Apotemnophilia: a neurological disorder. *Neuroreport*, 19(13), 1305-1306.
- Chancel, M., Ehrsson, H. H., & Ma, W. J. (2022). Uncertainty-based inference of a common cause for body ownership. *Elife*, 11, e77221.
- de Vignemont, F., Pitron, V., & Alsmith, A. J. (2021). What is the body schema?. *Body Schema and Body Image: New Directions*, p.9.
- Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870), 429-433.
- Holmes, N. P., & Spence, C. (2005). Visual bias of unseen hand position with a mirror: spatial and temporal factors. *Experimental Brain Research*, 166(3), 489-497.
- Kalckert, A., & Ehrsson, H. H. (2017). The onset time of the ownership sensation in the moving rubber hand illusion. *Frontiers in psychology*, 8, 344.
- Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS one*, 2(9), e943.
- Lane, T., Yeh, S. L., Tseng, P., & Chang, A. Y. (2017). Timing disownership experiences in the rubber hand illusion. *Cognitive research: principles and implications*, 2(1), 1-14.
- Laughlin SB, Lillywhite PG. Intrinsic noise in locust photoreceptors. *J Physiol* 332: 25– 45, 1982.

- Levi, D. M., & Klein, S. A. (1986). Sampling in spatial vision. *Nature*, 320(6060), 360-362.
- Lewald, J., Ehrenstein, W. H., & Guski, R. (2001). Spatio-temporal constraints for auditory–visual integration. *Behavioural brain research*, 121(1-2), 69-79.
- Lloyd, D. M. (2007). Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain and cognition*, 64(1), 104-109.
- Liu, Y., & Medina, J. (2017). Influence of the body schema on multisensory integration: Evidence from the mirror box illusion. *Scientific reports*, 7(1), 5060.
- Maravita, A., Spence, C., & Driver, J. (2003). Multisensory integration and the body schema: close to hand and within reach. *Current biology*, 13(13), R531-R539.
- Medina, J., Khurana, P., & Coslett, H. B. (2015). The influence of embodiment on multisensory integration using the mirror box illusion. *Consciousness and cognition*, 37, 71-82.
- Petkova, V. I., Björnsdotter, M., Gentile, G., Jonsson, T., Li, T. Q., & Ehrsson, H. H. (2011). From part-to whole-body ownership in the multisensory brain. *Current Biology*, 21(13), 1118-1122.
- Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs. The DO Hebb lecture. *Brain: a journal of neurology*, 121(9), 1603-1630.
- Ramachandran, V. S., & Rogers-Ramachandran, D. (1996). Synaesthesia in phantom limbs induced with mirrors. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1369), 377-386.
- Rentschler, I., & Treutwein, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature*, 313(6000), 308-310.
- Rieke, F., & Baylor, D. A. (2000). Origin and functional impact of dark noise in retinal cones. *Neuron*, 26(1), 181-186.
- Royal, D. W., Carriere, B. N., & Wallace, M. T. (2009). Spatiotemporal architecture of cortical receptive fields and its impact on multisensory interactions. *Experimental brain research*, 198, 127-136.

- Tsakiris, M. (2010). My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia*, 48(3), 703-712.
- Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology: Human perception and performance*, 31(1), 80.
- Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and cognition*, 15(2), 423-432.
- Welch, R. B., Widawski, M. H., Harrington, J., & Warren, D. H. (1979). An examination of the relationship between visual capture and prism adaptation. *Perception & Psychophysics*, 25, 126-132.

Appendix

IRB/HUMAN SUBJECTS APPROVAL



Institutional Review Board
210H Haulikien Hall
Newark, DE 19716
Phone: 302-831-2137
Fax: 302-831-2828

DATE: July 24, 2020
TO: Jared Medina, Ph.D.
FROM: University of Delaware IRB
STUDY TITLE: [1605575-1] Understanding Crossmodal Processing
SUBMISSION TYPE: New Project
ACTION: DETERMINATION OF EXEMPT STATUS
EFFECTIVE DATE: July 24, 2020
REVIEW CATEGORY: Exemption category # (3)

Thank you for your New Project submission to the University of Delaware Institutional Review Board (UD IRB). According to the pertinent regulations, the UD IRB has determined this project is EXEMPT from most federal policy requirements for the protection of human subjects. The privacy of subjects and the confidentiality of participants must be safeguarded as prescribed in the reviewed protocol form.

This exempt determination is valid for the research study as described by the documents in this submission. Proposed revisions to previously approved procedures and documents that may affect this exempt determination must be reviewed and approved by this office prior to initiation. The UD amendment form must be used to request the review of changes that may substantially change the study design or data collected.

Unanticipated problems and serious adverse events involving risk to participants must be reported to this office in a timely fashion according with the UD requirements for reportable events.

In-person research interaction with subjects cannot begin until the UD moratorium in response to the declaration of national emergency related to the COVID-19 pandemic is lifted. Please continue to reference <https://research.udel.edu/coronavirus> for the most up-to-date recommendations.

A copy of this correspondence will be kept on file by our office. If you have any questions, please contact the UD IRB Office at (302) 831-2137 or via email at hsrb-research@udel.edu. Please include the study title and reference number in all correspondence with this office.

INSTITUTIONAL REVIEW BOARD

www.udel.edu