LEABHARLANN CHOLÁISTE NA TRÍONÓIDE, BAILE ÁTHA CLIATH | TRINITY COLLEGE LIBRARY DUBLIN
Ollscoil Átha Cliath | The University of Dublin

Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Temporal and Spatial factors affecting the cross-modal integration of object information.

by

Andrew Thomas Woods

A dissertation submitted for the degree of Doctor of Philosophy of the University of Dublin, Trinity College, Dublin 2, Ireland.

This research was carried out in the Department of Psychology.

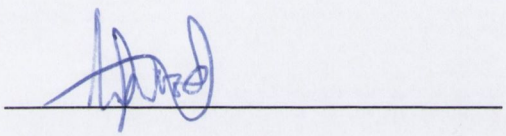
April 2005.

TRINITY COLLEGE
10 JUL 2006
LIBRARY DUBLIN
THESIS
7936

Declaration

I declare that this work has not been submitted previously as an exercise for a degree at this or any other university and that it is entirely my own work. The Trinity College Library may lend or copy this thesis without restriction.

Signed



Andrew Thomas Woods.

Department of Psychology
University of Dublin, Trinity College
Dublin 2, Ireland

Tel: +353 1 608 1886
Fax: +353 1 671 2006

Department of Psychology
Ethics Committee

21/04/2005

Dear Andy Woods,

The Department of Psychology Ethics Committee met to consider your application entitled

"Factors that influence the perception of felt objects in the presence of vision"

I am pleased to inform you that the Committee has approved your application.

Yours sincerely,

K Tierney

Kevin Tierney (Ph.D.)
Chairperson,
Department of Psychology Ethics Committee

Acknowledgements

I would first like to thank Fiona Newell for the large amount of time, effort and patience which must have gone into supervising me. I would not have been able to complete this thesis if it were not for Fiona, and have learnt a huge amount from her in the process. I would also like to thank Sile O'Modhrain for also supervising me and for providing insights into the more technical aspects of my thesis.

Heartfelt thanks go to all my friends and colleagues during my time in Trinity. Elaine, for being there for me, and helping me without question on anything imaginable; Jose, Andrew and Melissa for their friendship, and to the many who I have not the space to mention. Thank you also to the lab group for their constructive advise during all stages of my thesis. Especial thanks goes to Sarah, for being a great office mate, for reading drafts, and for being able to put up with my more creative pursuits.

Thank you also to Mum, Dad and Simon, whom I contacted at the slightest opportunity to bemoan my life woes, or to generally just be a nuisance. Mum also read a number of early drafts of my thesis, for which I am very grateful.

Publications resulting from the present work:

Articles:

Woods, A.T. & Newell, F.N. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *Journal of Physiology Paris.* 98, 147-59.

Woods, A. T., O'Modhrain, S. & Fiona N. Newell (2004). The Effect of Temporal Delay and Spatial Differences on Cross-Modal Object Recognition *Cognitive, Affective & Behavioral Neuroscience*, 4, 260–269.

Published Abstracts and Short Articles:

Woods, A.T., O'Modhrain, S. and Newell, F.N. (2005). Spatial factors involved in the multisensory integration of objects. Poster presented at the 2005 WorldHaptics conference, Pisa, Italy.

Woods, A.T. & Newell, F.N. (2004). Information sharing across the modalities is affected by horizontal spatial disparity. Poster presented at Touch, Temperature, Pain & Pleasure, University of Manchester, Manchester, England.

Woods, A.T. & Newell, F.N. (2004). Distance and direction from the observer affect the cross-modal perception of real objects. Poster presented at 5th Annual Meeting of the International Multisensory Research Forum, Barcelona, Spain.

Woods, A.T., O'Modhrain, S., & Newell, F. (June, 2003). Temporal factors and the sharing of cross-modal information. Poster presented at 4th Annual Meeting of the International Multisensory Research Forum, McMaster University, Hamilton, Ontario, Canada.

Woods, A.T., O'Modhrain, S. and Newell, F.N. (2003). Temporal factors in the sharing of cross-modal information. Proceedings of the 2003 EuroHaptics conference, Trinity College, Dublin, Ireland.

Thesis Summary

This thesis is concerned with how temporal and spatial factors influenced the cross-modal integration of object representations across modalities. Specifically tested was whether temporal synchrony and spatial co-location are crucial for the integration of visual and haptic object information. It was hypothesised that there would be mandatory integration when objects were presented to modalities in temporal synchrony and in the same location in space.

The results of Experiments 1, 2 outlined in Chapter 2, suggested that cross-modal object recognition was found to be best within a temporal window that included simultaneous or rapidly successive presentation of stimuli across the modalities. As the percepts of temporally co-incident stimuli did not interfere with each other as was predicted if they were mandatorily integrated (Experiment 3), this suggested also that vision and touch maintained discrete representations allowing for efficient discrimination. However, one reason why integration did not occur was because the objects may have needed to be spatially co-located.

Experiments outlined in Chapter 3 attempted to address this. Evidence that temporally synchronous and co-located objects were mandatorily cross-modally integrated was not found (Experiment 4). Increasing the assumption of unity of the stimulus-display also did not promote cross-modal integration (Experiment 5). It was thought that a difference in the perceived size of visual and haptic objects may have prevented these object percepts from being integrated. However, the results of Experiment 6 suggested that the objects were perceived as equivalent in size. A difference in visual and haptic encoding or sensorimotor patterns during encoding may have made the object percepts different in appearance such that they were not integrated. In Experiment 7, haptic encoding was equated with visual encoding

(haptic encoding was made more optimal), yet there was still yet no evidence that co-located objects were integrated. There was a concern that, the percepts of visual and haptic objects may indeed have been integrated, but not at the veridical spatial co-location. In Experiment 8, a small rightward bias in the location at which visual and haptic objects were co-located was identified, which suggested that integration may have occurred when haptic objects were located to the right. The results of Experiment 4, 5 and 7 suggested that evidence that the percepts of rightward located haptic objects were integrated with visual percepts was however, not consistent.

It was important to determine whether haptic object percepts underwent a location-specific distortion as this could also explain the partial evidence for the percepts of haptic objects located to the right being integrated with visual percepts. The results of the experiments in Chapter 4 suggested that haptic objects were indeed distorted depending upon their location in space. Thus, there was no direct evidence that the percepts of visual and haptic objects which were temporally synchronous and spatially co-located were integrated in the preceding chapters.

In Chapter 5, the findings from this thesis were discussed in terms of their importance for the different models of integration. It was argued that top-down effects such as not seeing the hand when exploring objects prevented object percepts from being integrated. Finally, the contributions that the findings of this thesis make to the literature are highlighted and suggestions are made about potential avenues of future research.

TABLE OF CONTENTS

CHAPTER 1 : GENERAL INTRODUCTION **1**

1.1. Behavioural investigations of visual and haptic perception *3*

1.1.1 Visual object recognition *3*

1.1.2 Haptic object recognition *4*

1.1.3 Similarities and differences between vision and haptics *7*

1.2. Cross-modal integration in the cortex *10*

1.2.1 Multisensory integration at the neuron level *10*

1.2.2 Multisensory integration in the cortex *11*

1.3 Factors affecting cross-modal integration *15*

1.3.1 Task factors *15*

1.3.2 Encoding factors *16*

1.3.3 Memory effects *18*

1.4 How is cross-modal integration achieved? *19*

1.4.1 Models of integration *20*

1.4.2 Criteria for cross-modal integration *21*

1.4.3 Temporal synchrony and its role in cross-modal integration *23*

1.4.4 Spatial co-location *24*

1.4.5 A methodology for exploring cross-modal integration *26*

1.5 Conclusion *27*

CHAPTER 2 : THE EFFECT OF TEMPORAL DELAY AND SPATIAL DIFFERENCES ON CROSS-MODAL OBJECT RECOGNITION **28**

Abstract **28**

2.1 Introduction **29**

2.2 Experiment 1 ..... **34**

2.2.1 Method *34*

2.2.1.1 Participants 34

2.2.1.2 Stimuli 35

2.2.1.3 Apparatus 36

2.2.1.4 Design 37

2.2.1.5 Procedure 39

2.2.2 Results 40

2.2.3 Discussion 42

2.3 Experiment 2 ..... **43**

2.3.1 Method 44

2.3.1.1 Participants 44

2.3.1.2 Stimuli and Apparatus 44

2.3.1.3 Design 45

2.3.1.4 Procedure 45

2.3.2 Results 46

2.3.3 Discussion 50

2.4 Experiment 3 ..... **51**

2.4.1 Method 51

2.4.1.1 Participants 51

2.4.1.2 Stimuli and Apparatus 51

2.4.1.3 Design 52

2.4.1.4 Procedure 52

2.4.2 Results 53

2.4.3 Discussion **55**

2.5 General Discussion **56**

CHAPTER 3 : THE INVOLVEMENT OF SPATIAL FACTORS ON OBJECT INTEGRATION **60**

Abstract **60**

3.1 Introduction **61**

3.2 Experiment 4 ..... **63**

3.2.1 Method *64*

3.2.1.1 Participants *64*

3.2.1.2 Stimuli *65*

3.2.1.3 Apparatus *65*

3.2.1.4 Design *67*

3.2.1.5 Procedure *68*

3.2.2 Results *69*

3.2.3 Discussion *72*

3.3 Experiment 5 ..... **73**

3.3.1 Method *76*

3.3.1.1 Participants *76*

3.3.1.2 Stimuli and Apparatus *76*

3.3.1.3 Design *77*

3.3.1.4 Procedure *78*

3.3.2 Results *78*

3.3.3 Discussion *79*

3.4 Experiment 6 ..... **80**

3.4.1 Method *82*

3.4.1.1 Participants *82*

3.4.1.2 Stimuli and Apparatus *82*

3.4.1.3 Design *83*

3.4.1.4 Procedure *83*

3.4.2 Results *83*

3.4.3 Discussion *84*

3.5 Experiment 7 ..... **85**
3.5.1 Method *86*
3.5.1.1 Participants *86*
3.5.1.2 Stimuli and Apparatus *87*
3.5.1.3 Design and Procedure *87*
3.5.2 Results *89*
3.5.3 Discussion *89*
3.6 Experiment 8 ..... **91**
3.6.1 Method *93*
3.6.1.2 Participants *93*
3.6.1.3 Stimuli and Apparatus *93*
3.6.1.4 Design *94*
3.6.1.5 Procedure *95*
3.6.2 Results *95*
3.6.3 Discussion *96*
3.7 General Discussion **97**
3.7.1 Reasons for the lack of integration effect *99*
3.7.2 Reasons why integration effects were not detected *101*
CHAPTER 4 : LOCATION-SPECIFIC HAPTIC-OBJECT DISTORTION .. **103**
Abstract **103**
4.1 Introduction **104**
4.2 Experiment 9 ..... **109**
4.2.1 Method *109*
4.2.1.1 Participants *109*
4.2.1.2 Stimuli *110*
4.2.1.3 Apparatus *110*
4.2.1.4 Design *112*

4.2.1.5 Procedure *112*

4.2.2 Results *113*

4.2.3 Discussion *115*

4.3 Experiment 10 ..... **116**

4.3.1 Method *117*

4.3.1.1 Participants *117*

4.3.1.2 Stimuli and apparatus *117*

4.3.1.3 Design and Procedure *118*

4.3.2 Results *119*

4.3.3 Discussion *122*

4.4 General Discussion **123**

4.4.1 Reasons why additional haptic object distortion was not detected *124*

4.4.2 Experimental validity *125*

CHAPTER 5 : GENERAL DISCUSSION **127**

Abstract **127**

5.1 Summary of findings *127*

5.1.1 Type of memory representations *128*

5.1.2 Temporal factors *128*

5.1.3 Spatial factors *129*

5.2 A model of cross-modal integration *132*

5.2.1 Automatic visual-recoding of haptic object percepts *132*

5.2.2 Amodal or multisensory integration of visual and haptic inputs *133*

5.3 Top-down factors affecting cross-modal integration *134*

5.3.1 The visual hand hypothesis *135*

5.3.2 Other evidence for the role of seeing the hand in multisensory integration *136*

5.3.3 Suggestions for future work *137*

5.4 Other potential factors affecting cross-modal integration. *138*

5.5 Limitations of the studies *139*

5.5.1 The role of the position of co-location *139*

5.5.2 Visual-haptic ventriloquist effects *140*

5.5.3 Reference frames *141*

5.5.5 An alternative theory *142*

5.5.5 Ecological validity of the theory *143*

5.6 Conclusions and contributions of this work to the literature *143*

REFERENCES **146**

LIST OF FIGURES

Figure 1. An illustration of how objects in the environment are cross-modally integrated. 2

Figure 2. An illustration of visual and haptic modality encoding biases 9

Figure 3. An illustration of an areal to neuronal continuum of multisensory integration. 13

Figure 4. An illustration of the stimuli set 35

Figure 5. An illustration of the apparatus used in Experiment 1 37

Figure 6. Plots showing percentage correct rejection scores for each spatial change and modality across the different ISIs in Experiment 2. 49

Figure An illustration of the apparatus used in Experiment 4 67

Figure 8. Plot depicting average correct matching performance scores for each location at which haptic objects were presented 70

Figure 9. An illustration of the apparatus used in Experiment 5 77

Figure 10. Plot depicting the 'same size' response scores across conditions in Experiment 6 84

Figure 11. An illustration of the apparatus used in Experiment 7 88

Figure 12. An illustration of the apparatus used in Experiment 8 94

Figure 13. An illustration of the apparatus used in Experiment 9 111

Figure 14. A illustration of the apparatus used in Experiment 10 118

Figure 15. Plot depicting the haptic object shape estimation error at each location for the object's x-axis arms and y-axis arms in Experiment 10 120

LIST OF TABLES

Table 1. *Table showing mean percentage correct responses (hit and correct rejection scores) and mean response times across the different ISIs in each order of modality in Experiment 1.* 41

Table 2. *Table showing mean percentage correct responses and mean RTs across the different ISIs in each modality in Experiment 2* 47

Table 3. *Table showing mean percentage correct responses and mean RTs across the different presentation order conditions in Experiment 3* 54

Table 4. *Mean and standard error scores for HITS, CR and RT for each level of POSITION in Experiment 4* 71

Table 5. *Mean and standard error scores for HITS, CR and RT for each level of POSITION in Experiment 5* 79

Table 6. *Mean and standard error scores for HITS, CR and RT for each level of POSITION in Experiment 7* 90

Table 7. *Mean error and standard error of measurements of the x-axis and y-axis of the haptic objects relative to the actual measurements for the objects in each position in Experiment 9* 114

TERMINOLOGY

Amodal memory: relating to how information is stored in the brain. Memory can be termed amodal when information pertaining to an object or event ia stored collectively, irrespective of the modalties that encoded the information.

Encoding: the process by which sensory receptors translate a detected physical event in the environment into a form that may be conveyed through the central nervous system.

Cross-modal integration: An object in the environment emits a range of information that can be encoded by different sensory receptors and transmitted through the central nervous system. Cross-modal integration occurs when information encoded by sensory receptors from different modalities combines together in the brain, usually forming a more accurate and more detailed description of the perceived object than the description formed through a single modality (Ernst & Bülthoff, 2004). According to Meredith (2002, p.33), this process is an "an emergent property of the brain."

Multisensory Integration: the neurological analogue of cross-modal integration. Information encoded by different sensory modalities may activate the same neuron leading to a neuronal response different from that would be expected if the modalities activated the neuron individually (see Stein & Meredith, 1993)

Percept: the collection of knowledge an observer has related to a perceived object.

Representation: see Percept.

Recoding: There is a growing debate in the cross-modal perception literature about how information perceived across modalities is stored. One account suggests that this information is stored amodally (see above). Another account suggests that information is stored independently by each modality (Reed, Shoham, & Halgren, 2004) and for information to be shared by the modalities, it may need to be translated (see Zhang, Weisser, Stilla, Prather, & Sathian, 2004). This recoding process may incur a cost in terms of cross-modal performance (Newell, Ernst, Tjan, & Bülthoff, 2001).

Chapter 1 formed the basis of a peer-reviewed paper entitled: "Visual, haptic and cross-modal recognition of objects and scenes", which was recently published in the Journal of Physiology (Paris).

Chapter 2 was peer-reviewed and published in a special issue of the journal Cognitive, Affective and Behavioural Neuroscience, as a paper entitled: "The effect of temporal delay and spatial differences on cross-modal object recognition."

Woods, A. T., & Newell, F. N. (2004). Visual, haptic and cross-modal recognition of objects and scenes. *Journal of Physiology (Paris), 98*, 147-159.

Woods, A. T., O'Modhrain, S., & Newell, F. N. (2004). The effect of temporal delay and spatial differences on cross-modal object recognition. *Cognitive, Affective and Behavioral Neuroscience, 4*, 260-269.

CHAPTER 1: GENERAL INTRODUCTION

Recognising an object for the first time is a complicated process, yet it is achieved with surprising accuracy. In visual processing, saccades scan across an object, analysing colour, size and shape. In haptics, the hand can feel the surface texture, and encodes tiny bumps and groves whilst simultaneously calculating how much pressure is needed to stop the object from falling to the floor. What is felt is combined with what is seen and a search through memory for translucency, coldness, smoothness and a hollow, cylindrical shape reveals you are holding a glass. Both the visual and haptic systems together can provide clues to the identity of the object; both perceive a number of similar and dissimilar object features, and through combining this information the object can be successfully identified. On its own, vision could have deduced that it was a glass, but would not have provided the information necessary for the prevention of drinking scalding coffee (unless just in time you noticed the steam rising off the liquid). Similarly touch could also have identified the object as a glass, but without a great deal of care, it would not forewarn you about a nasty jagged edge. Thus the modalities work together creating a rich, cross-modal representation of the glass and its contents.

This brief anecdote demonstrates the co-dependence between vision and touch in recognising objects. For the purpose of object recognition, both vision and haptics can be considered as image-based, in that both modalities can acquire shape information for recognition, albeit using different means. Depending on the nature of the goal, vision can dominate touch, touch can dominate vision, or if the circumstances are right, they both contribute to the percept. The modalities both encode similar aspects of an object by using different means, and also each encode a variety of modality specific information; an emergent property of the brain is the integration of this information (see

Figure 1) into a richer representation of that object in memory (for a review of the mechanics behind touch the reader is directed to Meredith, 2002), that is both of greater accuracy and of more detail than that would be formed from a single modality (Ernst & Bülthoff, 2004).

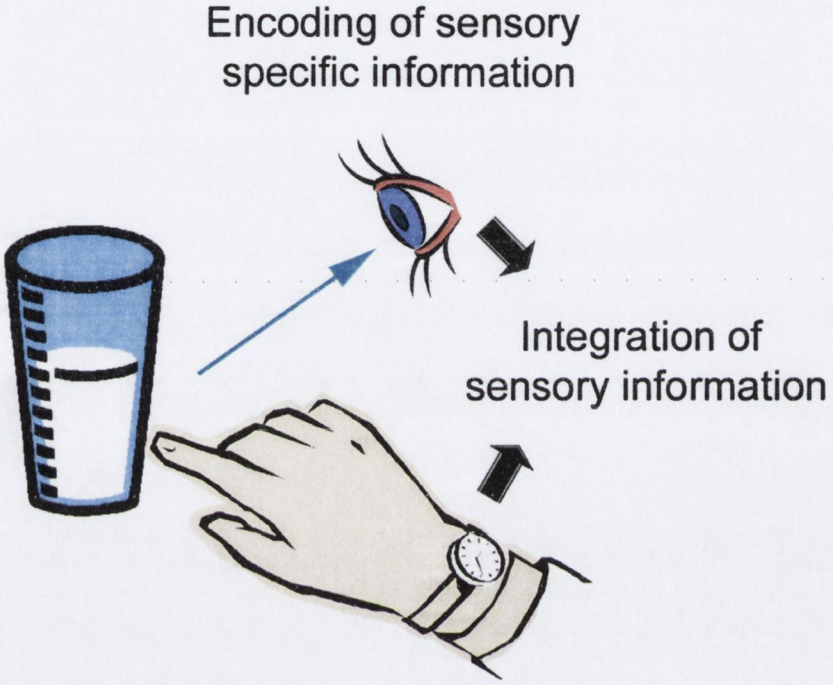


Figure 1. A stimulus in the environment emits a range of information that can be measured by the different sensory receptors of the central nervous system. In the diagram, photons bounce off an object, e.g., glass, and into the eye allowing it to be seen. The surface of skin on the finger deforms when it comes into contact with the glass, revealing a range of different object properties. These sensory specific codes may then be combined to form a rich representation of the object. The photons and resultant skin deformations do not interact with each other in the environment, and it is only after their successful encoding by the nervous system that they may be combined and a single, transparent, cold glass is perceived. Thus "the resultant forms of processing represent an emergent property of the brain" (Meredith, 2002 p. 33).

But how does the brain know to combine this information? It would be inefficient if the brain wrongly combined information about two *different* objects together, and likewise to treat information regarding a single object encoded across the modalities as multiple objects. There must therefore be some criteria for when information from across the modalities is combined. Insights may be found in observing natural cross-modal object recognition. Generally an object in the environment is observed

simultaneously by the senses. Thus it may be sufficient that object information encoded at the same time across the senses be considered as arising from a single object. It is also the case that an object will also generally be perceived to be in the same location by the modalities, so co-location, or a combination of co-location and simultaneous encoding, may be the requisites for the cross-modal integration of object information.

This chapter reviews the literature on cross-modal integration between touch and vision, with reference to the above factors. To start, a recent literature on processing within these two modalities will be reviewed separately, defining their similarities and differences and how they function to encode information. Then integration will be discussed at the level of the neuron, followed by candidate cortical areas involved in cross-modal object recognition. Factors influencing integration, and potential mechanisms which may be responsible for the combination of information, will lastly be covered.

1.1. Behavioural investigations of visual and haptic perception

1.1.1 Visual object recognition

An outstanding achievement of human vision is the capacity for rapid and seemingly effortless recognition of objects. Viewing the environment, anything of interest initiates a saccade, moving the projected image of the object from the periphery of the retina to the fovea. The object is now the centre of attention and accordingly is analysed with the most sensitive area of the eye. Depth cues are used to deduce the object's features, and texture and colouring give clues to material it is made of, or, for example, whether it is edible. Vision takes very little time to extract information about the object, and is generally labelled as being able to simultaneously analyse a large number of object features (Biederman, Rabinowitz, Glass, & Stacy, 1974; but for a

review see Tarr & Bülthoff, 1995). Vision must overcome the problem of object constancy to be able to recognise objects in the environment. Typically, a known object may be largely recognisable from any viewing angle, at a variety of distances, and under a range of different viewing conditions (e.g., poor lighting). The object may still be recognisable when it is partially hidden from view, or even when its view is distorted, as would happen for example by the wearing of unsuitable glasses. In short, the visual system is able to overcome the problem of object constancy as it is able to recognise the object from a potentially large variety of different images which may be projected onto the retina– an amazing achievement of the perceptual system (see Edelman & Bulthoff, 1992, for a more in depth discussion on object constancy).

1.1.2 Haptic object recognition

It is naturally presumed that touch is a slow, inadequate system for exploring the world and research on passive touch tends to confirm this notion. However, when we are allowed to explore an object freely with our hands, to explore it actively, suddenly touch seems so much more competent. Gibson (1962) offers some insights to the feats of touch: even though a single object which is being manipulated by hand stimulates many different receptors across five fingers and a palm, it is still viewed as a single object, as being stable and it appears to have a uniform surface stiffness. Reed (1988 p. 199) expands upon Gibson's ideas stating that "*a single object will generate myriads of different 'proximal stimuli', depending on the intention and activities of the perceiver*". The brain must integrate a great variety of diverse information from across the five fingers and palm; it also must combine the information from four receptor types (Merkel disks, Meissner Corpusles, Ruffini endings and Pacinian Corpuscles), each type detecting differing stimulations. From this, an adequate representation of the held object is formed (for a review of the mechanics behind touch the reader is directed to

Craig & Rollman, 1999). As with vision, touch may also recognise an object from a potentially huge number of 'viewing positions' in which information about the object is encoded, demonstrating that touch must also overcome object constancy. The well-known Aristotle illusion nicely demonstrates that during normal operation without vision, object constancy is achieved by touch (McKeon, 1941). When an object (e.g., a pencil) is pushed between two adjacent fingers one object is perceived, with two sensory impressions made by the two opposite sides of the object. The information from touch therefore specifies one object. But if that same pair of fingers is crossed so that non-adjacent areas of the skin are opposed (irregular operation), then the illusion of feeling two objects is often perceived.

In one of the classical experiments on touch, Gibson (Gibson, 1962) asked participants to identify the shape of different biscuit cutters that were either pressed into the persons palm or were allowed to be freely explored. Gibson found that active touch achieved a 95% success rate at identifying cutter shapes while a passive imprint of the cutter upon the person's skin only achieved 29% accuracy (chance being at 16.7%). In a similar experiment, Heller (1984) limited the time participants had to explore objects through touch and found that to match the performance of 5 seconds of active touch, participants needed over 30 seconds of passive stimulation. Along similar lines, Klatzky, Lederman & Metzger (1985) asked people to identify one hundred common objects through active touch and found that almost all of them could be identified within a few seconds. Interestingly, Jansson & Monaci found that participants took longer and were worse at identifying objects if active touch was constrained by attaching plastic sheaths to the fingers (Jansson & Monaci, 2005). Thus, it seems that passive touch or restricted active touch typifies our belief that touch is a slow, inadequate system for exploring objects. Active touch on the other hand represents the

modality's true ability, being both a fast and an efficient means of extracting object information.

Active touch, which is also known as Haptics (from the Greek word 'haptikos' meaning 'to touch'), consists of all the elements of passive touch (mechanoreceptor sensation) plus information from the kinaesthetic system and other higher order cognitive processes. An example of such a higher order cognitive process is the purposeful direction of suitable hand movements, which allow the most sensitive part of the hand to come into contact with the object property being explored. Lederman & Klatzky (e.g., 1993; 1987) observed that participants used highly stereotypical movement patterns when they explored objects, and through a series of studies these were categorised and explored. The movements were termed 'Exploratory Procedures' (EP's), with each EP specific to the object property being explored by the observer. For example, 'lateral motion' is used to explore an objects surface texture whilst 'contour following', a more time consuming EP, is used to explore an objects edges in detail for object recognition.

Lederman & Klatzky identified eight EP's, which differ according to the type of hand interaction with an unseen object. Each EP can be differentiated according to four parameters: movement type (static or dynamic); direction of force applied (normal or tangential); object region explored (edges or the centre of the object) and position relative to workspace (freely held or supported by a surface). Furthermore, EP's differ according to the range of object properties they can explore, their degree of proficiency and measurement precision of these properties, and their speed of data acquisition. EP compatibility may arise when two or more EP's share some of the same parameter values. When this occurs, exploring an object with an EP will reveal object property information pertaining to the other EP's it is compatible with, but perhaps with less

accuracy and speed. Accordingly, EP's differ according to the number of other EP's they are compatible with. This range of compatibilities acts to further divide EP's so that they have either a general or a specific function. When confronted with an object to be recognised, a person will first employ a general EP that is fast and provides adequate information across multiple dimensions of the object. Information deduced by the general EP will then be used to decide which of the more task specific EP's to use for subsequent explorations, in order to test hypotheses about the object or to define it further. Thus, a cycle of EP selection and encoding repeats until the object is identified or sufficient knowledge has been extracted from an unknown object. Lederman & Klatzky (1993) termed this as a two-process account of haptic exploration.

Vision and touch seem to be very different perceptual systems. Vision can scan over and across objects quickly and with ease, whilst touch must initiate a series of EP's to gather the analogous information. Both modalities are able to recognise objects from a wide variety of information pertaining to the object; they both overcome the problem of object constancy. The next section will compare vision and touch, considering their similarities and differences.

1.1.3 Similarities and differences between vision and haptics

Comparing vision and touch, a distinction can be made between the process of extracting object information and the type of object information that each modality can encode.

The processes involved in the visual and haptic extraction of object information across vision and touch differ in many ways. These differences include: a) the rate, b) the scope, and c) the relative ease of extracting object information. For extracting information, in vision a series of saccades rapidly pick up object information, whilst for touch, information is encoded at a much slower rate through a series of "successive

impressions" (Gibson, 1962). Both modalities are similar however in terms of both being able to encode object shape information and both can only explore one object in detail at a time, although vision can often achieve this in a fraction of the time that it would take touch. For vision, the scope of this extraction (i.e., the working range), with an appropriately sized set of objects is the horizon. For touch the scope is usually peri-personal space that is the reachable distance around the body. However, peri-personal space can be extended by using tools or looking at a body part in a mirror or video monitor (Holmes, Calvert, & Spence, 2004). For ease of extracting object information, vision is able to deduce rapidly object position/relative size information in just a few saccades, whereas for touch, an EP must be initiated to gain such information. Thus, a property of vision is the rapid extraction of relative information with little effort. This information is still accessible for touch, but requires a greater effort to encode and uses more system resources (the search for and the allocation of a specific EP).

Vision and touch differ according to their speed and accuracy of data extraction along object dimensions such as texture, shape and size (e.g., Lederman & Klatzky, 1987). Both modalities are capable of measuring an object's macrogeometric features (a term coined by Zangaladze, Epstein, Grafton, & Sathian, 1999) encompassing for example, object orientation, size and gross shape, and to an extent both are able to measure microgeometric features (for example, surface roughness and fine texture). It seems though that vision will preferentially encode an object's macrogeometric (global shape) properties whilst touch often preferentially encode microgeometric (local shape) information (Lakatos & Marks, 1999), suggesting that modality-encoding biases exist, where one modality may be used over another for the exploration of certain object dimensions. In confirmation, vision appears to be more involved in the perception of object geometry than touch (Klatzky, Lederman, & Matula, 1993) and touch has been

found to be better at measuring fine object texture than vision (Johansson & LaMotte, 1983), with the speculation that touch will be used rather than vision for the exploration of this dimension. There are also a set of dimensions that are particular to each modality. Vision for example is the only modality able to explore pigment-texture (e.g., the dappled colourings of polished marble) and colour (Gibson, 1962) whilst touch is the only modality to be able to directly explore object mass, surface friction, hardness and temperature (see Figure 2, which speculates on the visual and haptic modality encoding biases).

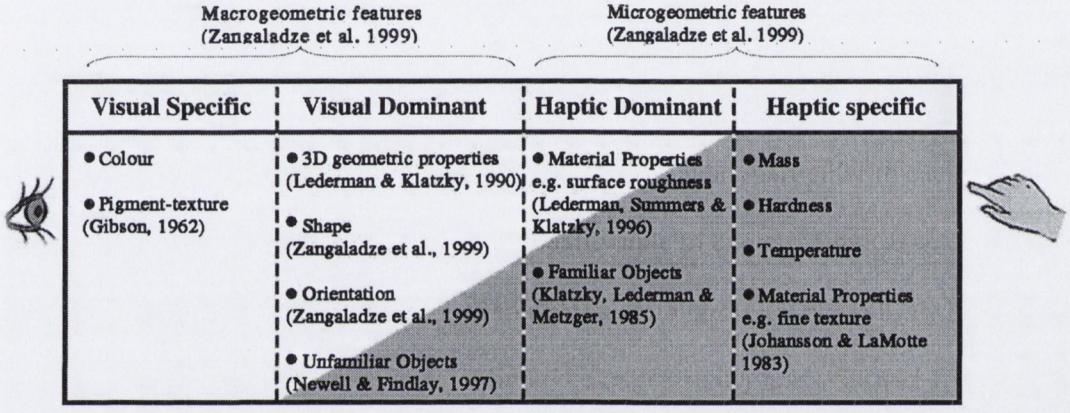


Figure 2. Information about an object either can be modality specific or can be shared across modalities. Task dependent modality biases have been mapped in the Figure. For example, only vision can perceive colour information (Visual Specific area, coloured white) whereas only touch can perceive mass (Haptic Specific area, coloured grey). Other features, such as orientation, can be perceived by both modalities. However, each modality can differentially influence the perception of these features. For example, Zangaladze, Epstein, Grafton & Sathian (1999) suggest that orientation perception is dominated by vision (the Visual Dominant area). On the other hand, Lederman, Summers & Klatzky (1996) argue that the perception of some material properties is dominated by touch (the Haptic Dominant area). Here, the relative influence of each modality on the perception of different types of object features has been characterised, based on some studies in the literature. Note however that this Figure is by no means comprehensive and acts mainly as a guide for modality biases in perception.

According to Gibson (1962), similarities, such as both modalities being able to sample object macrotexture, size, and microtexture are only apparent through active exploration, with both able to "yield the same information" and "the same phenomenal experiences". Thus, the simultaneous active exploration of the same object by vision and touch will yield a similar subset of information. As only one object is perceived, it

seems safe to assume that visual and haptic information has been bound together (e.g., Calvert, 1998) allowing for the creation of a rich cross-modal representation of the object in memory (Newell, Ernst, Tjan, & Bülthoff, 2001). Possible neurological ways for this binding/integration will now be covered.

1.2. Cross-modal integration in the cortex

In this section multisensory integration will initially be defined at the level of a single neuron, and then candidate cortical areas for multisensory integration will then be discussed.

1.2.1 Multisensory integration at the neuron level

The neurological analogue of cross-modal integration is multisensory integration. Both multisensory 'response enhancement' and 'response depression' are seen in the cortex, (where the neuronal response to stimulation from two or more modalities is greater or less than the response from stimulation from a single modality: e.g., Stein, Jiang, & Stanford, 2004). Meredith and Stein (1996) found that the greatest multisensory neuronal response in the superior colliculus (SC) of cats was found when a concurrent visual and auditory stimulus were in 'spatial register' with each other. Increasing spatial disparities between the visual and auditory stimuli produced respectively smaller response enhancements. Response depression occurred when the spatial disparity between the location of the visual and auditory stimuli was such that one or both of the stimuli fell outside each modality's corresponding receptive field (RF) for a particular neuron. Thus, the optimum multisensory neuronal enhancement occurs when the location information from each modality is the same. Response enhancement seems also to be dependent on different sensory information arriving at a neuron within a specific temporal window (Meredith, Nemitz, & Stein, 1987).

Interestingly, Meredith et al. found that response enhancement in the SC was not dependent on the concurrence of stimulation to the modalities but rather on the concurrence of 'discharge trains' from each modality (the period in which the greatest response is elicited by a neuron). Response suppression occurred when the two discharge trains did not overlap in time. The authors even suggest that "since visual and somatosensory latencies are so different... their response onsets may never coincide (if simultaneously presented)", (Meredith et al., 1987 p. 3227). A window of approximately 100ms was proposed to exist during which different sensory information converging on a neuron has the greatest chance of eliciting an enhanced multisensory neuronal response. It also seems that the findings mentioned for multisensory neurons in the SC do have some generality across the cortex (Stein & Meredith, 1993). However, cortical receptive fields in some cortical areas such as the anterior ectosylvian cortex of the cat may be larger than those found in the SC (Stein & Wallace, 1996).

1.2.2 Multisensory integration in the cortex

Broadly speaking, two accounts of multisensory integration are discussed within the neurological literature. The first view was evident in the 1960's to 1980's where vision and touch were seen as structurally and functionally independent systems (e.g., Felleman & van Essen, 1991) that broadly adhered to Fodor's (1983) rules of modularity. Accordingly, a large volume of literature is dedicated to processing in each of these modalities separately. Cross-modal integration was not considered to occur directly, but through an intermediatory or translation process such as through imagery. Hadjikhani and Roland (1998) for example, propose that multisensory cortical areas such as the claustrum (see also Ettlinger & Wilson, 1990) perhaps act as translator mechanisms between the modalities rather than multisensory storage areas. Recent findings on neuronal or cortical plasticity, however, have challenged this notion of

modularity: it is now accepted that cortical areas traditionally considered as sensory specific are not necessarily so and can, for example, be recruited by other areas in the case of sensory deprivation (Ramachandran, Rogers-Ramachandran, & Stewart, 1992).

The second, more recent account suggests that dedicated multisensory areas exist within the brain to process and perhaps store information from two or more modalities. Researchers have used a variety of techniques and paradigms to identify candidate areas within the brain involved in shape recognition. For example, Banati, Goerres, Tjoa, Aggleton & Grasby (2000) conducted a PET study during a novel shape matching task (the arc-circle test was used, where semi-circles had to be matched to circles of the same curvature) and found that the anterior cingulate, inferior parietal lobules and claustrum areas were selectively activated during cross-modal matching. Using fMRI, Amedi, Malach, Hendler, Peled & Zohary (2001) investigated cortical areas involved in cross-modal, familiar object recognition. They found that the lateral occipital complex (LOC), a cortical area well known to be involved in visual object recognition, was also active during haptic recognition of familiar objects. This activation was reduced but not eliminated when visual imagery was controlled for, suggesting that the LOC is directly implicated in visual and haptic recognition. The authors suggested that the LOC is a candidate cortical area for the convergence of multisensory information about objects. Another recent fMRI study involving haptic to visual priming of novel objects also provided evidence for the role of area LOC in cross-modal object recognition (James et al., 2002).

The question still remains however as to the architecture of these multisensory areas at the neuronal level. Meredith (2002) proposes that sensory convergence might be considered as a continuum (see Figure 3 for an illustration). He argues that at one extreme of the continuum convergence is at the areal level where sensory specific

neurons from several modalities converge in the same cortical area. At the other end of the continuum convergence is at the neuronal level where a single neuron is able to respond to inputs from one or both sensory streams. He argues that many cortical systems may lie somewhere along this continuum responding to both sensory specific and multisensory inputs.

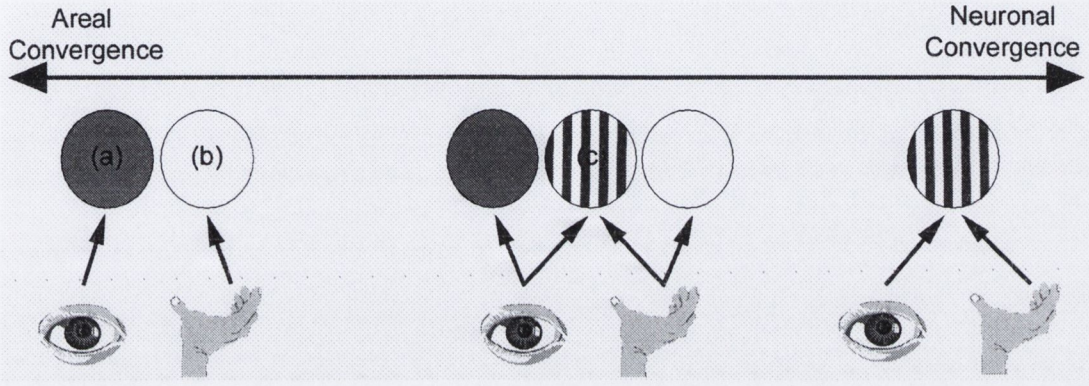


Figure 3. An illustration of areal to neuronal convergence continuum based on Meredith (2002). Three sets of neurons (circles) are evident in the above figure and each receives input from touch and/or vision. In the illustration, those neurons receiving input from vision are shaded grey whilst the neurons receiving input from touch are white. Hashed neurons receive input from both modalities. Areal convergence is where different modality specific neurons populate the same cortical or subcortical area (the left of the figure) and neuronal convergence is where a single neuron receives input from two or more modalities (the right of the figure). Meredith argues that most multisensory areas are characterised by both areal and neuronal convergence as illustrated in the centre of this continuum.

For example, area LOC may be populated by neurons that are selective to sensory information from one modality only. Therefore, areas such as the LOC are multisensory only because they are populated by multiple, sensory specific neurons. On the other hand, individual neurons in LOC may be multisensory and selective to particular features irrespective of modality. Such multisensory neurons have been found in sub-cortical structures such as the superiour colliculus (Patton, Belkacem-Boussaid, & Anastasio, 2002; Stein & Meredith, 1993) and in cortical areas such as parietal areas involved in representing extra-personal space (Graziano & Gross, 1995). It is feasible, therefore, that there exist multisensory neurons which are selectively

tuned to features of an object encoded through any modality. To the best of this authors knowledge however, there are no studies reporting the existence of such neurons in the human brain.

Recent studies have reported that primary sensory areas can also be involved in cross-modal tasks. For example, Zangaladze, Epstein, Greafton & Sathian (1999) found that applying transcranial magnetic stimulation (TMS) to the occipital cortex, an area previously thought to be dedicated for visual processing, hindered performance on a haptic grating orientation task in normally sighted participants. Conversely, visual tasks involving stimuli with tactile qualities have been shown to activate somatosensory cortex. Zhou & Fuster (1997) recorded activity in individual neurons in monkey somatosensory cortex and found that visual stimuli that are behaviourally associated with tactile information activated a number of these neurons (in this case, bars with horizontal or vertical gratings, perceivable by the haptic system as texture). The monkey was required to remember a visually presented grating orientation for a short period and then make a decision about which of two haptically presented gratings they had just seen. Zhou and Fuster suggest that neurons in somatosensory cortex are involved in the cross-modal short-term memory of object features. Pascual-Leone & Hamilton (2001) go one step further by suggesting uni-modal cortical areas may simply be specialists in processing a specific type of information (see also Freides, 1974)- as such, if that type of information is predominantly encoded by one modality, that modality will seem to dominate the corresponding unimodal area. Pascual-Leone & Hamilton suggest that the removal of input from such a dominant modality will lead to increasing activation of the associated 'unimodal' cortical area by other modalities. They demonstrated this effect in a radical experiment where participants were

blindfolded for 5 days in succession. The visual cortex was found to activate to tactile stimulation after 2 days of being blindfolded.

Interestingly, a number of behavioural effects can be predicted for each of these extreme accounts. One prediction is that the efficiency in accessing information stored in multisensory areas should be independent of the encoding modality. On the other hand, cross-modal performance should be reduced if uni-modal areas need to communicate information between each other because accessing of information across modalities may require a translation process.

Observations of neural activation alone might not reveal phenomenological experiences or the efficiency of the mechanism involved in performing a cross-modal recognition task. Behavioural studies can better reveal how information is shared across modalities. For example, neuronal activation would not reveal the nature of the subjective experience in the individual nor whether the person is experiencing an illusory percept or a veridical percept. The next section discusses factors that influence cross-modal integration of object information, to set the bedrock for a discussion on the potential mechanisms that permit integration to occur.

1.3 Factors affecting cross-modal integration

A number of factors can affect the nature and extent of cross-modal integration. Some factors are specified by the task at hand, other factors are involved in the encoding of the objects, and it also seems likely that the method in which object percepts are stored in memory may also influence integration.

1.3.1 Task factors

Familiarity with both task and stimuli may affect integration. When learning to drive, for example, initially each cross-modal action is performed explicitly until

automatic co-ordination between vision and touch is achieved. At first, changing into third gear requires a glance at the gear stick to find out exactly where it is, a press on the clutch with your foot and then a poorly timed, clumsy movement of the gear stick. As much information as possible is gathered about the task of changing gear through touch and vision (and through audition when the car screams its protests) but with practice, such overt cross-modal integration is not needed and the gear can be changed through touch and proprioception alone. Familiarisation can lead to the optimisation of information encoding such that other sensory information is no longer explicitly encoded but can be retrieved from memory for task completion. Welch & Warren (1980) refer to such factors as "historical", and suggest that if the history between two modalities is that they reliably encode the same information as each other, this may also bias the observer into perceiving a single object when in fact two objects are present. It thus seems that familiarity can act at a number of levels to influence cross-modal integration.

1.3.2 Encoding factors

Rock & Victor (1964) argued that for shape perception vision can dominate touch, thus coining the term 'visual capture'. However, more recent studies have suggested that 'capture' is an extreme form of cross-modal processing and that decisions are generally based on the reliability of information in each modality. Ernst and Banks (2002) systematically adjusted the reliability of visual information in a visual-haptic height judgement task. Using a cross-modal cue conflict paradigm, the authors decreased the reliability of the visual information and found that haptic information increasingly influenced the participant's percept. This makes sense if it is considered that the less accurate the information from a particular modality, the smaller the weighting given to that modalities input to a multisensory representation. As visual

information was being made less reliable here, the magnitude of the weighting of the visual input to a multisensory perception would be reduced, increasing the influence of the haptic information on the multisensory percept. Thus, the integration of visual and haptic information is achieved in a statistically optimum way (the maximum likelihood estimate or MLE) with the use of sensory information from either vision or haptics being dependent on the estimated reliability of that information. Therefore, the quality of the encoded visual and haptic information influences object integration. Also, factors that would influence the quality of sensory information, such as light levels and the wearing of gloves would thus affect cross-modal integration.

Situations may arise when vision and touch erroneously encode different information on the same object. This may be the result of encoding errors that occur for numerous reasons, including effects associated with the aging process, injury or experimental manipulation (e.g., by wearing prism glasses). Vision and haptics are able to correct for this by dynamically calibrating one to the other, so that at any one time, both perceptual systems are able to produce accurate representations of the world. For example, Ernst, Banks and Bülthoff (2000) found that the haptic system can bias how different visual depth cues are combined so that preference is given to a visual cue that reports the same tilt as done so by haptics. These effects were still evident after 24 hours, demonstrating that haptic biasing of the combination of individual visual cues last for a prolonged period. Along similar lines, Atkins, Fiser & Jacobs (2001) performed a series of experiments where participants had to judge the depth of cylinders when a number of different depth cues were available. Two visual cues and one haptic cue were used. After a training phase where one visual cue and the haptic cue portrayed the same object depth information, and a second visual cue portrayed different information, the authors found that subsequent participant judgements of

cylinder depth were biased towards the first visual cue. Atkins et al. go on to suggest that touch is used as a standard to which the reliabilities of differing visual cues may be compared, such that the more reliable a visual cue, the more it influences the final percept of an explored dimension. Thus different cues are combined according to the reliability of the encoded information in order to achieve a robust representation.

1.3.3 Memory effects

The method of storing object information also would likely affect cross-modal integration. If object information is held within a multisensory memory system to which both vision and haptics have equal access, integration may occur automatically as percepts overlap when derived simultaneously from the same spatial location (this is discussed in the following section). If on the other hand objects are held independently from each other, some process to co-ordinate integration and recoding may be necessary for the transfer of information across modalities.

In support of the modalities commonly storing object information, both vision and haptics have been shown to be affected by a range of common illusions (e.g., the Ebbinghaus and Müller-Lyer illusions: Franz, Gegenfurtner, Bulthoff, & Fahle, 2000; Gallace & Spence, 2004). Also objects and words may be primed across modalities (Easton, Srinivas, & Greene, 1997; Reales & Ballesteros, 1999) which strongly suggests that the modalities share the same object representations.

But the evidence is not so clear-cut as not all illusory effects are found across modalities. The Poggendorff illusion has not reliably be found in touch (for a discussion, see Gentaz & Hatwell, 2004) and interestingly, the haptic Zöllner illusion is opposite to that found in vision (Suzuki & Arashida, 1992). However these illusions may arise for reasons not linked to how an object representation is stored in memory (Coren & Girgus, 1978). Stronger evidence for the independence of visual and haptic

percepts comes from the finding that articulatory suppression hinders haptic performance at recognition tasks, but does not hinder vision to the same extent (e.g., Gilson & Baddeley, 1969; Mahrer & Miles, 2002). If it is presumed that this recoding process is vulnerable to articulatory suppression, this finding firstly demonstrates that recoding takes place, and also implicates phonological mechanisms in the recoding. However, it is perhaps more important to consider verbal recoding as a strategy which facilitates both normal cross-modal integration or even normal haptic processing (e.g., Millar, 1975; Newell, Woods, Mernagh, & Bülthoff, 2004), as individuals who lack the ability to verbalise may obviously still integrate information across the senses (e.g., people born deaf and non-human primates, Stein & Meredith, 1993).

It may be that "the integration of information from different sources cannot be accounted for by a single, overarching mechanism" (Millar, 2000) and that instances of both multisensory and modality specific memory systems are found across the cortex (for a review see Mesulam, 1998). Indeed, there are multiple levels of processing within modalities. There are, for example, independent dorsal and ventral pathways for the processing of 'what' and 'where' information in vision (Haxby et al., 1991), and this independent processing has also recently been found in haptics (Chan & Newell, submitted).

Despite the range of factors that may affect cross-modal integration, the brain manages to integrate objects across the senses with apparent ease. It may be that cross-modal integration follows some simple rules or principles, and this is the topic of the next section.

1.4 How is cross-modal integration achieved?

Recent models of cross-modal integration will be first discussed, followed by a discussion on the rules or principles that seem pertinent to cross-modal integration.

1.4.1 Models of integration

Three main models of cross-modal integration have been speculated on in the literature.

Millar (2000) suggested that object information is stored within interrelated networks in the brain. Exposure to an object will lead to a pattern of activation within these interrelated networks. Each modality encodes a range of specialised yet complementary information which may overlap with all or other modalities. This overlap of information is what allows the brain to consider an object, which is perceived across multiple modalities, as a single object. Information may overlap for a number of reasons, ranging from quite simple criteria such as temporal co-incidence and stimulus intensity, to more complex reasons such as "continuity versus change" information which is pertinent to object recognition (e.g., features such as smoothness and edge information could be redundantly encoded across vision and touch).

Integration in Millar's interrelated networks model may occur according to the Bayesian Maximum Likelihood model, recently proposed by Ernst & Bülthoff (2004) where information is integrated within a neuronal population according to the Maximum Likelihood Estimate (MLE, Ernst & Banks, 2002). In the maximum likelihood model, information is combined in a feed forward manner within a multisensory population of neurons and the resulting spread of activation is characteristic of the MLE model of integration. Simple rules may be all that is required for information to be integrated here, and Ernst and Bülthoff suggest that spatial co-location and temporal synchrony are suitable rules.

Cross-modal integration may also be the result of the translation of haptic information into visual representations mediated by, for example, visual imagery (see Zhang, Weisser, Stilla, Prather, & Sathian, 2004). The same criteria that affect the

neurological MLE model are likely to affect also this visual recoding model. However, more complex top-down factors may also affect whether or not information from across the senses is integrated, or for example, whether haptic information is recoded.

Lastly, vision and haptics object processing may be functionally independent (Reed, Shoham, & Halgren, 2004). It is likely, for such a model, that complex top-down criteria would determine whether information is to be translated and integrated. Needless to say however, more simply criteria, such as objects being commonly located in time and space may also be a useful indicator for such a model that information should be integrated.

Although there are a range of different criteria which each model requires for objects to be integrated, it is interesting that temporal synchrony and spatial co-location seem to be crucial within all models.

1.4.2 Criteria for cross-modal integration

The factors necessary for the cross-modal integration of object information are speculated on here. Factors must be sensitive enough to combine fleeting information from across the senses- e.g., a car horn and a rapid movement in the periphery of your vision as you cross the road will hopefully alert you to move out the way of a careering car. Such factors must also prevent information relating to independent events from being integrated- e.g., it would not be advantageous to integrate the car horn noise with the sight of a flying bird. Potential factors which facilitate cross-modal recognition are discussed below.

Welch & Warren's (1980) intersensory bias model proposes a range of precursory factors necessary for the perception of a single object across multiple modalities. The authors attempt to account for both everyday multisensory perception (Welch, 1999) and the consequences of disrupting this perception, by for example, altering the

information that each modality encodes about an object (see also Bertelson & de Gelder, 2004). The crux of the model is that during a task such as visual-haptic cross-modal object recognition, a number of dimensions particular to an object, such as shape, movement, and stimulus onset are redundantly encoded by both modalities. As just one object is encoded, information referring to the object's different dimensions or features are generally invariant across the senses. On the other hand, when information is derived from different objects it is likely that there will be a large volume of non-overlapping information across modalities (c.f. Millar, 2000). Thus, the level of similarity across redundantly encoded information may be the criterion for a single object to be perceived (the "unity assumption", Welch & Warren, 1980). The authors also suggest a range of cognitive factors, such as past experience (i.e., a 'historical' factor) and attention that act to modulate this.

But is this complex operation really necessary? It may be sufficient that just one or two cross-modal object dimensions are contrasted. It may even be the case that these rules are a natural product of the way information is stored in the brain, and there is no need for separate resource demanding mechanisms to process and detect invariance across modalities. As already discussed, multisensory integration is dependent on stimulation across modalities activating a neuron within a temporal window. A neuron will also only activate from stimulation if its corresponding visual and haptic receptive fields overlap. Thus the co-location of stimulation is also crucial for a multisensory response. It seems likely that the neurological precursors of the response enhancement of 'temporal synchrony' and 'spatial co-location' may be sufficient for the occurrence of cross-modal integration (this has been suggested as a mechanism co-ordinating exogenous crossmodal attention: Driver & Spence, 1998).

1.4.3 Temporal synchrony and its role in cross-modal integration

Typically a single object is encoded simultaneously across the senses, so the timing of stimulus presentation is invariant. It has been shown developmentally that infants first respond to temporally coincident stimuli then later to spatially redundant factors (Lewkowicz, 1999). It may even be that temporal co-incidence is the sole underlying factor that determines cross-modal integration. If this were so, two spatially independent events may be integrated if there was a high degree of temporal congruency between them. A number of illusions do seem dependent on temporal co-incidence and perhaps the best known examples of cross-modal illusions concern the interaction between audition and vision, such as the McGurk and ventriloquism effects (Howard, 1966; McGurk & MacDonald, 1976). For example, the McGurk effect occurs when discrepant visual information (lip movements) and auditory information (phoneme sounds) is integrated leading to the experience of a qualitatively different percept than that would be provided by either the visual or the auditory stimulus individually (e.g., seeing /gi/ lip movements whilst hearing a /bi/ sound may be experienced collectively as a /di/ sound). Demonstrating the importance of temporal co-incidence, the McGurk effect breaks down when a delay over 180ms is introduced between the onset of the cross-modal stimuli (e.g., Munhall, Gribble, Sacco, & Ward, 1996).

Temporal order judgement (TOJ) tasks have been used to explore the temporal window in which stimuli may be presented across the modalities yet still be integrated. In the TOJ task, simple stimuli (e.g., flashes and vibration and beeps) are presented at varying stimulus onset asynchronies across modalities and it is the order of stimulus presentation that must be determined. When the order cannot be determined, it has been argued that this is because the stimuli have been binded together (e.g., Zampini,

Shore, & Spence, 2003). By determining the extent to which stimuli may be temporally separated whilst their order is undeterminable, the temporal window of cross-modal integration can be deduced. Hirsh & Sherrick (1961) found a temporal window of 20ms in which the order of presentation could not be accurately established for all modality combinations (between vision, touch and audition), suggesting that the temporal window of cross-modal integration for such stimuli is 20ms in size.

It is interesting that stimuli in the McGurk effect may be presented within a much larger temporal window (180ms, Munhall et al., 1996) than objects in the TOJ task (20ms, Hirsh & Sherrick, 1961). One difference is that the stimuli in the McGurk illusion are much more complex than the simple stimuli in the TOJ tasks. It may be that complex stimuli have larger temporal windows of integration than those of more simple stimuli (raised by Spence & Squire, 2003).

A consequence of a delay between the presentation of a stimulus from one modality from presentation of a stimulus in another is that a delay would require the first presentation to be held in memory, and this would be subject to decay effects and possibly interference effects (Gilson & Baddeley, 1969; Logie & Marchetti, 1991; Sullivan & Turvey, 1972). A useful step before exploring the effects of temporal synchrony on integration would be to determine the rate of object percept decay within and across modalities. This would give clues as to how the modalities maintain object information, providing a suitable bedrock for further study. Also, the complexity of stimuli should also be considered when determining the rate of object percept decay within and across modalities.

1.4.4 Spatial co-location

Temporal congruency may not be sufficient for integration in all conditions. For example, vision and touch may simultaneously encode information about different

objects during the exploration of a complex scene (this may occur, for example, if the hand occludes what it is feeling from view and vision attends to another object). It would not be beneficial to integrate cross-modal information in this situation. It may be that the brain makes use of both temporal co-incidence and spatial co-location to decide if integration should take place (e.g., Ernst & Bülthoff, 2004; Lalanne & Lorenceau, 2004; Stein & Meredith, 1993; Wallace et al., 2004). For example, referring back to TOJ research, recently Spence, Shore & Klein, (2001) found that temporal window size is modulated by spatial co-location. In Hirsh & Sherrick's study the objects were not presented in the same spatial location across the senses. Spence et al., found in a TOJ task that stimuli which were not spatially co-located needed to be separated by 37ms to be correctly ordered. However, when the stimuli were spatially co-located they needed to be temporally separated by 68ms for their order to be correctly determined. Thus, when objects are spatially co-located, the temporal window is larger than it would be if the objects were spatially separated. Response times can also be enhanced when stimuli are spatially and temporally aligned. It takes a shorter time to identify whether a tactile stimulation on the arm is of a low or high frequency when a moving stimulus is seen to collide with the arm at the same time and location as the tactile stimulation, compared to when the moving stimuli collides either to the left or right of the tactile stimulation (Gray & Tan, 2002). Low spatial congruency between visual and haptic events can also lead to a degradation of an object's representation in memory (Kerzel, 2001; Logie & Marchetti, 1991; Mahrer & Miles, 2002). Kerzel (2001) found that memory for the velocity of a visual stimulus was affected by different movement of the participant's own arm. An arm movement that occurred simultaneously yet was relatively slower than the velocity of a visual stimulus resulted in a memory of a slightly slower visual

stimulus than the actual speed of the visual stimulus. The opposite effect was found with a faster arm movement.

Although temporal synchrony and spatial co-location have been implicated in cross-modal integration, their exact role in the process, and whether or not they are solely sufficient to determine visual and haptic integration of object information has yet to be determined. It also important to consider that stimuli with high level information may rely on perhaps more complex combinations of temporal and spatial congruency for optimum integration (Calvert, 1998).

1.4.5 A methodology for exploring cross-modal integration.

It has been discussed how the brain makes use of temporal synchrony and spatial co-location to differentiate between the same object being perceived across the senses, and different objects being perceived. It follows that if these factors are crucial for integration, then presenting a visual and a haptic object in the same location at the same time should lead to them being integrated in the brain. The consequence of presenting two different objects in such a way will lead to their percepts interfering with each other, making it harder to realise there are indeed two objects (Miller, 1972; Zampini et al., 2003). By varying the level of temporal co-incidence and spatial location of objects from each other, and by measuring how well participants are able to deduce whether or not two different objects are perceived, the extent of the influence of spatial and temporal factors in cross-modal integration may be determined. Thus, percept interference may be an excellent dependent variable to explore spatial and temporal factors.

1.5 Conclusion

Vision and haptics differ in terms of how they encode information, and also what they may encode. But they both can encode object information. A range of literature has been reviewed showing that information is often integrated across the modalities in such a way that the resulting percept is more robust than if a single modality had been used for perception. It is likely that temporal synchrony and spatial co-location of object presentation across modalities are sufficient cues for the occurrence of cross-modal integration, and this is what this thesis intends to determine. A clearer understanding of how object representations are formed by vision and touch will provide a suitable bedrock for the subsequent exploration of temporal and spatial factors.

CHAPTER 2: THE EFFECT OF TEMPORAL DELAY AND SPATIAL DIFFERENCES ON CROSS-MODAL OBJECT RECOGNITION

Abstract

In a series of experiments, the matching of objects across visual and haptic modalities was investigated across different time delays and spatial dimensions. In all of the experiments, simple shaped figures were used as stimuli that varied in either the x or the y dimension or in both dimensions. In Experiment 1, cross-modal matching performance was found to decrease as a function of the time delay between the presentation of the objects. No difference was found in performance between the visual–haptic (VH) and haptic–visual (HV) conditions. Cross-modal performance was better when objects differed in both the x and y dimensions rather than in one dimension alone. In Experiment 2, the relative contribution of each modality to performance across different interstimulus delays was investigated. No differential effect of delay between the HH and VV condition was found, although general performance was better for the VV condition than for the HH condition. Again, responses to xy changes were better than changes in the x or y dimensions alone. Finally, in Experiment 3, performance in a matching task with simultaneous and successive presentation conditions was investigated. No difference between simultaneous and successive presentation conditions was found however. The findings suggest that the short-term retention of object representations is similar in both the visual and haptic modalities. Moreover, these results suggest that recognition is best within a temporal window that includes simultaneous or rapidly successive presentation of stimuli across the

modalities and is also best when objects are more discriminable from each other.

2.1 Introduction

The fundamental differences between the encoding mechanisms for the visual and haptic senses would suggest that these modalities might also use different underlying mechanisms for the storage of information. For example, in a cross-modal recognition task, encoding time is generally longer for haptics than for vision (Newell, 2004). Once an object has been encoded then active touch is both a proficient and efficient method for recognising common objects (Klatzky, Loomis, Lederman, Wake, & Fujita, 1993), and for exploring and recognising novel objects (Newell et al., 2001). However, the fact that vision and touch encode and store object properties in different ways does not necessarily imply the existence of different storage systems.

The studies presented here are concerned with how objects are stored in visual and haptic memory and whether these memory systems are functionally similar and thus allow for efficient cross-modal recognition. Specifically investigated here was the effect of temporal delay between learning the stimulus in one modality and retrieving its representation in the other. Also of interest was whether spatial differences along one or two dimensions between objects would affect cross-modal recognition.

A large volume of literature has explored visual and haptic memory systems, and two broad accounts have emerged. The first of these is that object memory is multisensory such that object representations are easily shared across modalities, provided the same properties are encoded. Here no cost between cross-modal performance and within-modal performance would be expected. Furthermore, any observable effects on recognition performance with manipulations such as time delays

should be common to all modalities. The second account would be that objects are stored as modality-specific representations requiring a recoding from one memory system to the other. Such recoding would result in a cost in performance in terms of time and error rates relative to within-modal performance. Moreover, there would be no reason to expect similar effects on recognition performance across modalities with manipulations such as time delay. In other words, the memory systems may be functionally independent such that performance in one might not be related to performance in the other.

Many recent studies support the idea that visual and haptic memory for objects is shared. For example, neuroimaging evidence has emerged to show that the cortical area involved in visual object recognition is also involved in haptic object recognition. Previous studies have reported that the lateral occipital complex (area LOC) is active during visual object recognition tasks (see Grill-Spector, Kourtzi, & Kanwisher, 2001 for a review). More pertinently however, Amedi, Malach, Hendler, Peled and Zohary (2001) reported that activation in the LOC to objects is not specific to the visual modality since LOC is also active during haptic object recognition. Similarly, James, Humphrey, Gati, Servos, Menon and Goodale (2002) reported activation in area LOC to the visual presentation of novel objects that were previously learned through touch. Furthermore, this area was not active to auditory information about objects (Amedi, Jacobson, Hendler, Malach, & Zohary, 2002) suggesting that area LOC (or LOtv as termed by Amedi et al.) is involved in recovering the geometrical shape of objects.

Some recent behavioural studies have also suggested that object representations are shared across modalities. For example, Easton, Greene and Srinivas (1997) reported cross-modal priming effects for 2D patterns and 3D objects suggesting that vision and haptics share representations of the objects based on their shape and

structure. Reales & Ballesteros (1999) also found that when an object was presented in one modality (vision or haptics) the recognition of the same shape in the other modality was facilitated. Cross-modal object representations can also complement each other in order to achieve object constancy with changes in viewpoint (Newell et al., 2001). Moreover, bimodal encoding of objects can result in recognition performance that is superior to uni-modal encoding of objects (Newell, Bülthoff, & Ernst, submitted), suggesting that a rich representation of objects is created in memory across the modalities. Taken together, these studies support the idea that object representations might reside in a common storage location which both vision and touch can easily access.

However, other studies suggest that memory for touch is not equivalent to visual memory. Research on tactile memory provides tentative support for a tactile memory independent of other sensory memory systems and provides a framework for the characterisation of haptic memory. For example, Gilson and Baddeley (1969) found evidence for a tactile memory that is comprised of two elements, one affected by articulatory suppression, and the other unaffected. They applied a single tactile stimulus to a specific location on the underside of a participant's forearm and the participant's task was to point to the location of that stimulus. The authors introduced a delay between stimulus and response ranging from 0 to 60 seconds. In one condition participants conducted an articulatory suppression task during the delay and in the other condition they were silent. Gilson and Baddeley found that accuracy levels were maintained for up to 15 seconds in the silent condition. After this 15 second delay, accuracy then diminished. Articulatory suppression resulted in worse performance than the silent condition but these differences emerge only after 15 seconds delay. Using the same paradigm however, Sullivan and Turvey (1972) found that accuracy was reduced

after a delay of only 5 seconds. There were two task conditions, one with an unfilled delay between stimulation and recall, and the other where the delay was filled with a subsidiary arithmetic task. The authors found that task accuracy was lower for the condition with the arithmetic task. Mahrer and Miles (2002) found that articulatory suppression impairs accuracy in a task where two temporally separated sequences of tactile stimuli applied to different fingers were compared. This result suggests, as in the previous studies, that there is a verbal component to the memory of tactile sequences. Although together these studies involve an investigation of tactile memory, it might be that tactile stimulation of body parts involves processes that are independent of those involved in haptic object recognition and may not inform us as to the precise nature of representations of objects per se more the nature of body representations. Furthermore, studies on passive touch may reveal very little about the processes involved in active touch such as those involved in object recognition tasks (Gibson, 1962; Heller, 1984).

Some further studies have investigated active haptic memory for objects. For example, Kiphart, Hughes, Simmons & Cross, (1992) found that the haptic memory of complex objects decayed only after a delay of 15 seconds (i.e., a similar decay function to that described by Gilson and Baddeley). In a study on 3-dimensional object recognition in children, Millar (1974, experiment 2) found that haptic matching performance for nonsense objects was better with inter-stimulus delays of 1 second relative to delays of 5 and 30 seconds, suggesting that haptic memory starts to decay immediately after the exploration of an object. On the other hand, Garvill and Molander (1973) reported no effects of inter-stimulus delays from 1 second to 30 seconds in either haptic, visual or cross-modal matching of 3-dimensional nonsense objects. Differences between these studies may be accounted for by differences in task demands or stimulus complexity which may influence the encoding, storage and

retrieval of object information. For example, Kiphart et al. reported using complex objects made from 15 Lego bricks and differences between objects were characterised by changes in the position of 3 of these bricks. In comparison, the shapes used in the Garvill and Molander study were simple. Thus, the differences between these studies may reflect differences in the stimulus complexity, although it is difficult to quantify the differences in stimuli across these studies. Additionally, the degree of familiarity with the objects being explored and whether or not they can be verbalised may also affect performance (Millar, 1981).

In these experiments, the effect of temporal delay on within and cross-modal matching performance in an object recognition task was investigated. Unlike most previous studies on haptic and visual memory performance which used object stimuli sets where objects were categorically different from each other (e.g., a square versus a circle) in these experiments a set of 'L-shaped' object stimuli were used that systematically varied along one or two stimulus dimensions. Such objects have been previously used by Millar and Al-Attar (2000) to investigate vertical and bisection illusions in vision and touch. They found that these L-shapes elicit the same the same illusion in touch as in vision, suggesting that differences between these shapes provide the same spatial effects. Furthermore, verbal recoding has been found to improve tactile memory especially for similar tactile items (Millar, 1975), therefore it was necessary to ensure that the differences between the stimuli were difficult to verbalise. Since stimuli constituted metric rather than categorical differences such differences are thus difficult to verbalise. The main interest here was to investigate whether objects that varied from each other in more than one dimension would be more easily discriminated across modalities and whether objects that are more easily discriminable are also remembered more accurately over time.

2.2 Experiment 1

A cross-modal matching task was used to measure object recognition performance in this experiment. Investigating memory performance in object recognition was particularly of interest, therefore, traditional tests of short-term memory (e.g., memory span tests or Brown-Peterson designs) were not appropriate here. It was of interest whether a time delay between matching objects had the same effect on objects that were encoded visually and later presented haptically versus objects that were encoded haptically and then presented visually. If visual memory is independent from haptic memory then different effects of delay on cross-modal performance would be expected. If, on the other hand, object memory is multisensory then similar effects of delay irrespective of encoding modality would be expected. Furthermore, it was of interest whether effects of delay are influenced by stimulus complexity. It might be assumed, as suggested by previous literature, that the representations of more discriminable objects would be retained across large delays between presentation of the stimulus to be encoded and presentation of the comparison stimulus (e.g.,Garvill & Molander, 1973) whereas the representation of less discriminable objects might be more susceptible to decay.

2.2.1 Method

2.2.1.1 Participants

16 undergraduate students (4 male and 12 female) from Trinity College Dublin participated in the experiment for course credits. Their ages ranged from 17 to 26 years, with a mean age of 22.2 years. All of the participants had normal or corrected-to-normal vision and none reported any somatosensory impairment.

2.2.1.2 Stimuli

25 'L' shaped objects were created as stimuli. The objects differed from each other according to the length of their x- and y-axes arms. The smallest object in the set measured 2cm in the y-axis and 2cm in the x-axis. The largest object measured 6cm by 6cm. The length along each axis increased in increments of 1cm. All possible pairings of x- and y-axis lengths were created between 2cm and 6cm, to create the complete set of 25 objects. The largest object occupied approximately 8 degrees of visual field at a distance of 57cm from the participant. Figure 4 illustrates the set of stimuli used in this experiment.

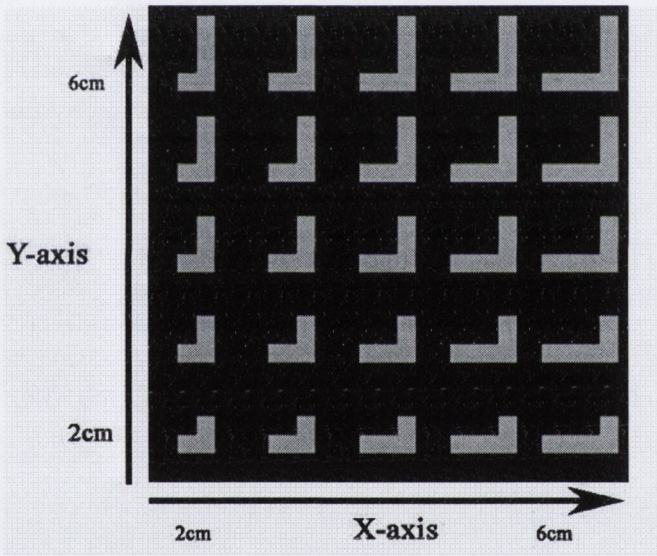


Figure 4. An illustration of the stimuli set. All the stimuli were 'L-shaped'. The lengths of their y and x-axis extensions were measured from the inside edge of each arm (the minimum extension was 2cm, the maximum 6cm). Stimuli which are horizontally or vertically next to each other on the matrix differ by 1cm from each other.

Two sets of stimuli based on these size criteria were created; a Haptic and a Visual set. The haptic set was made from four fused sheets of 0.5cm thick laminate and always measured 2cm in depth. The visual set was presented on a computer screen and

was made using the CorelDraw© graphics package. The visual set was created so that they matched the physical dimensions of the haptic set exactly.

2.2.1.3 Apparatus

The apparatus consisted of a computer monitor placed on a platform which was used to occlude the haptic stimuli from view. The centre of the monitor was then positioned approximately at eye level for each participant and positioned 57cm away from the participant. Figure 5 illustrates the experimental set-up. Visual stimuli were presented centrally on the 21 inch CRT monitor and haptic stimuli were presented underneath the monitor behind the front of the platform and thus hidden from view. The haptic set of stimuli was affixed onto the end of a supporting structure so that each object was suspended 10cm above the table to allow the participants to freely explore each object. The stimuli were held in place by two prongs, which projected from the supporting structure and fitted into two holes in each of the stimuli. These holes were located at the junction of each stimulus's arms allowing each length of the 'L' shaped stimulus to be perceived. The supporting structure's base fitted into a specially made socket on the table to keep the stimuli in a stable vertical orientation. Once positioned the haptic stimulus was directly aligned with the front of the computer monitor.

The experiment was run using DMDX software for the PC (Forster & Forster, 2003). A head mounted microphone was used to record participants' responses during the experiment.

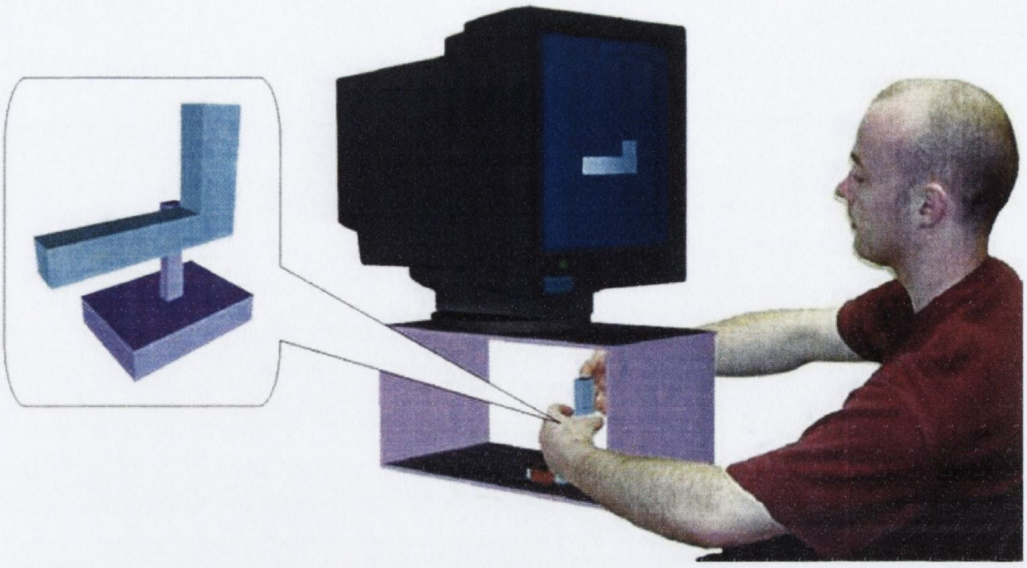


Figure 5. An illustration of the experimental set up including the monitor and the supporting structure. An illustration of the stimulus is shown in the inset.

2.2.1.4 Design

The experiment was based on a mixed-subject design with two-within and one-between subject factors. The between-subject variable was the order of modality (ORDER) between learning and test. There were two levels to this factor; visual learning followed by haptic testing (VH) and haptic learning followed by visual testing (HV). As the main purpose of this research was to compare multisensory object memory performance across different delays and not to compare memory performance across modalities per se, it was not necessary to have VV and HH controls. The within-subject factors were the inter-stimulus interval (ISI), and spatial differences (SPATIAL). For ISI the levels were 0, 15 and 30 seconds and for SPATIAL they were differences in the x-axis, y-axis or xy-axes between the stimuli. On 'different' trials, there were equal numbers of 1cm, 2cm and 3cm size differences along the x- or y-axes. For differences along both x- and y-axes, the size difference was always 4cm and differences in scale between two objects were avoided (e.g., if one object was 2cm by

2cm then the other could not be 4cm by 4cm). The sum of x-axis and y-axis changes equalled the sum of xy-axis changes. This allowed for a direct comparison to be made between recognition performance to differences in the x- or y-axis to differences in the xy-axes changes.

Two dependent variables measured were hit rate (HIT, whether two *same* shaped objects are perceived as identical) and correct rejection rate (CR, whether two *different* shaped objects are perceived as non-identical). In this study, when considering the effects of ISI and MODALITY ORDER on matching performance, for power issues the mean of HIT and CR scores were used as a derived dependent variable as there was no reason to suspect that HIT and CR scores would differ in terms of the effect ISI and MODALITY ORDER[1]. However, this study was also interested in the effects of SPATIAL differences, ISI and MODALITY ORDER on matching performance. As the objects in each object pair had to differ in shape for this analysis, only trials where the objects were *different* could be used here. Thus, two analyses were performed, one looking at the effects of ISI and MODALITY ORDER on matching performance, and the other looking at the effects of ISI and SPATIAL differences and MODALITY ORDER on matching performance. Also measured was voice-activated response time (RT, for correct trials only). All trials were randomly presented across the participants.

---

[1] For *same* trials, the stored representation of an object had to be identified as being identical in shape with a second new object representation for a correct HIT response. For different trials, the stored representation had to also be identified as different in shape with a second new object representation for a correct CR response. Thus both trial types were identical up until the point where a comparison had to be made between the first and second object representation. Because of this, the first object representation for both trials types should be similarly influenced by the effects of ISI. It seems likely then that both same and different trials would therefore be similarly affected by the effects of ISI. Please note that it is not argued that objects here would be cross-modally integrated. It is likely that if objects were integrated, than HIT and CR scores may then differ from each other (see Chapter 3).

2.2.1.5 Procedure

Before the experiment, participants were randomly assigned to the two experimental groups of the ORDER condition, namely VH and HV learning and testing conditions.

The experiment was based on a same/different matching protocol and participants were instructed to decide whether two successively presented stimuli were the "same' or "different" as fast and as accurately as possible.

Prior to the experiment, participants in the HV condition were instructed to match objects by first feeling and then looking at the objects in each trial. Participants in the VH condition were instructed to first look at and then feel an object. Prior to haptic presentation, all participants received a 500msecs warning tone (a series of beeps) which alerted them to prepare to feel the stimulus. They then heard a long continuous tone during which instructed them to feel the haptic stimulus. The tone lasted for 3 seconds and on offset participants had to remove their hands from the object. Thus, haptic exploration of an object lasted for 3 seconds which was considered adequate timing to encode the properties of the object through touch (Kiphart et al., 1992). A visual presentation of a stimulus was preceded by a fixation cross for 500ms which alerted the participant and prepared them for the subsequent presentation of a visual stimulus. A visual object remained on the screen for 250 msecs[2]. The delay between the presentation of the first and second stimulus in a trial was either 0, 15 or 30 seconds. There was an inter-trial interval (ITI) of 15 seconds during which the experimenter changed the haptic stimulus for the next trial.

---

[2]The presentation timings of both the visual and haptic stimuli during learning were carefully determined in a pilot study to ensure that there were no encoding differences across the modalities. Such differences might have caused differential rates of memory decay over time because the representation in one modality may be weak compared to the representation in the other. Thus, pilot studies were used to determine the presentation timings across modalities that resulted in equal cross-modal performance for the 0 second ISI.

For both the VH and the HV conditions, the second stimulus was 50% likely to be identical to the first. On the trials where the two stimuli were different, the second stimulus could differ along the x-axis, the y-axis or both (see design section). There was a total of 108 trials in the experiment (3 SPATIAL conditions, 3 ISIs, repeated six times, totalling 54 'different' trials and a further 54 'same' trials). To familiarise the participant with the procedure, 10 practice trials were given at the start of the experiment. The voice-activated response recorder was calibrated for each participant. The experiment lasted for approximately 80 minutes.

2.2.2 Results

The main measure of performance was the mean hit and correct rejection scores for each participant across the MODALITY ORDER and ISI conditions as shown in The main measure of performance was the mean hit and correct rejection scores for each participant across the MODALITY ORDER and ISI conditions as shown in

Table 1 (1.3% of the data was lost due to background noise[3]). A mixed-factorial ANOVA was conducted, using MODALITY ORDER as a between subject factor and ISI as a within subject factor. No effect of MODALITY ORDER was found [$F(1,14)=1.665$, n.s.]. A main effect of ISI was found [$F(2,28)=7.183$, $p<0.005$]. A posthoc Newman-Keuls analysis revealed that delays of 0 seconds produced significantly better matching performance than delays of 30 seconds ($p<0.005$) but not delays of 15 seconds. There was a non-significant trend towards a difference in

[3] The computer determined whether or not the participant made a response ("same" or "different") by contrasting the volume of the participant's utterances to an established baseline. The baseline was individually determined for each participant before the experiment. Data were lost when noises spurious to the experiment exceeded this baseline. The percentage of data lost cited above refer to the data lost for the entire experiment. For all experiments in this chapter, a fan in the experimental room and accidental noises emitted by the participant both seem likely to have led to the loss of data.

performance between delays of 15 and 30 seconds (p=0.052). There was no interaction between the factors [$F(2,28)< 1$].

Table 1.
Showing mean percentage correct responses (hit and correct rejection scores) and mean response times (RTs) across the different ISIs in each order of modality in Experiment 1. Standard error of the means is shown in parenthesis.

Modality Order	Vision - Haptics			Haptics - Vision		
ISI	0	15	30	0	15	30
% Correct	86.89	81.94	76.82	82.70	79.51	75.22
Response	(1.40)	(2.16)	(1.84)	(1.71)	(4.13)	(2.23)
RTs	1929.7	2044.8	2226.6	770.3	878.1	869.9
(in ms)	(272.9)	(245.6)	(282.8)	(40.1)	(80.6)	(79.8)

The mean response times (across hits and correct rejections) for each MODALITY ORDER across the different ISIs are shown in Table 1. A two-way, mixed subject design ANOVA was conducted with MODALITY ORDER as the between-subject factor (VH and HV) and ISI (0, 15 and 30 seconds) as the within

subject factor[4]. A main effect of ORDER was found [$F(1,14)=20.329$, $p<0.01$). Response times to the HV condition were significantly faster than response times to the VH condition. A main effect of ISI was found [$F(2,28)=8.304$, $p<0.01$]. There was no significant interaction between the factors [$F(2,28)=1.805$, n.s.].

A three-way, repeated measures ANOVA was conducted on the correct rejection responses (i.e., trials where the objects differed) with MODALITY ORDER, ISI and SPATIAL differences as factors. There was no effect of MODALITY ORDER [$F(1,14)<1$] or of ISI [$F(2,28)=1.658$, n.s.]. The effect of SPATIAL differences was significant [$F(2,28)=43.200$, $p<0.001$]. The mean percentage performance for each of the SPATIAL differences was as follows: 69.9% (3.1% standard error) accuracy for x-axis differences; 66.4% (2.9% S.E.) accuracy for y-axis differences and 90.6% (2.0% S.E.) accuracy for xy-axis differences. A Newman-Keuls test revealed that performance was significantly better to changes in the xy dimensions than changes in either the x-axis ($p<0.001$) or y-axis ($p<0.001$) dimensions and no difference between the x- and y-axis dimensions. There were no significant interactions between the factors [ORDER x ISI- $F(2,28)<1$; ORDER x SPATIAL- $F(2,28)=2.04$, $p>.05$; ISI x SPATIAL- $F(2, 28)=1.24$, $p>.05$; ORDER x SPATIAL x ISI- $F(4,56)<1$].

2.2.3 Discussion

The effect of delay between cross-modal presentations of object stimuli was found to be the same whether the initial object was encoded visually or haptically. This finding suggests that visual and haptic memory for objects is equivalent and that representations decay at the same rate. The fact that response times were slower in the

---

[4]The SPATIAL factor was analysed separately because the 'same' trials were identical for all levels therefore HIT data could not be separated across the levels.

VH condition than the HV condition is simply a reflection of haptics being relatively slow in encoding shape properties when compared to vision. A more interesting finding is that there was no difference in discrimination performance between the cross-modal conditions.

Also, changes in the xy dimensions was found to result in better performance than changes in either the x- or y-axis dimension alone with no difference between these uni-dimensional conditions. This shows, perhaps not surprisingly, that the more dimensions the two objects differ along, the easier it is to discriminate between the objects. What is interesting however is that there was no interaction between ISI and spatial differences. The effects of ISI were expected to be attenuated when the objects were more discriminable. It may be however that the experiment lacked sufficient power for effects to be observed.

These findings suggest that the visual and haptic modalities share the same memory characteristics. The results might also suggest evidence that objects are encoded into a common representation shared by both modalities and that performance reflects the characteristics of this multisensory memory. It cannot be concluded, however, that memory is multisensory without first comparing recognition performance between the modalities. Therefore a further experiment was conducted to investigate differences between the modalities.

2.3 Experiment 2

This experiment compared the matching performance between the visual and haptic modalities directly. In the previous experiment no effect of order of the encoding and comparison modalities in the matching task was found, suggesting that memory is the same across modalities. However, performance may also have been related to one memory system if, for example, sensory capture had occurred. Thus, in

all cases the representation of the object may have been recoded into visual memory in order for cross-modal recognition to have occurred. In this case, differences between visual and haptic matching tasks in terms of performance across time delays would be expected to be found. On the other hand, visual and haptic memory may indeed be similar, therefore the same effect of time delay would be expected on performance irrespective of modality. Finally, as in Experiment 1, the effect of time delay across modalities with objects of varying discriminability was tested to see if the representation of objects that are more easily discriminated would be more robust over longer time delays.

2.3.1 Method

2.3.1.1 Participants

24 undergraduate students (14 male and 10 female) from Trinity College Dublin participated in the experiment for course credits. Their ages ranged from 19 to 47 years, with a mean age of 24 years. All of the participants had normal or corrected-to-normal vision and none reported any somatosensory impairment.

2.3.1.2 Stimuli and Apparatus

An exact copy of the haptic set of stimuli was created in order to have identical, duplicate haptic objects for use in the 'same' trials. In all other aspects the stimuli were the same as those used in Experiment 1.

The apparatus was adapted slightly in order to be able to present two haptic stimuli at the same time. Two supporting structures were created, placed one behind the other, onto which haptic objects could be placed. The distance between these supports was 10cm which allowed for palpation of each object without interference

from the other. The centre of the computer monitor screen was now aligned with the midpoint between the two haptic stimuli.

2.3.1.3 Design

The experiment was based on a three-factor, within-subject design with modality (MODALITY), and ISI as factors. The SPATIAL factor was nested under these factors. For the MODALITY factor the levels were intra-modal visual (V-V) and intra-modal haptic (H-H). The levels for the ISI factor were again a delay of either 0, 15 or 30 seconds. Again differences along the x-axis, y-axis, or xy-axes were used together as levels in the spatial factor.

The dependent variables were again participant error rate (ERROR) and voice-activated response time (RT). All trials were randomly presented across the participants.

2.3.1.4 Procedure

With the following exceptions the procedure was the same as in Experiment 1. All trials were within modalities. Also, there was an inter-trial interval of 15 seconds for the H-H trials and 1.5 seconds for the V-V trials. During the haptic presentation of the stimuli, the participants were instructed to first feel the object closest to them. The second haptic stimulus was presented 10 cm behind the first and participants were instructed to reach behind the first object to feel the second. Haptic presentation was always 3 seconds and visual presentation was 0.25 seconds (as in Experiment 1). The second visual stimulus was presented in a location on the monitor that was offset from the location of the first stimulus by 1cm in a random direction. This was done to ensure that participants were not using apparent motion to match between the visual stimuli.

The number of trials was reduced overall from Experiment 1 in order not to over-burden the participants. Thus there were slightly different numbers of trials depending on the differing number of presentations of each of the x, y and xy dimensions across participants but the same number of trials was retained per level across the experiment. Trials were blocked per modality and the order of these blocks was counterbalanced across participants. There were 18 'different' trials and 18 'same' trials per modality yielding a total of 72 trials in the experiment. The experiment took approximately 45 minutes to complete.

2.3.2 Results

The mean of each participant's hit and correct rejection scores was calculated and a two-way, within subject ANOVA was conducted using MODALITY and ISI as factors (2.1% of the data was lost due to background noise). See Table 2 for the mean scores across the MODALITY and ISI factors. A main effect of MODALITY was found [$F(1,23)=9.959$, $p<0.01$]. Performance to the V-V modality was better than performance to the H-H modality. A main effect of ISI was also found [$F(2,46)=4.057$, $p<0.05$]. Posthoc Newman-Keuls analyses revealed that performance to an ISI of 0 seconds was better than performance to a delay of 15 seconds ($p<0.05$) and 30 seconds ($p<0.05$). Performance did not differ between ISIs of 15 and 30 seconds. There was no interaction between the MODALITY and ISI factors [$F(2,46)< 1$].

Table 2.

Table showing mean percentage correct responses (hit and correct rejection scores) and mean RTs across the different ISIs in each modality in Experiment 2. Standard error of the means is shown in parenthesis.

Modality Order	Vision - Vision			Haptics - Haptics		
ISI	0	15	30	0	15	30
% Correct Response	93.19	87.50	87.36	83.47	80.66	79.27
	(1.76)	(2.01)	(2.09)	(2.26)	(2.33)	(3.25)
RTs	951.5	959.7	1038.0	2095.2	2344.8	2481.8
(in ms)	(28.8)	(40.4)	(68.7)	(151.6)	(129.0)	(174.1)

The mean response times (hits and correct rejections) for each of the modalities across the different ISIs are shown in Table 2. A two-way within subjects ANOVA was conducted on the responses times across the MODALITY and ISI factors. A main effect of MODALITY was found [$F(1,23)=100.87$, $p<0.001$] with faster response times to the VV condition than the HH modalities. A main effect of ISI was also found [$F(2,46)=5.72$, $p<.01$]. Response times increased with longer inter-stimulus delays such that RTs to 0 second ISI were faster than to 30 seconds ISI ($p<0.05$). There was no difference in responses between either the 0 and15 second ISI or between the 15 and 30 seconds ISI. There was no interaction between the factors, although the interaction was approaching significance [$F(2,46)=2.778$, $p=0.073$].

A separate three-way ANOVA was conducted on the correct rejection responses with MODALITY, ISI and SPATIAL differences as factors. There was an effect of MODALITY [$F(1,22)=7.306$, $p<0.05$] with better performance to the VV condition than the HH condition. An effect of ISI was also found [$F(2,44]=3.777$, $p<0.05$]. A post-hoc Newman-Keuls analysis revealed that performance to the 0 second ISI was significantly better than performance to an ISI of 30 seconds ($p<.05$). There was no difference between 0 and 15 seconds ISI and between 15 and 30 seconds ISI. A main effect of SPATIAL differences was found [$F(2,44)=9.214$, $p<.0005$]. A posthoc Newman-Keuls analysis on the SPATIAL effect revealed that performance was better to differences between the xy axes than to differences between either the x ($p<0.01$) or y-axes ($p<0.01$) and there was no difference between responses to the x- and y-axes.

Interestingly, an interaction was found between SPATIAL differences and ISI [$F(4,88)=12.200$, $p<0.001$]. This interaction is depicted in Figure 6(a). Post-hoc Newman-Keuls tests revealed that there was no effect of ISI with changes in the xy-axis dimensions. For changes in the x-axis dimension, performance to the ISI of 30 seconds was significantly worse than performance to both 0 seconds ($p<0.001$) and 15 seconds ISI ($p<0.001$) and no difference between 0 and 15 seconds ISI. Finally, for changes in the y-axis dimension only, performance was worse to an ISI of 15 seconds relative to ISIs of 0 seconds ($p<0.05$) and 30 seconds ($p<0.05$) with no difference between 0 and 30 seconds ISI.

A significant interaction was found between MODALITY and ISI [$F(2,44)=3.856$, $p<0.05$]. This interaction is depicted in Figure 6(b). Post-hoc Newman-Keuls tests revealed that there was no difference between the ISI conditions in the visual modality but in the haptic modality performance was significantly worse for ISI of 15 seconds and 30 seconds relative to 0 seconds ($p<0.05$). The MODALITY and

SPATIAL interaction was not significant [$F(2,44)<1$]. The three way interaction was significant [$F(4,88)=3.412$, $p<0.05$].

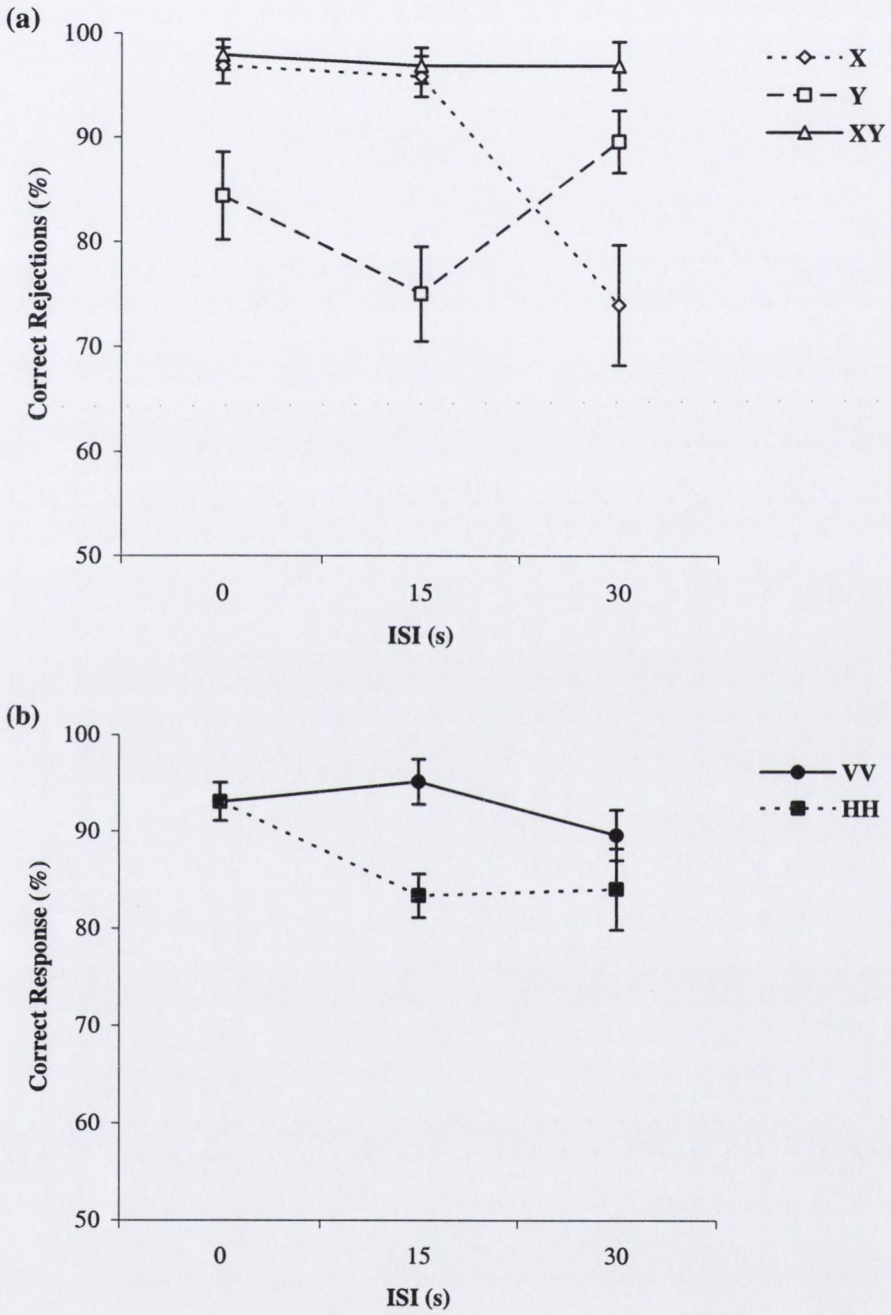


Figure 6. Plot showing percentage correct rejection scores for each spatial change (a) and modality (b) across the different ISIs in Experiment 2.

2.3.3 Discussion

This experiment found that the effect of delay is the same for visual matching tasks as for haptic matching tasks. Overall performance was better when there was a short delay between the stimuli than when there was a longer delay and this effect was not different across the modalities. This finding suggests that the effect of delay found in the previous experiment across modalities was not due to any differences between the modalities since both show the same effects of ISI.

An interesting effect was also observed with respect to the correct rejection responses across the different spatial dimensions. The results show that when objects are more discriminable they are also less affected by delay. Only differences in the x- and y-axis dimensions were affected by ISI whereas differences in the xy-axis dimensions together were not affected by ISI, even as long as 30 seconds. It is suggested that these differences in the effect of time delay across object discriminability might account for the inconsistencies in the literature on the effects of delay for object matching tasks. This finding also makes intuitive sense; if the task is to discriminate between two objects from different categories (e.g., a pen and a coffee mug) then you might expect this task to be less sensitive to delay simply because many feature differences may be more robust to time delays. On the other hand, a task involving objects from within the same category (e.g., between two different coffee mugs) may need to rely on small differences between the objects and may therefore be more sensitive to time delay. Previous studies have found that the recognition of highly similar objects is not only more difficult but is also more sensitive to small image changes such as changes in viewpoint (Edelman, 1995; Newell, 1998).

2.4 Experiment 3

Experiment three tested simultaneous presentation against successive presentation of objects across modalities. In the real world, an object of interest is encoded at the same time across the visual and haptic modalities. Thus, temporal synchrony might be key to integrating information across the modalities into a single multisensory representation. If objects are represented into the same multisensory memory system as a result of simultaneous presentation then you would expect performance in a matching task to be worse for simultaneous presentation than for successive presentation. The reasoning is as follows; if simultaneous presentation results in mandatory fusion across the modalities then the consequence of this is that two different objects might be represented as one. Therefore, it would be more difficult to discriminate objects that are different. On the other hand, separate memory systems may still exist independently of presentation timing. In this case then, objects should be equally discriminable irrespective of presentation order. As in the previous experiments the effects of spatial differences across the presentation conditions were also tested for.

2.4.1 Method

2.4.1.1 Participants

22 undergraduate students (7 male and 15 female) from Trinity College Dublin participated in the experiment for course credits. Their ages ranged from 18 to 28 years, with a mean age of 20 years. All of the participants had normal or corrected-to-normal vision and none reported any somatosensory impairments.

2.4.1.2 Stimuli and Apparatus

The stimuli and apparatus were the same as in Experiment 1.

2.4.1.3 Design

The experiment was based on a two-factor, within-subject design with PRESENTATION order of the modalities and SPATIAL differences as factors. Unlike in Experiment 1, here there were three levels to the PRESENTATION ORDER factor; VH, HV and simultaneous presentation. The SPATIAL levels were the same as those described in Experiment 1 (i.e., differences in x-, y-, and xy-axis dimensions). Trials were blocked according to the PRESENTATION ORDER levels and the sequence of the blocks was counter-balanced across participants. The dependent variables were participant error rate (ERROR) and voice-activated response time (RT).

2.4.1.4 Procedure

With the following exception the procedure was the same as that outlined in Experiment 1: both the visual and haptic stimuli were presented for 3 seconds during the experiment. Thus, in the simultaneous condition, the stimuli were presented for 3 seconds in total. An ISI of 0 seconds was used for the VH and HV trials and compared performance in these conditions to simultaneous presentation of the stimuli across modalities. Participants' eye movements were informally monitored by the experimenter during the simultaneous condition to ensure that participants were simultaneously viewing whilst touching the objects. The experimenter was afforded a direct view of the participant's eyes as he stood in front of the participant next to the monitor during the trials. If the participant were not looking at the visual stimulus at the correct time it would be obvious to the experimenter that this was the case. It was also obvious to the participant that the experimenter could tell whether or not the task was being completed as instructed, which possibly acted as an incentive for the participant to perform the task correctly.

There were 108 trials in the experiment; 36 VH, 36 HV and 36 simultaneous trials (there were 18 'same' trials and 18 'different' trials for each of these blocks). To familiarise the participant with the procedure, 3 practice trials were given at the start of each block of the PRESENTATION ORDER condition. In all other ways the procedure was the same as described in Experiment 1. The experiment took approximately 80 minutes to complete.

2.4.2 Results

As in the previous experiments, the mean of each participant's hit and correct rejection scores was calculated (5.6% of the data was lost due to background noise). The mean percentage correct (i.e., hits plus correct rejections) scores for the different PRESENTATION ORDER conditions are given in Table 3. A one-way repeated measures ANOVA was conducted with PRESENTATION as the factor (VH, HV and simultaneous). No main effect of PRESENTATION was found [$F(2,42)=1.181$, n.s.].

Table 3.
Table showing mean percentage correct responses (hit and correct rejection scores) and mean RTs across the different presentation order conditions in Experiment 3. Standard error of the means is shown in parenthesis.

	Presentation type		
	SIM	VH	HV
% Correct Response	84.96 (1.56)	86.11 (1.39)	83.81 (1.62)
RTs (in ms)	2275.4 (130.0)	2151.0 (151.8)	951.1 (35.2)

Table 3 also shows the mean response time performance across the presentation conditions. A one-way ANOVA was conducted on the RT data and found a main effect of PRESENTATION [$F(2,42)=79.335$, $p<.001$]. A post-hoc, Newman-Keuls analysis revealed that performance to the HV condition was significantly faster than to the VH ($p<0.01$) or the simultaneous condition ($p<.001$). There was no difference between the HV and simultaneous conditions. This relatively fast performance in the HV condition is not surprising because participants were faster at encoding and responding to the test stimulus when it was presented as a visual rather than a haptic stimulus.

A separate two-way ANOVA was then conducted on the correct rejection responses with SPATIAL differences and PRESENTATION ORDER conditions as factors. A main effect of SPATIAL differences was found [$F(2,42)=49.827$, $p<0.001$]. The mean percentage performance for each of the SPATIAL differences was as follows: 72.7% (1.9% standard error) accuracy for x-axis differences; 70.9% (2.3% S.E.) accuracy for y-axis differences and 93.4% (1.5% S.E.) accuracy for xy-axis differences. Performance to the changes in the xy-axis dimension was significantly better than performance to changes in either the x or y dimensions (Newman-Keuls, $p<0.05$). There was no effect of PRESENTATION ORDER [$F(2,42)=2.272$, n.s.] and no interaction between the factors [$F(4,84)=0.829$, n.s.].

2.4.3 Discussion

The finding that there was no difference in performance between simultaneous and successive presentation of objects in the cross-modal matching task suggests that the modalities may work independently and allow for efficient comparison across objects. If objects were stored as multisensory representations then performance would have been expected to be worse for simultaneous presentation than successive presentation. No differences were found between the presentation conditions. This finding suggests that the representations of the objects within each modality are sufficiently discrete to allow for cross-modal comparison, even when presented at the same time to the different modalities.

No difference was also found between the spatial conditions across each of the presentation orders. That is to say, differences in x-, y- or xy-axis dimensions were not any more or less difficult when presented simultaneously than when presented in succession. Overall, however, as in the previous experiments, the results show that

performance to differences in the xy-axis was better than to differences in x- and y-axis dimensions separately.

2.5 General Discussion

The differences in matching performance across inter-stimulus time delays between vision and haptics were investigated. Experiment 1 found that cross-modal matching performance decreased with an increase in time delay between one object stimulus and the next. Furthermore, the rate of performance decrease was the same irrespective of the modality to which the first stimulus was presented. This result suggested that both the visual and haptic modalities share the same memory characteristics in terms of the effects of delay between stimulus presentations. This suggestion was tested directly in Experiment 2 by comparing performance with increasing ISI across each of the modalities. Again no difference was found in the effect of ISI across the modalities, although visual performance was better overall than haptic performance. Finally, Experiment 3 investigated whether the findings in the previous experiments were due to a single multisensory representation of objects or whether separate, modality specific representations are maintained. If objects were represented in a single multisensory representation with temporal synchrony, it was predicted that simultaneous presentation of object stimuli across the modalities would result in a cost in matching performance in comparison to successive presentation of the objects. No difference between simultaneous and successive matching across modalities was found suggesting that representations are sufficiently modality specific. It is concluded that objects are mainly stored in separate modalities that share similar functional properties of their memory systems. This similarity in properties may allow for efficient access between object representations across the modalities.

In all the experiments objects which differed by more than one dimension were found to be more discriminable. This is perhaps not a surprising finding. More interesting, it was found that in each modality there was no effect of inter-stimulus delay on objects that were more discriminable than objects that changed on either the x- or y-axis only (Experiment 2). It was suggested that inter-stimulus similarity may affect performance over different delays and that this may explain the inconsistencies in the literature between effects of delay on cross-modal recognition tasks (Garvill & Molander, 1973; Kiphart et al., 1992). Spatial differences, however, interacted with inter-stimulus intervals in the within-modal experiment only (Experiment 2) and not in the cross-modal experiment (Experiment 1), although changes in the xy-axis resulted in better performance in both experiments. Experiment 1 failed to find a main effect of ISI on the correct rejections, and this result may have obscured any interactions with the spatial factor. Given that an effect of ISI was found with both hits and correct rejection responses together, it certainly seems that for an analysis using only correct rejection responses the experiment lacked sufficient power for these effects to emerge.

In Experiment 3 an interference effect with simultaneous presentation was expected to be found if the objects were stored as a single, multisensory representation due to temporal synchrony. Instead no difference were found between presenting stimuli to the modalities simultaneously or successively. It was argued that this finding suggests that modality specific representations might be maintained such that comparisons between the stimuli can occur. However, it might still be the case that temporal synchrony across the modalities results in a multisensory representation. In Experiment 3, interference effects may have been avoided because of the difference in the spatial location between the visual and haptic stimuli or a difference between the stimuli themselves, irrespective of the timing of the presentation of the stimuli.

Generally, a single object of interest will not only be simultaneously perceived by the modalities but it will also occupy the same spatial location and share the same spatial characteristics. Thus, in the experiment the objects may not have been integrated because of spatial disparity. The role of spatial disparity in object integration across the modalities is currently being investigated.

The finding that visual and haptic memory for objects share similar functional properties fits with current neuroimaging literature demonstrating similar cortical mechanisms for visual and haptic object recognition (Amedi et al., 2001; James et al., 2002). One interesting observation about these imaging studies is that activation to haptic object recognition resides in a visual area (LOC) implicated in visual object recognition and not in a somatosensory area. This begs the question of whether haptic objects are 'recoded' into visual co-ordinates and stored as visual representations. If so, then these findings may mean that haptic memory for objects is the same as visual memory because objects are recoded into visual memory. Indeed in Experiment 2 within-modal haptic memory was found to be less efficient than within-modal visual memory suggesting that a recoding was occurring in the previous cross-modal experiments (since there was no differences between the modality orders). Furthermore, the findings from Experiment 3 suggest that visual recoding of haptic stimuli is not mandatory since no interference was found with simultaneous, cross-modal presentation. Neuropsychological literature also suggests that haptic representations are not necessarily recoded into visual co-ordinates since impaired memory for recognition in one modality is generally dissociated from performance in the other modality (Farah, 1990; Reed, Caselli, & Farah, 1996). Nevertheless, visual recoding seems to be a good candidate model for cross-modal efficiency although this may depend on the task. Certainly, in tactile tasks where participants are encouraged to

use non-informative vision or a visually determined reference frame, then judgements of spatial differences improve in touch (Millar & Al-Attar, 2004; Newport, Rabb, & Jackson, 2002). One way to test the contribution of visual memory in cross-modal object recognition tasks is to compare matching performance across ISIs in congenitally blind participants and sighted participants. If haptic memory performance is not affected by visual experience then it would be expected that there would be no difference between the participant groups. Such studies can further our understanding of the role of visual processing in haptic and cross-modal object recognition.

CHAPTER 3: THE INVOLVEMENT OF SPATIAL FACTORS ON OBJECT INTEGRATION

Abstract

This chapter was concerned with the effects of spatial factors on cross-modal integration. In the preceding chapter it was suggested that cross-modal objects were not integrated as they were not co-located in space. In Experiments 4 & 5, visual and haptic objects were either co-located in space or were vertically or horizontally spatially separated. It was expected that the percepts of different co-located objects, if integrated,, would be harder to determine than spatially separated objects. However, there was no evidence that co-located objects were mandatorily cross-modally integrated. Increasing the assumption of unity of the stimulus-display in Experiment 5 also did not promote cross-modal integration. Experiment 6 tested whether potential differences in object size prevented integration, but found this not to be the case as visual and haptic objects were found to be perceptually equivalent in size. Object percepts may also have differed from each other as visual and haptic encoding was not equivalent. When the modalities were matched in terms of encoding however (haptic encoding was made optimal), objects were still not integrated (Experiment 7). Experiment 8 found that the point at which visual and haptic objects were co-located was biased, suggesting that perhaps objects in previous experiments were integrated, but not at the location where initially thought. So, although there was some indirect evidence that visual and haptic objects were indeed mandatorily integrated,, it seems more likely that some unidentified factor acted to prevent objects being integrated in these experiments.

3.1 Introduction

The preceding chapter discussed the suggestion from the results that representations of objects perceived simultaneously by vision and touch do not combine into a single percept, and it was concluded that visual and haptic representations are stored independently. However, it may be that these representations are not independent after all: a difference of spatial location between the visual and haptic objects may have prevented the integration of the object percepts across modalities. Previous studies have suggested that the co-location of objects is an important factor in cross-modal integration (Ernst & Bülthoff, 2004; Làdavas, Farnè, Zeloni, & di Pellegrino, 2000; Lalanne & Lorenceau, 2004; Soto-Faraco, Lyons, Gazzaniga, Spence, & Kingstone, 2002; Spence et al., 2001; Stein & Meredith, 1993; Wallace et al., 2004). The aim of this chapter is to investigate whether it is this common location of objects across the senses which permits object information to be integrated. If objects which are co-located across modalities have their percepts integrated together, it would be expected that two different objects presented at the same spatial location but encoded by different modalities would be harder to differentiate as their percepts should be mandatorily integrated. On the other hand, the percepts of spatially separated objects should not be integrated together, and it should be easier to determine they are different objects.

There is some support for an effect of co-location on multisensory integration in the literature. Miller (1972) suggests that different objects which are co-located across the senses may be mistaken as a single object. Also, Zampini, Shore, & Spence (2003 p. 202) state that "The likelihood that two stimuli from different sensory modalities will be bound into a single coherent object representation is greater when they come from the same rather than from different locations." However, as far as this author is aware,

no one has directly tested to see whether complex multisensory objects will integrate together when they are co-located in space.

The effects of spatial co-location have been investigated for very simple visual and haptic stimuli. For example, in a temporal order judgement study, Spence, Shore & Klein (2001) found that visual flashes and tactile vibrations were required to be separated in time by longer intervals for their order of presentation to be determined than for stimuli which were spatially separated (86ms versus 37ms). Presumably, the co-location of these stimuli allowed them to be integrated together, making their temporal order more difficult to discriminate. Other studies investigated the effects of spatial co-location on integration across other modalities, such as audition and vision. Soto-Faraco et al. (Experiment 2, 2002) were interested in the integration of dynamic moving information between vision and audition and found automatic and strong integration effects when stimuli were both spatially and temporally aligned, compared to when stimuli were spatially separated by only 10cm or 50cm (i.e., more movement discrimination errors were made when stimuli were spatially co-located compared to when they were not co-located). It is interesting however that despite a spatial separation, Soto-Faraco et al. still found some evidence for an integration effect, suggesting that there may be a spatial window within which cross-modal objects may be integrated, albeit not optimally. For more complicated auditory-visual stimuli, spatial co-location has been found to be even less important in cross-modal integration (e.g., Jack & Thurlow, 1973). Jack and Thurlow (1973) used auditory speech and synchronized visual lip movements as stimuli and varied the spatial separation between stimuli. However the authors found strong location-bias effects for stimuli separated by upto 30 degrees. Zampini et al., (2005) suggest that the influence of co-location on cross-modal integration effects is quite dependent on the modalities which encode

information. It will be interesting to see whether, as predicted, that relatively complex visual and haptic stimuli that are co-located in space will be integrated together. If this is so, this would be evidence that co-location is an important factor in the cross-modal integration of visual and haptic information.

3.2 Experiment 4

In this experiment the location of haptic objects relative to the position of a visual object was varied. The visual objects presented along the participant's midline, at eye level, and were projected onto the same fronto-parallel plane as the haptic objects. By using mixed reality (where virtual and real objects are simultaneously presented to the observer), it was possible to present the haptic objects in such a position that the visual objects were superimposed on to the position of one of the haptic objects. Here, it was crucial that the visual and haptic objects were perceived only by their respective modalities. By hiding the haptic stimulus-display from sight, vision was prevented from directing haptic search (Klatzky, Lederman et al., 1993). Crucially, this manipulation also prevented vision from perceiving the haptic objects particularly when they were spatially separated from the visual object.

It is important to point out however, that when using mixed reality as suggested above, that the haptic stimuli are real distal objects whilst visual stimuli are, depending on the technology used, rendered 2D or 3D representations of real 3D objects (in the experiments here, visual stimuli were always 2D). It is possible that this discrepancy could influence whether or not visual and haptic object percepts would be integrated or not. For example, this discrepancy could alert the observer that two different objects are being explored, which potentially could inhibit cross-modal integration (Welch & Warren, 1980).

In the following experiment visual and haptic objects were either co-located or the haptic objects were spatially separated from the visual object. It was predicted that cross-modal integration was likely to be maximal when objects were co-located, however, as mentioned in the introduction, there may be a window within which integration may still occur. Two main directions of mislocation were tested and the distance between the visual and haptic objects was manipulated. Half of the spatially mislocated objects were separated along the vertical axis and half were separated along the horizontal axis. Participants performed a cross-modal matching task where they had to match the shape of the visual object with the shape of the haptic object. On half the trials the visual and haptic objects were the same shape and for the other half of object pairs the objects were of a different shape. The dependent variable was cross-modal matching performance across co-located and mislocated positions of the objects. The hypothesis was that if cross-modal integration relies on objects being co-located, then co-located objects will be difficult to discriminate when they were of a different shape, relative to mislocated objects. It may also be the case that matching performance would improve with greater separation between the location of the visual and haptic objects along both the horizontal and vertical axes.

3.2.1 Method

3.2.1.1 Participants

15 undergraduate students (9 male and 6 female) from Trinity College Dublin participated in the experiment for course credits or payment (8 euro per hour). Their ages ranged from 20 to 27 years, with a mean age of 23.1 years. 14 of the participants reported to right-handed, all had normal or corrected-to-normal vision and none reported any tactile impairment.

3.2.1.2 Stimuli

Six suitable haptic object stimuli were randomly taken from the complete set used in Experiment 1 and were used in this experiment. Six visual objects were also randomly taken from Experiment 1. Visual and haptic objects were paired together: three of visual and haptic object pairs were identical in shape and three pairs differed in shape. One of these pairs differed by 2cm[5] along the x-axis arm; another differed along the y-axis arm by 2cm; the last by 2cm along both arms. Visual stimuli were displayed through a head mounted display (HMD) and were scaled in order to match the haptic objects in size. This scaling was determined in a pilot study (N=5), in which participants adjusted the size of visually presented objects to exactly match the size of haptic stimuli. The average subjective scaling across these participants was used to resize the visual objects in the present experiment. The largest object in the visual set when viewed through the HMD occupied approximately 9 degrees of visual angle.

3.2.1.3 Apparatus

The apparatus was designed so that the participants could not see the haptic apparatus or the haptic stimuli, just their virtual counterparts. Visual stimuli were presented on a HMD which was fixed in place 42cm above a table. The stimuli were always perceived with binocular vision. Problems associated with eye dominance were avoided as the visual stimuli were presented in 2D- information encoded by both eyes was thus identical. Haptic stimuli were presented on a magnetic board located 30cm from the HMD. Each haptic stimulus was attached to a moveable magnetic disk that could be displaced to any location on the board (see Figure 7). A haptic object could be

[5] From the results of Experiment 1it was found that objects which differed by 1cm were found to be difficult to discriminate (37.5% accuracy), whilst objects differing by 3cm were quite easy to discriminate (90.1% accuracy). Objects that differed from each other by 2cm were found to lead to a discrimination accuracy of 76.1%. It was therefore decided to use differences of 2cm in order to avoid ceiling and floor effects.

attached to the disk by fitting it onto two prongs that protruded from the magnetic disk. The disk was 20cm in diameter and was approximately 1cm thick. Nine different locations for the haptic objects were marked out on a reference card attached to the board. For each participant, at the start of the experiment the position of the card on the board was carefully adjusted so that the centrally located position on the card (labelled "middle" in Figure 7) occupied the exact location of the projected visual stimulus. For each participant, the card's position thereafter remained fixed for the rest of the experiment. At the subjective central location objects were presented at eye level along the participant's body-midline. Of the remaining 8 positions on the card, 4 were horizontally separated from the central position in 10cm steps, and four were similarly vertically separated forming an overall '+' configuration of positions. Before the experiment, participants adjusted the height of their chair so they could comfortably view the visual stimuli through the HMD for the entire duration of the experiment.

The HMD (i-Glasses SVGA 2D) viewing range was 26° diagonal, its resolution was 800 x 600 pixels and the frame rate was 120 Hz. The experiment was run using DMDX software for the PC (Forster & Forster, 2003) and participant responses were recorded through two custom made foot pedals.

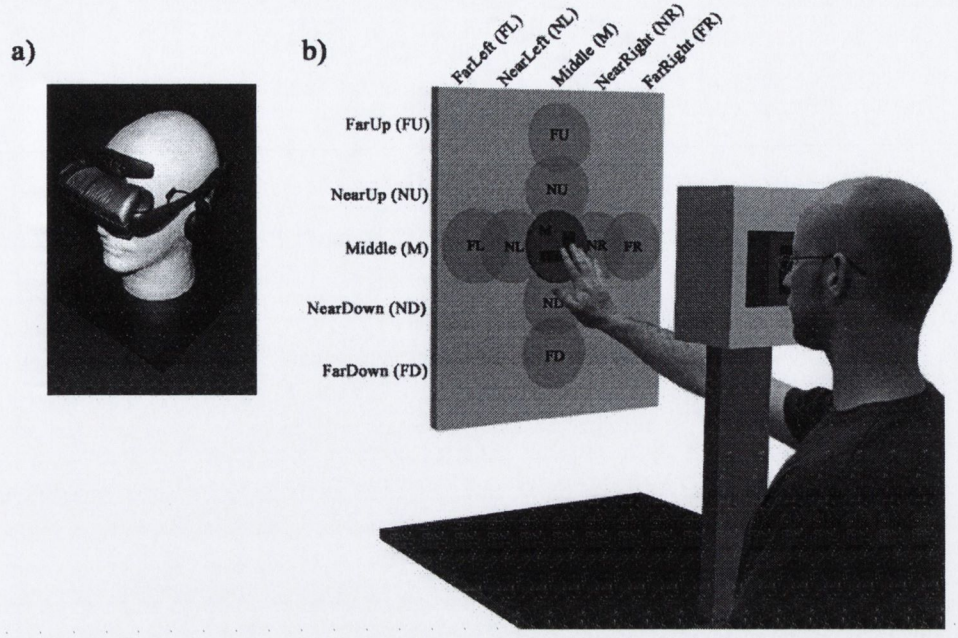


Figure 7. a) A photo of the HMD. Straps were used to attach the HMD to the participant. b) An illustration of the apparatus used in Experiment 4. Here the participant viewed a visual object through a HMD (note participant depicted is not wearing the HMD) whilst simultaneously freely exploring a haptic stimulus with both hands (one handed exploration is seen in the figure to make the depiction simpler). The participant was required to judge whether these shapes were the same shape or different. The red circle highlights the position where the visual and haptic objects were subjectively co-located in space). Other possible locations for the haptic stimuli are shown as labelled green circles in the illustration.

3.2.1.4 Design

The experiment was based on a one-factor, within-subjects design. The independent factor was the position of the haptic object (POSITION). There were 9 levels of POSITION, each corresponding to the different locations the haptic objects could occupy. The dependent variables were participant correct response rate (HITS and CORRECT REJECTIONS) and response time (RT).

There were a total of 54 trials in this experiment (9 positions x 3 repeats x same\different). These trials were presented randomly across participants. 50% of the trials included a mismatch of object shapes (as specified in the *stimuli* section).

3.2.1.5 Procedure

The experiment was based on a same/different matching paradigm where the participant's task was to decide whether two stimuli, presented simultaneously across modalities, were the same or different as fast and as accurately as possible.

Prior to the experiment the location of the central point of the haptic display was moved until it was subjectively aligned with the centre of the visual display. This was achieved by asking each participant to carefully adjust the position of a haptic object (located on the card in position 'middle') so that it exactly matched the position of an identically shaped visual object. Moving the haptic object moved the disk to which it was attached, which in turn moved the reference card. Once the participant was satisfied that visual and haptic stimuli were co-located in space, the reference card was anchored into a fixed position. This prevented further accidental movement of the haptic stimulus during the experiment.

In the experiment, participants were required to locate the position of the haptic object by first finding the position of the disk on the board using two hands. They were instructed to locate the position of the disk within 2.5 seconds (this was ample time to locate the disk) and then rest their hands on the rim of the disk and await further instruction. After haptic exploration, a 0.5 second fixation cross was displayed through the HMD. As soon as the fixation cross disappeared, the visual object was presented and the participants had to start feeling the haptic object at the onset of the visual stimulus. To provide a further cue to the participant to start bi-manually feeling the haptic object, a tone was sounded when the fixation cross disappeared. After 3 seconds the tone stopped and the visual object disappeared, indicating to the participant to remove their hands from the haptic stimulus. Participants responded by pressing the right foot pedal if they thought the objects were matched in shape and the left foot pedal

if they thought the objects' shapes were different. The inter-trial interval (ITI) was approximately 10 seconds in duration, during which the experimenter changed the haptic object and placed the disk in the correct spatial position for the next trial.

10 practice trials were given at the beginning of the experiment. The experiment took approximately 45 minutes to complete.

3.2.2 Results

The mean percentage HIT and correct rejection (CR) scores for the different positions of the haptic objects are listed in Table 4. A two-way, within-subjects ANOVA was conducted using POSITION as one factor (9 levels) and TRIAL TYPE (HIT or CR trials) as the second factor. There was an effect of POSITION [$F(8,112)=2.04$, $p<0.05$], and this is shown in

Figure ***8***. A Posthoc Newman-Keuls analysis found that the combined correct response scores for the different positions did not differ from each other. However, although non-significant, correct response scores for matching objects at middle position were higher (i.e., participants were more sensitive to matching objects) than scores at the near-right position ($p=0.064$). In turn, scores to the near-right position were lower than the far-up position ($p=0.054$) and the far-down position ($p=0.080$). All other positions did not differ from each other.

There was no effect of TRIAL TYPE [$F(1, 14)<1$] and there was no interaction between factors [$F(8,112)<1$].

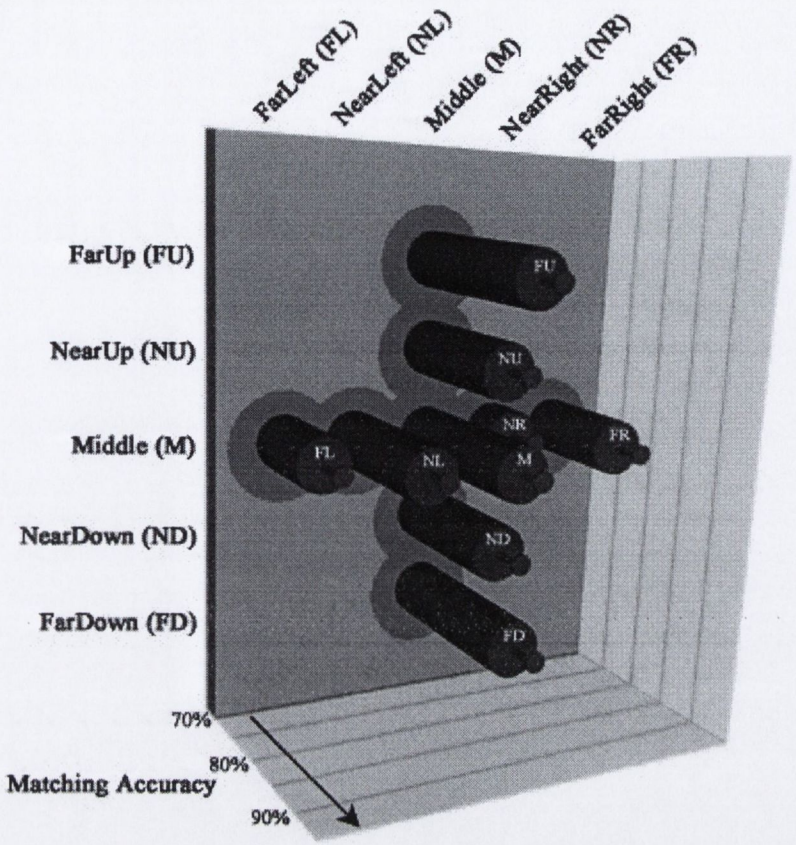


Figure 8. Plot depicting average correct matching performance scores for each location at which haptic objects were presented. The greater the size of the cylinder at each position at which haptic objects were presented, the greater the matching performance accuracy. Error bars (in red) depict the standard error of the mean.

Table 4.
Mean and standard error scores (in brackets) for HITS, CR (correct rejections) and RT (average of the response times for HITS and CR trials) for each level of POSITION in Experiment 4.

Position	% HITS (StdError)	CR (StdError)	RT in seconds (StdError)
Middle	91.7 (±3.9)	88.9 (±5.4)	3.19 (±0.27)
Far-up	91.7 (±3.9)	91.7 (±5.3)	3.31 (±0.26)
Near-up	77.8 (±5.4)	94.4 (±5.1)	3.27 (±0.31)
Near-down	94.4 (±5.3)	80.6 (±5.5)	3.41 (±0.27)
Far-down	86.1 (±5.4)	91.7 (±5.1)	3.23 (±0.23)
Far-left	80.6 (±5.3)	80.6 (±6.4)	3.28 (±0.30)
Near-left	86.1 (±4.4)	88.9 (±4.4)	3.20 (±0.28)
Near-right	83.3 (±6.9)	72.2 (±5.8)	3.23 (±0.23)
Far-right	88.9 (±4.4)	75.0 (±6.2)	3.45 (±0.26)

Mean reaction times were computed by averaging RT's for correct responses only for each level of POSITION. These are given in Table 4. RT was entered as the dependent factor in a one-factor, within-subject ANOVA with POSITION as the dependent variable. There was no effect of POSITION [$F(1,8)=0.92$, n.s.].

3.2.3 Discussion

Participants were not less sensitive to object changes when objects were co-located, which suggests that co-located objects are not mandatorily integrated by the brain. However, there are a number of reasons why matching performance to objects in this position is not reduced relative to mislocated objects. In Experiment 5, 6, and 7, a number of reasons were examined why the predicted effects were not found here. Firstly, the visual and haptic objects may have been too dissimilar in appearance that integration did not occur (Experiment 5). It was also vital to test whether the visual and haptic objects were actually perceived as identical in size since size differences may have cued the difference between objects (Experiment 6). There may have been also a difference in optimal encoding between the visual and haptic objects thus preventing the objects from being integrated across modality (Experiment 7).

In summary, the location of the haptic objects did seem to influence sensitivity to matching object shape across modalities. Although non-significant, participants were less sensitive to matching objects when haptic objects were located to the near-right of co-location relative to when objects were co-located, or the haptic object was positioned in the far-up or far-down locations. It was speculated that this result reflects a bias in the subjective co-location of visual and haptic objects. This effect of horizontal mislocation on matching performance to cross-modal objects and the subjective judgements of the co-located position is explored further in Experiment 8.

3.3 Experiment 5

It seems unlikely that the lack of a predicted integration effect in Experiment 4 would imply that spatial co-location and temporal congruency are not necessary for integration as there is a consistent view in the literature that these factors are important in integration (e.g., Lalanne & Lorenceau, 2004; Stein & Meredith, 1993; Wallace et al., 2004). Much more likely is that integration did not occur in Experiment 4 as some additional requirement was not met. For example, objects may have to be very similar for integration to occur (Welch & Warren, 1980). However, in the previous experiment, the objects were quite different from each other. The present experiment is a repeat of Experiment 4, but the visual and haptic objects were rendered to be more similar to each other in appearance. It is also possible that the predicted integration effect in Experiment 4 may have been obscured in some way. For example, experimental power may have been insufficient in Experiment 4. Here experimental power was increased by including more trials and reducing the number of spatial locations.

In natural object recognition, both vision and touch generally explore the same object. However, in Experiment 4 the visual and haptic objects differed as the haptic objects were 3D, whilst the visual objects were 2-dimensional graphical images. Even though in the same trials visual and haptic objects were exactly matched according to shape, it may be the difference in appearance that acted to prevent the predicted effects of integration in Experiment 4. If this were so, a likely candidate is that some top-down process may have acted to inhibit integration, preventing object percepts of apparently different co-located objects from being integrated together.

Welch & Warren's (1980) model of intersensory bias suggests two reasons why such a difference of appearance would prevent integration. First, the difference in appearance would add to the number of discrepant properties encoded about the objects across modalities. According to the intersensory bias model, integration occurs when the sum of equivalent cross-modal object properties minus the sum of non-equivalent object properties across modalities exceeds some constant. The difference between the stimuli in the present experiment may act to reduce the likelihood of integration, because the assumption of unity across the modalities may be violated. This could explain the lack of an integration effect in Experiment 4.

Second, the stimuli used in the present experiments are unfamiliar. Welch (1999) argued that if objects that are normally associated across modalities or if the stimulus situation is compelling (Welch & Warren, 1980), then cross-modal integration occurs. For audio-visual integration for example, Jackson (1953) found that more appropriate visual-auditory stimuli pairings, such as the sight of steam from a steam-kettle and the spatially separate sound of a steam-puff, led to a much stronger ventriloquist effect than for more inappropriate pairings, such as the sight of a stationary bell paired with a ringing noise. Thurlow & Jack (1973) also found stronger ventriloquist effects when visual stimuli were more appropriate to the sound source. The authors found a strong effect of ventriloquism for the sound of a tone (AT) coupled with a video of a finger pressing a button (VB), and also for the speech of a man counting (AS) coupled with a video of the man counting (VC). Conversely, when stimuli pairs were made inappropriate by swapping the stimuli pairs (i.e., AT-VC & VB-AS), the observed ventriloquist effect was much reduced. However, the effect of compellingness on ventriloquism was not replicated in a similar study by Radeau & Bertelson (1977), who coupled the sound of bongo drums with either a video of the

hands of a person playing a drum (the realistic condition) or a video of diffuse light which was modulated by the sound. The authors found no difference between these conditions. Although there is mixed evidence on the effects of compellingness for visual-auditory ventriloquism effects, it seems quite likely that effects of compellingness would play a role in visual-haptic cross-modal integration. This is suggested as, prima-facie, discrepant visual-haptic pairings occur less often in nature than do discrepant visual-auditory pairings (we can usually tell if a visual-haptic pairing is discrepant as we can see our hands whilst we haptically and visually explore objects. However there is no visual-auditory analogue of this useful "sight-of-hand" cue). In the following experiment the stimulus-display was rendered to be compelling as possible in order to avoid the problem of unfamiliar or uncompelling cross-modal associations.

As mentioned already, it is also possible that integration may have occurred despite these factors, but that the effects were hidden due to either a relatively high error rate for matching objects or for low statistical power. For example, there were only 3 different stimulus pairs in Experiment 4. In the following experiment the number of stimulus-pairs was increased to avoid issues associated with using small stimulus sets. Also, because of the low number of stimuli participants may have been overexposed to each stimuli (haptic objects were each presented 9 times, once at each of the 9 positions, potentially giving participants time to become quite familiar with the objects). The number of positions at which the objects are presented were also reduced in this experiment.

In order to address the issue of compellingness digital photographs of the haptic objects were used as visual stimuli, and were presented to the participant through the HMD. The number of stimuli were doubled and each stimulus was presented fewer times than in Experiment 4 by restricting the positions in which visual and haptic

objects were spatially different to just along the horizontal axis. These differences aside, this experiment and Experiment 4 are identical. The prediction again is that differently shaped object-pairs that are similar in appearance and are co-located will be harder to discrimate compared to cross-modal object-pairs that are not mislocated.

3.3.1 Method

3.3.1.1 Participants

24 undergraduate students (9 male and 15 female) from Trinity College Dublin participated in the experiment for course credits or payment (8 euro per hour). Their ages ranged from 18 to 46 years, and the mean age was 22.7 years. 18 participants reported to be mainly right-handed. All of the participants had normal or corrected-to-normal vision and none reported any tactile impairment.

3.3.1.2 Stimuli and Apparatus

The apparatus and the haptic stimuli were the same as those used in Experiment 4. However, visual stimuli here were made from digital photographs of haptic stimuli that were presented to the participant through the HMD. Visual stimuli resized so that they were identical in size to the haptic stimuli (i.e. they were identical in size to the visual stimuli in Experiment 4). Here, the position of the haptic objects was restricted to the horizontal axis on the magnetic board, as shown in Figure 9. Also, as the number of object stimuli were doubled in this experiment, a total of 12 object shapes was used.

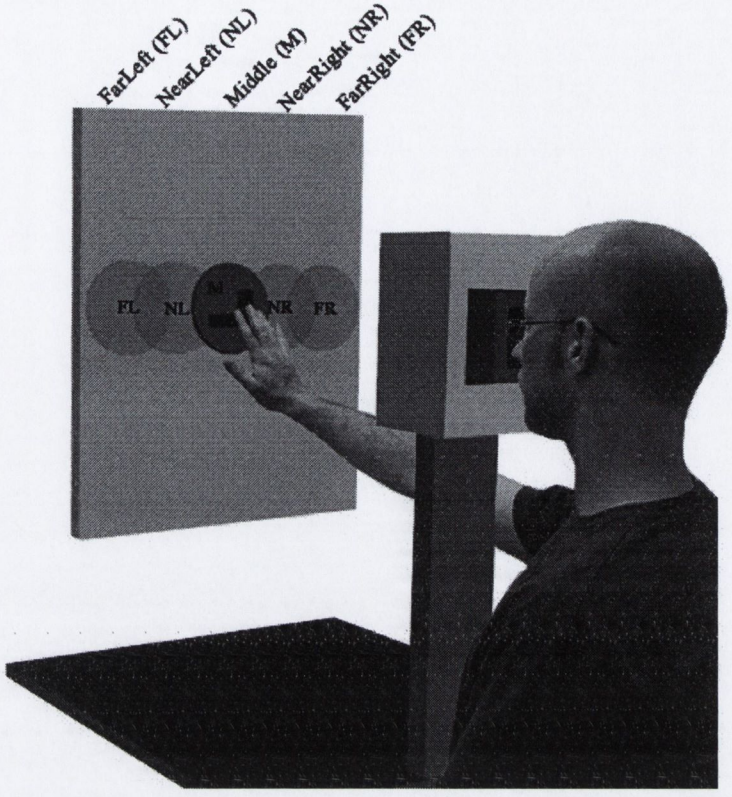


Figure 9. An illustration of the apparatus used in Experiment 5. In this Experiment, haptic and visual objects were either co-located (both visual and haptic objects were superimposed as indicated by the red circle), or were separated along the horizontal axis. Participants felt haptic objects with two hands.

3.3.1.3 Design

A two-way, within-subject design was used for this experiment with the haptic object's location in space (POSITION) as one factor, and the TRIAL TYPE (hits or correct rejections) as the second factor. There were five levels to the POSITION factor, each corresponding to the different horizontal positions the haptic object could occupy (far-left, near-left, middle, near-right, far-right). As in Experiment 4, for half of the trials the objects were matched, and for the other half they were different. Trials were presented randomly across participants.

The dependent variables were correct matching responses (HITS and CORRECT REJECTIONS) and response times (RT).

3.3.1.4 Procedure

The procedure was identical to that used in Experiment 4. For each of the 5 positions on the board there were 12 trials (six same and six mismatched). In total there were 60 trials. 10 practice trials were given at the start and the experiment took approximately 40 minutes to complete.

3.3.2 Results

The average percentage hit and correct rejection (CR) scores for the different positions of the haptic objects are listed in Table 5. A two-way, within-subjects ANOVA was conducted using POSITION as one factor (5 levels) and TRIAL TYPE (HIT or CR trials) as the second factor. There was no effect of POSITION [$F(4,92)=1.77$, $p=0.14$]. There was an effect of TRIAL TYPE [$F(4,23)=10.71$, $p<0.005$], with HIT scores being larger than CR scores. There was no interaction between factors [$F(4,92)=1.56$, $p=0.19$].

A second one-way ANOVA with POSITION using average RT for correctly matched objects across the different positions was not significant [$F(4, 92)<1$].

Table 5.
Mean and standard error scores (in brackets) for HITS, CR, and RT (average RTs for HITS and CR trials) for each level of POSITION in Experiment 5.

Position	% Hits (StdError)	% CR (StdError)	RT in seconds (StdError)
Middle	88.2 (±2.9)	71.5 (±2.9)	3.10 (±0.16)
Far-left	88.9 (±2.6)	77.8 (±3.0)	3.24 (±0.13)
Near-left	90.3 (±2.8)	74.3 (±4.4)	3.25 (±0.21)
Near-right	91.0 (±2.8)	79.2 (±3.2)	3.14 (±0.16)
Far-right	81.3 (±3.4)	77.1 (±3.3)	3.13 (±0.14)

3.3.3 Discussion

There was no difference found between matching performance to the cross-modal objects co-located in space compared to cross-modal objects mislocated in space. This finding suggests that objects presented across modalities did not integrate together, even though the criteria of spatial co-location, temporal co-incidence, and high compellingness were met. It was speculated that some top-down knowledge that the objects were different may have acted in Experiment 4 to prevent integration. As the assumption of unity in this experiment was made stronger by using visual stimuli that

were more appropriate to being integrated with the haptic stimuli, it was predicted that the top-down effect would not inhibit integration here. It seems however, that such a top-down effect did not influence the results and that there may be some further factor required for integration to occur.

It is possible the visual and haptic objects differed according to size and that this difference prevented integration. In experiment 6 the size difference between the visual and haptic objects was manipulated to see if there is an optimal size at which the objects are perceived as being the same. Visual and haptic encoding differences may have also prevented integration, and this is tested for in Experiment 9.

The difference between the results in Experiment 4 and here may have been due to the information available in the visual stimuli. A visual background 'frame' was not present in Experiment 4, where visual objects had a black background. Frame effects at the body-midline have been shown to modulate the perceived position of visual objects in space (Dassonville, Bridgeman, Kaur Bala, Thiem, & Sampanes, 2004), which possibly could account for the differences of effect between experiments.

3.4 Experiment 6

In the following experiment the idea was tested that visual and haptic objects may not have been integrated in Experiment 4 and 5 because they were perceived as different in size. Although the visual objects were matched in size to the haptic objects from previous subjective data collected in a pilot study, it is possible that during the actual experiments the visual and haptic objects may have been perceived as different in size. Even though it is very likely that small size differences would not be perceivable by the observer due to sensory capture effects (e.g., Heller, Calcaterra, Green, & Brown, 1999; Hershberger & Misceo, 1996; Rock & Victor, 1964), the combination of

possible size differences alongside small shape differences between the stimuli in different trials may have reduced the assumption of unity sufficiently to prevent integration from occurring (Welch & Warren, 1980).

Size differences would most likely have arisen in the pilot study where it was determined how large the visual objects had to be to match the haptic objects in size (see the Method section, Experiment 4). The small number of participants in the study may have not been sufficient to determine accurately the scaling function. Also, given that similar cross-modal size matching studies in the past have found quite varying levels of correspondence between visual and haptic size estimations (Hohmuth, Phillips, & VanRomer, 1976; Kelvin, 1954; Millar, 1972; Schultz & Petersik, 1994; Seizova-Cajic, 1998; Teghtsoonian & Teghtsoonian, 1970), it is possible that the scaling function determined in the pilot study may have been inappropriate for the purposes of the matching studies

To determine whether the visual objects differed in perceived size from the haptic objects, the objective scaling required to exactly match visual and haptic objects in size was calculated and this was compared to the subjectively sized visual objects. It was found that the visual objects were indeed 32.6% smaller than the haptic objects. It is interesting to note that Wuillemin, van Doorn, Richardson & Symmons (2005) similarly found that virtual visual objects subjectively matched in size to haptic objects were actually smaller than their haptic counterparts. Such subjective size discrepancies may simply be a by-product of mixed reality set-ups.

In order to ensure that the size of matched objects in the previous experiments was perceived as the same, the performance across differently sized cross-modal objects was compared. Participants were instructed to judge whether differently sized but same shaped objects were the same size or not.

3.4.1 Method

3.4.1.1 Participants

14 students (3 male and 11 female) from Trinity College Dublin volunteered to participate in the experiment. Their mean age was 24.3 years. One of the participants reported to be left-hand dominant. All of the participants had normal or corrected-to-normal vision and none reported any tactile impairment.

3.4.1.2 Stimuli and Apparatus

The 12 visual stimuli used in Experiment 6 were copied 4 times, adjusted in size, and assigned accordingly into four sets of object sizes relative to the size of the haptic object. These sizes were SMALL, SUBJECTIVE 'SAME', OBJECTIVE 'SAME' and LARGE (SUBJECTIVE 'SAME' sized objects were the same size as the objects in previous experiments). These differently sized visual objects compared in size to their haptic counterparts in the following way: SMALL was 67% objectively smaller than the haptic objects; SUBJECTIVE 'SAME' was 32 % smaller; OBJECTIVE 'SAME' was the same objective size; and LARGE was 18% larger (the maximum size visual objects could be due to the screen size of the HMD). The approximate visual angles subtended by the horizontal of the largest object in each set were as follows: SMALL, 5 degrees; SUBJECTIVE, 9 degrees; OBJECTIVE, 13 degrees and LARGE, 16 degrees. The 12 haptic stimuli were identical to those used in Experiment 5. The haptic and visual pairs of objects were all identical in shape.

The apparatus was identical as that used in Experiment 4, except that the haptic objects always occupied the central location so that visual and haptic objects were co-located in space. Responses were made by depressing the right foot pedal for "same" and left for "different".

3.4.1.3 Design

A one-way, within-subject design was used with relative SIZE as the factor (SMALL, SUBJECTIVE, OBJECTIVE and LARGE). The dependent variables were the number of times the participant judged the pair of objects to be the same size and response times (RT). Trials were randomly presented across participants.

3.4.1.4 Procedure

The procedure was identical to that used in Experiment 4, except here the participant's task was to identify whether the simultaneously presented objects were identical *in size* or not. Thus in this experiment, all object pairs were *identical* in shape, but could differ in size. As in previous experiments, haptic objects were explored using both hands.

Participants were presented with 12 object pairs in each size condition. Thus there were 48 trials in total in this experiment. Five practice trials were given at the beginning of the experiment. In total the experiment took approximately 30 minutes to complete.

3.4.2 Results

Data from two participants had to be removed from the analysis as it became clear that they had misinterpreted the task (both matched objects according to shape and not to size). The mean percentage 'same' matching responses were calculated across conditions and presented in Figure 10. A within-subject ANOVA was conducted with the number of responses as the dependent variable and relative SIZE as the independent variable (SMALL, SUBJECTIVE, OBJECTIVE, and LARGE). A main effect of relative size was found [$F(3,33)=2.98$, $p<0.05$]. A posthoc Newman-Keuls analysis found that there were significantly more 'same' judgements made to the

SUBJECTIVELY same sized object pairs than when the visual stimulus was LARGE ($p<0.05$). Performance to the other cross-modal size differences did not differ from each other (see Figure 10).

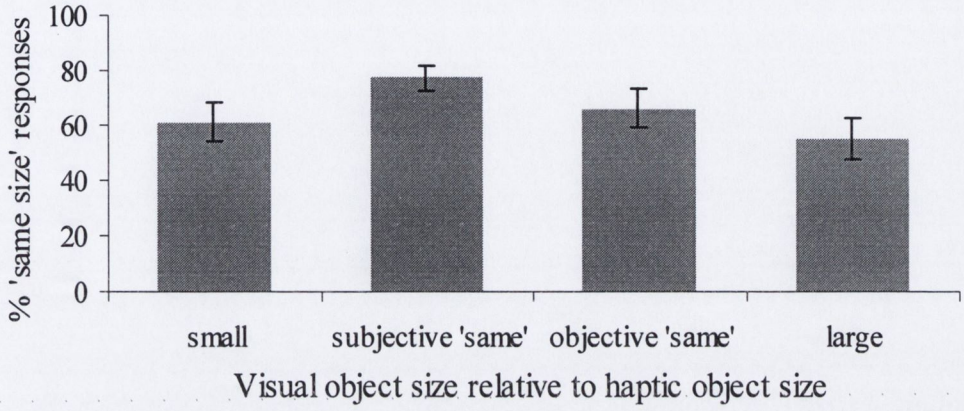


Figure 10. Plot depicting the extent to which objects in each size category matched the size of the haptic objects to the size of the visual objects as a function of the relative size of the visual stimuli (a score of 100% would indicate that the object pairs were perceived as perfectly matched in size, whilst a score of 0% would indicate that the object pairs were different in size).

A one-way ANOVA was performed on the RT data with relative SIZE as the factor. There was no effect of size [$F(3,33)=1.56$, $p=0.22$]. The mean RTs for each level of SIZE was a follows: 3.00s (S.E. 0.24s) for the relatively SMALL visual objects, 3.2s (S.E. 0.29) for the SUBJECTIVE same set, 3.3s (S.E. 0.30s) for the OBJECTIVE same visual size and 3.3s (S.E. 0.30s) for the objectively LARGER visual stimuli.

3.4.3 Discussion

As expected, the results suggest that the visual and haptic objects used in previous experiments were perceived as identical in size as the objectively matched object pairs for size. It is therefore quite unlikely that integration was prevented from occurring in the previous experiments because of size effects. Here there was no difference between

matching performance for visual stimuli that were relatively SMALL (61.1%), SUBJECTIVELY same (77.1%), and OBJECTIVELY the same size (66.0%) as the haptic objects (LARGE objects matching performance was at 54.9%). It seems that the results of the pilot study previously discussed correctly determined the subjective size of visual objects so that they were matched in size to the haptic objects.

It was surprising however, that there was a high frequency of 'same-size' response across all trials in the experiment. The number of 'same' responses across the size categories never dropped below 50% and on average people rated over 64% of the cross-modal object pairs as being equivalent in size. To illustrate the range of visual object sizes that were rated as identical to their haptic counterparts, visual objects objectively the same size as haptic objects subtended an area on the retina that was approximately 2.7 times as great as the area subtended by visual objects which were relatively small. It is possible, as discussed in the introduction, that the sizes of the haptic objects were effectively captured by vision thus reducing the probability that actual size differences would be noticed (e.g., Heller et al., 1999; Hershberger & Misceo, 1996; Rock & Victor, 1964). Nevertheless, the sheer magnitude of this capture was unexpected. It thus seems that there is considerable flexibility in the relative size of visual and haptic objects that may be integrated (Klatzky, Lederman et al., 1993).

In conclusion, it seems that the objects used in the previous experiments were not perceived as different in size, making it unlikely that size differences between visual and haptic objects prevented integration from occurring.

3.5 Experiment 7

Integration may rely on objects being optimally encoded across the senses. In Experiments 4 & 5, visual and haptic objects were positioned in such a way that they faced the observer. However, as recently demonstrated by Newell, Ernst, Tjan, &

Bülthoff (2001), haptics is viewpoint dependent and the surface of an object that is best encoded through touch is the surface that faces away from the observer. Thus, haptic encoding in Experiments 4 & 5 may not have been optimal.

As discussed in the introduction, Gibson (1962) argued that it is only through active exploration that both modalities may "yield the same information" about an object. The nature of the exploration in Experiments 4 & 5 may have accentuated the differences in what was encoded by vision and touch thus resulting in better discrimination of all objects at all positions. Perhaps a caveat needed for Gibson's statement is that exploration must be both active and natural for integration to occur.

Newell, Ernst, Tjan, & Bülthoff (2001) suggest that the encoding of common surfaces by the modalities leads to efficient cross-modal recognition. However, they argue that surfaces of objects must be optimally encoded for matching to occur. In this experiment haptic objects were rotated 180 degrees leading to the encoding of one surface by vision, and the encoding of the back of the surface by touch. If integration depends on optimal encoding of object surfaces across modalities then it was predicted that matching performance for co-located objects would be compromised for the different object pairs.

3.5.1 Method

3.5.1.1 Participants

14 undergraduate students (4 male and 10 female) from Trinity College Dublin participated in the experiment for course credits or payment (8 euro per hour). Their ages ranged from 18 to 29 years, and the mean age was 20.2 years. All of the participants had normal or corrected-to-normal vision, none reported any tactile impairment. 11 of the participants reported to be right-hand dominant.

3.5.1.2 Stimuli and Apparatus

The stimuli were identical to those used in Experiment 5. The apparatus was the same, except that a new supporting structure was used to hold the disk and objects in a position facing away from the participant (see Figure 11). An adjustable clamp held the disk in place on the structure, allowing the height of the disk from the table to be adjusted. The lateral distance between the disc and the participant could be adjusted by sliding the supporting structure's base on the table either to the left or to the right of the participant. The table was marked with 5 positions along the horizontal axis to which the structure could be positioned. This permitted the haptic object to be precisely located along the same 5 locations in space as in Experiment 5. To ensure the apparatus was always perpendicular to the participant, guiding rails were used to constrain the movement of the haptic object between one trial and the next. As in previous experiments, visual objects were seen through an HMD and were virtually projected so that the visual object and haptic object were superimposed when the haptic object was located at the central position.

3.5.1.3 Design and Procedure

A one-way, within-subjects design was used with POSITION as the main factor. There were 5 levels corresponding to the 5 positions that the haptic object could occupy relative to the location of the visual image. The dependent variables were participant response rate (HITS AND CORRECT REJECTIONS) and response time (RT). Trials were randomly presented across participants.

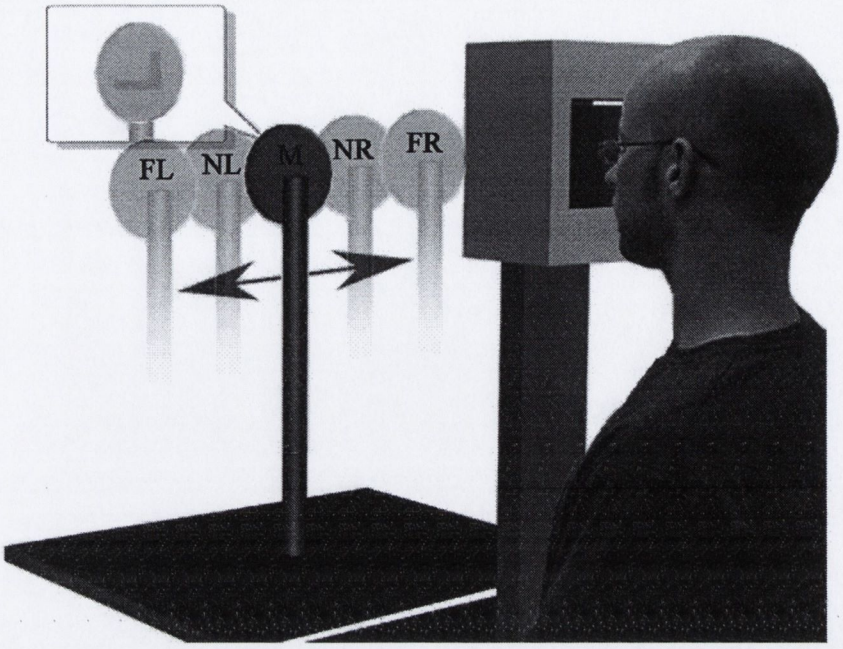


Figure 11. An illustration of the apparatus used in Experiment 7. Haptic objects were placed on the back surface of the disk shown, facing away from the participant. The participant had to reach around the disk with both hands to feel the haptic object. The haptic object was placed at one of five different positions relative to the location of the visual object (FR = far-right, NR = near-right, M = middle, NL = near-left, FL = far-left).

As in previous experiments the experiment was based on a same/different matching paradigm where the participant's task was to decide whether two stimuli presented simultaneously across modalities were the same or different as fast and as accurately as possible. The procedure was similar to Experiment 5, but there were a few exceptions. Firstly, to ensure that the visual and haptic objects were co-located at the central haptic object position, participants were asked to first feel a haptic stimulus and instruct the experimenter to move it accordingly until the participant perceived that it was placed in the position where the visual object was projected. A reference marker on the supporting structure was then positioned in such a way to indicate the subjective point of co-location for each participant before the experiment commenced. The experimenter moved the position of the haptic object between trials by sliding the supporting structure across the table so that the reference point and the appropriate

position marker were aligned. Throughout the experiment, participants explored the haptic object with two hands.

3.5.2 Results

As in previous experiments, the mean of each participant's hits and correct rejection scores and response times were calculated (see Table 6 for descriptive statistics). A two-factor within-subject ANOVA on the POSITION and TRIAL TYPE factors was performed on the scores. There was no effect of POSITION [$F(4,52)=1.63$, $p=0.18$] or TRIAL TYPE [$F(1,13)<1$], and there was no interaction between the factors [$F(4,52)=1.15$, $p=0.35$]

A one-way, within-subjects ANOVA on the RT data across the positions also revealed no significant effects [$F(4,52)=1.42$, $p=0.24$].

3.5.3 Discussion

The results of this experiment suggest that the integration of cross-modal objects when both visual and haptic objects were optimally encoded did not occur at the point of co-location in space. Also, the results neither verify nor deny whether haptic objects which were facing away from the observer are indeed perceived as being more similar to the visual objects (Newell et al., 2001) than in previous experiments. The response rates and response times found in this experiment are equivalent to those found in previous experiments.

Table 6.

Mean and standard error scores (in brackets) for HITS, CR and RT (average RTs for HITS and CR trials) for each level of POSITION in Experiment 7.

Position	% Hits (StdError)	% CR (StdError)	RT in seconds (StdError)
Middle	85.7 (±5.8)	85.7 (±3.4)	3.46 (±0.25)
Far-left	85.7 (±4.9)	77.4 (±4.8)	3.37 (±0.24)
Near-left	86.9 (±4.0)	81.0 (±3.9)	3.44 (±0.25)
Near-right	89.3 (±4.1)	88.1 (±2.7)	3.67 (±0.22)
Far-right	79.8 (±5.3)	83.3 (±3.9)	3.41 (±0.25)

3.6 Experiment 8

In Experiment 4 matching performance to cross-modal objects was found to vary over horizontal separations between the objects. Trials where haptic objects placed at the near-right position were (non-significantly) more difficult to differentiate from visual objects than haptic objects which were placed in the other horizontal positions. This result was not expected since it was predicted that cross-modal objects co-located on the middle position would be more difficult to discriminate. However, it may have been that the actual point of co-location in Experiment 4 differed from the point of co-location that was subjectively determined at the start of the experiment. Here it was tested whether there was a haptic rightward bias in where the subjective co-location of visual and haptic objects lay.

But why would the subjective point of co-location differ from the actual point of co-location during the experiment? One suggestion is that when the arm moves that this proprioceptive information increases uncertainty in terms of location in haptic space. However, it is unlikely that proprioceptive imprecision would bias the location of the point of co-location to the right since, as van Beers, Sittig, & Denier van der Gon, (1998) recently found proprioception is comparable in terms of location accuracy to vision (when proprioception uses radial rather than tangential movements of the shoulder). In any case, increased error would not explain a consistently rightward error of location as the previous data suggest.

Other studies have suggested that the point of co-location actually drifts in a leftward direction, during an experiment (Brown, Rosenbaum, & Sainburg, 2003). The direction of drift has also been found to be quite subjective (Desmurget, Vindras, Grea, Viviani, & Grafton, 2000; Tillery, Flanders, & Soechting, 1994; Vindras, Desmurget,

Prablanc, & Viviani, 1998). However, the above studies were all performed with one hand only. In Experiments 4 & 5 reported here, participants used two hands to explore stimuli. If drift is related to single hand exploration then it seems unlikely that the point of co-location would systematically drift to the right for all of the participants.

Distortions in the representation of space have previously been demonstrated, such as effects of pseudoneglect. Although pseudoneglect effects generally lead to the midline of a bisected space or object being moved towards the left (Bowers & Heilman, 1980; Sampaio & Chokron, 1992), Sampaio & Philip (1991), on the other hand, found a rightward pseudoneglect effect for a bar bisection task which relied on proprioceptive information only, compared to a leftward pseudoneglect effect for a task relying on kinaesthetic and proprioceptive information. Handedness (Sampaio & Chokron, 1992) and scanning direction (Baek et al., 2002) also act to modulate the size of the pseudoneglect effect, and as the majority of the participants in the experiments were right handed, this may have had an effect. The possibility remains that the point of co-location between the cross-modal objects may not have been at the midline but displaced to either the left or the right. The following experiment was designed to investigate if distortion of the point of co-location occurred in the previous experiments.

In the following experiment, participants were required to position a haptic object until its visual counterpart superimposed the location of the haptic object. The deviation between the subjective position of the object and the objective centre was calculated.

3.6.1 Method

3.6.1.2 Participants

10 undergraduate students from Trinity College Dublin participated in the experiment for course credits or payment (8 euro per hour). Eight of the participants were female. One of the participants reported to be left-hand dominant. The ages of the participants ranged from 18 to 41 years, and the mean age was 22.9 years. None of the participants reported any tactile impairment and all reported normal or corrected-to-normal vision.

3.6.1.3 Stimuli and Apparatus

The stimuli were the same as those used in Experiment 4. The apparatus was the same as that used in Experiment 4 with a few exceptions. The layout of the apparatus is depicted in Figure 12. A co-ordinate grid was drawn in millimetres onto the board where the haptic stimuli were presented. A triangular piece of card was attached to the bottom of the disc that held the haptic objects. This triangle was right-angled and positioned such that its right-angled corner was directly placed beneath the right-angled intersection of the haptic stimuli. It was used to indicate precisely where the haptic object was positioned by the participants. To move the haptic stimuli, participants slid the disk which held the haptic stimuli across the surface of the board. As in previous experiments, visual objects were always displayed through the HMD and were projected in the centre of the virtual board.

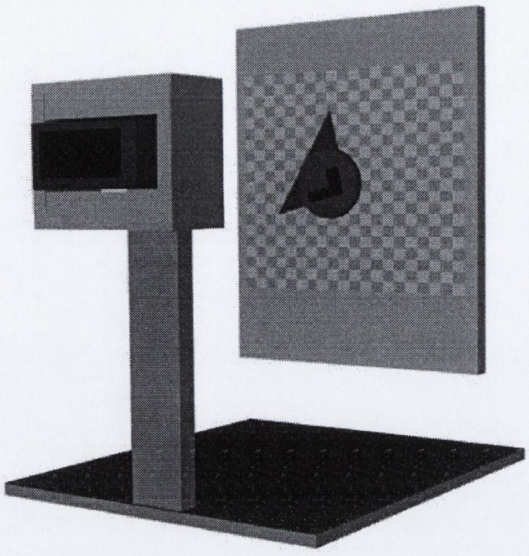


Figure 12. An illustration of the modified apparatus, including the co-ordinate grid and the triangular card which was attached to the back of the disc holding the haptic object.

3.6.1.4 Design

The experiment was based on a one-way within-subject design with STARTING POSITION as the factor. The dependent variable was haptic object horizontal mislocation (ERROR) measured in centimetres. Before the participant moved the object the haptic objects could either be located to the left or to the right for each trial, thus left and right were the two levels of STARTING POSITION.

The mislocation of the haptic object was calculated by measuring the distance between the point of intersection of the objects x- and y-axis arms (calculated by measuring the x and y co-ordinates of the two visible points of the triangle on the haptic stimulus) from the objective vertical axis of co-location determined at the start of the experiment.

All trials were randomly presented across the participants and the experiment was run using DMDX software for the PC (Forster & Forster, 2003).

3.6.1.5 Procedure

The participant's task was to reposition the haptic object that either was initially located to the far right or far left of the magnetic board so that it occupied the exact location as its visual counterpart viewed through the HMD and projected onto the virtual board. Participants used two hands to do this.

The basic procedure for each trial, the practices, and the initial calibration was identical. It took approximately 10 seconds for the experimenter to place a haptic object on the board either to the left or right of the board. These starting positions were 21cm away from the objective centre in either the left or right direction. Once the experimenter placed the object, the trial began. A fixation cross was first displayed for 500ms followed by the visual stimulus, which remained on the screen until the end of the trial. As soon as the visual object appeared, the participant located the haptic stimulus and then slid it across the board to occupy the exact same location as the visual object. Participants were required to perform this task as fast and as accurately as possible. This position of the haptic object was noted by the experimenter by measuring the distance of two of the points of the card triangle on the co-ordinate grid relative to the objective centre of the board. At the start of the experiment, participants were given two practice trials.

Each of the 12 haptic and visual object pairs was presented twice in the experiment, one at each of the starting positions. Trials were presented in a random order. In total there were 24 trials, and the experiment took around 30 minutes to complete.

3.6.2 Results

Mean horizontal subjective spatial-location errors were calculated by subtracting the horizontal location at which participants placed objects during the experiment, from

the objective vertical axis of visual-haptic object co-location. The direction of haptic object movement determined the sign of the horizontal subjective spatial-location error. A positive score would indicate that participants had overshot the actual position of co-location when repositioning haptic objects whilst a negative score would indicate that participants had undershot the actual position of co-location relative to the starting point. Thus, objects initially located to the left which had a positive subjective spatial-location error would be located to the right of the actual point of co-location, and objects initially located to the right which had a negative subjective spatial-location error would also be located to the right of the actual point of co-location. The mean horizontal location error for objects that were originally located to the left was +0.49cm (0.53cm S.E.) and for objects originally located to the right, the mean was -2.74cm (0.66cm S.E.). Thus, for both starting positions, objects were placed to the right of the point of co-location.

A dependent-samples 2-tailed t test was conducted, contrasting horizontal spatial-location errors for objects initially located to the left to objects initially located on the right. Horizontal location estimation error was affected by starting position [$t(9)=4.47$, $p<0.005$]. Sign-tests were used to determine whether horizontal spatial-location error was significantly greater than objective co-location (i.e., 0cm error). This analysis revealed that horizontal spatial-location error was not greater than that expected by chance for objects which were initially placed to the left but horizontal spatial-location error was greater than that expected by chance for objects which were initially placed to the right ($p<0.005$).

3.6.3 Discussion

The results suggest that there was a rightward bias in the perception of co-location between the visual and haptic objects in this experiment. Irrespective of the objects'

starting positions, objects were positioned slightly to the right of the vertical axis of co-location that was determined at the start of the experiment.

This co-location bias towards the right may reflect an error that is introduced by proprioception, pseudoneglect, or even handedness. Importantly, this position bias has consequences for the effects found in Experiment 4. It seems possible that the reduced performance found to the near right position in Experiment 4 may be due to co-located cross-modal object positions, as was predicted. Unfortunately, this effect was not replicated in the follow-on experiments (e.g., Experiment 5) therefore it is difficult to determine how robust it is.

3.7 General Discussion

This chapter tested to see whether objects that were spatially co-located across modalities would be integrated using a same/different matching paradigm. It was hypothesised that if co-location was a requisite for cross-modal integration, participants should be less able to discriminate nonmatching object shapes which were co-located as these objects should be integrated, compared to objects that were spatially separated. Experiment 4 however did not find any evidence for a reduced sensitivity when objects were objectively co-located, suggesting that cross-modal integration does not rely on objects across the senses being spatially co-located.

One reason which may have prevented integration from occurring in Experiment 4 was that the object stimuli were rendered quite different from each other. The visual objects were 2-dimensional graphical depictions of the objects without any depth information. In experiment 5 photographs of the haptic stimuli were used as visual objects, which was suggested would strengthen the assumption of unity of the stimulus-display (Welch & Warren, 1980). Here, however, there was no evidence that visual and haptic objects were being integrated since performance did not depend on the position

of the haptic object. It is not clear why there was an effect of haptic object position with graphical images in Experiment 4 but not with photographs in Experiment 5. The number of positions was reduced and the trial number increased in Experiment 5 which may have constituted a practice effect that reduced any effects of co-location.

Potentially the visual and haptic objects may have differed in perceptual size even when they were the same shape. In Experiment 5 it was identified that the size of the visual objects were in fact objectively smaller in size than their haptic counterparts. It was vital to verify that the data from a pilot study (which was ran to determine the subjective equivalent size of the visual objects to the haptic objects) had accurately determined visual object size. It follows that if visual and haptic objects were perceived as differently sized, the objects' percepts may not have been integrated across modalities because they were not sufficiently similar. However, the results from Experiment 6 suggest that the pilot had been accurate in determining visual object size, and it was concluded that the visual and haptic objects were indeed seen subjectively as the same size.

Potentially, integration may not have occurred as vision and haptics differed in terms of encoding (Newell et al., 2001). In Experiment 7, haptic objects were rotated so that both vision and touch were able to encode object information optimally, but did not find any effect of integration for co-located objects.

Interestingly, previous literature and the results from Experiment 4 suggest that there may have been a rightward bias of the location of where visual and haptic objects were co-located. If haptic objects located to the right of visual objects were perceived as being co-located, then this may explain the reduced matching performance to object pairs since the cross-modal objects may have been potentially integrated at these rightward locations. The results from Experiment 8 suggested that this may indeed

have been the case. For example, the location of the haptic objects that was judged to be co-located with the visual objects was, in fact, slightly to the right of centre. The fact that in Experiment 4 reduced matching performance was found for rightward-located haptic objects (not replicated in Experiments 5 and 7 however) lends support to the idea that subjectively co-located objects are integrated and therefore more difficult to discriminate. In Chapter 4, it was tested whether or not location actually has an effect on haptic object shape.

In general, the findings from this chapter suggest that cross-modal integration does not consistently rely on the perceived spatial co-location of the visual and haptic objects. In Chapter 2, it was discussed how a similar finding, that temporally disparate objects do not integrate, suggested that visual and haptic percepts may thus be maintained independently, albeit in memory systems with very similar information processing properties to each other. The findings from Chapter 3 also generally support the conclusion that the modalities can maintain discrete object representations. However, it still cannot be ruled out the possibility that the modalities do share object representations and that the experiments were not sensitive enough. The findings, as they stand, are contrary to a wide range of previous research (e.g., Ernst & Bülthoff, 2004; Làdavas et al., 2000; Lalanne & Lorenceau, 2004; Soto-Faraco et al., 2002; Spence et al., 2001; Stein & Meredith, 1993; Wallace et al., 2004) and so it is quite likely that there are other reasons, besides from co-location being not useful in cross-modal integration, which explain inconsistent effects reported in this chapter. The following sections include some suggestions for what these reasons might be.

3.7.1 Reasons for the lack of integration effect

Firstly, despite the efforts across these experiments the objects may still have been perceived sufficiently different from each other such that they did not satisfy the

unity assumption (Welch & Warren, 1980); even when visual objects were photographs of haptic objects (Experiment 5); and when encoding was rendered as equivalent across vision and touch (Experiment 7). For example, even though the visual objects in Experiment 5 were photographs of their haptic counterparts (and would thus seem a lot more similar in appearance to the visual objects), the visual objects were still presented as 2D images since they were not in stereo. Visual encoding may thus not have been optimal as objects were not perceived in stereovision (e.g., stereovision-haptic curvature matching along the Z dimension has a lower Weber fraction than that of the same task where vision is not stereo, illustrating that stereovision is better able to encode object information, Jansson & Öström, 2004), making object percepts seem different from each other (c.f. Experiment 7). Also, recently, the importance of dynamic motion in integration has been highlighted by Soto-Faraco & Kingstone (2004). Small movements of the haptic objects may have prevented the visual and haptic objects from being integrated. The haptic objects, although rigidly held in place by the experimental apparatus, could nevertheless still be moved by very small amounts during exploration (e.g., there was a certain unavoidable 'give' in the supporting apparatus that permitted very slight movements of the haptic objects). Similarly, the visual objects could also have appeared to move very slightly: although the HMD was anchored in place, if participants shifted in their seats or attempted to move their heads there would be an associated, yet very small shift of the projected location of the virtual visual object on the stimulus-display with respect to the observers head or eye position. A compensative measure which could be taken to avoid this issue in future research would be to anchor the virtual object directly to the stimulus array. Any movements elicited during haptic exploration would then be reflected in the visual object. Although it is likely that matched visual movements would not be directly proportional in

magnitude to haptic movements, movements would be matched in terms of onset, offset, direction etc., making the cross-modal experience more realistic.

3.7.2 Reasons why integration effects were not detected

Second, as the large number of errors in each experiment suggest, objects may have been integrated in the studies, however, the effects of co-location may not have been detected due to lack of sensitivity in the design. Although the power in Experiment 5 was increased relative to Experiment 4, it may not have been raised sufficiently to find an effect of co-location.

High error rates in the matching performance may have hidden an effect of integration at the point of co-location. As a similar error rate was observed over all the positions in Experiments 4, 5 and 7, one possibility is that a ventriloquist-like effect occurred such that, irrespective of haptic object position, haptic objects were always integrated with visual objects (in the traditional auditory-visual ventriloquist effect, auditory and visual information derived from different spatial locations are typically perceived as both orientating from the same visual location, as occurs with a ventriloquist and his dummy, Howard, 1966). Under within-modal haptic object recognition haptics may encode objects relative to a hand centred frame of reference (Kappers, 2004). According to such a frame, the shape of the object would be the same, irrespective of location. However, when visual objects are presented this may mean a change in reference frame for haptics, possibly resulting in a ventriloquist effect. This possibility will be considered in greater detail in the final chapter.

The lack of a consistent finding of co-location effects may have occurred as a result of the participants not seeing their hands during haptic exploration. By preventing vision from determining the exact location of the hand in space, this may have influenced the perceptual system making it rely on proprioceptive cues to deduce

hand position. It is known however that vision can dominate in the perceived location of the hand over proprioception (Hay, Smith, Pick, & Ikeda, 1965). It may be that all haptic objects at all positions were subjectively perceived as co-located with the projected visual object, hence, integration may have occurred irrespective of haptic object position. Again, this will be discussed further in the final chapter.

CHAPTER 4: LOCATION-SPECIFIC HAPTIC-OBJECT DISTORTION

Abstract

This chapter was concerned with location-specific haptic object distortion. In the preceding chapter, haptic objects located to the right of the body-midline in Experiments 4, 5 and 7 were found to be harder to match to visual objects, compared to haptic objects in other locations. It may be that objects in the rightward location were not distorted, relative to objects in other locations. In Experiments 9 & 10 the lateral location of a haptic object in respect to body were manipulated and participants had to reproduce perceived haptic object shape visually. The distance at which objects were separated from the body was also manipulated to determine the optimal placement that haptic objects should be from the body such that they were least distorted. Experiment 9 identified that haptic objects were subject to distortion, but the effects were not modulated by position. In experiment 10 haptic objects were placed in a square matrix in front of the participant and evidence for a location-specific haptic object distortion were found. These findings suggest that haptic objects in Experiments 4, 5 and 7 located to the right were harder to match to visual objects when located to the right as haptic objects at other locations were subject to location-specific distortion and were thus more easy to discriminate.

4.1 Introduction

In chapter 3 the ability to match visual and haptic objects according to shape was found to depend on the haptic object's horizontal location in space (Experiment 4, and non-significantly in Experiments 5 and 7), and this effect seemed only partly due to a bias of position of the point of visual haptic co-location (Experiment 8). Here, it is identified whether or not this differential performance was also due to a location-specific distortion of the representation of the haptic object. It follows that if the haptic objects were distorted depending on their position in space, the similarity of visual and haptic percepts would also differ depending upon the haptic object's position in space.

As the only constraint imposed on haptic exploration in previous chapters was that objects were to be explored with two hands, it is possible that the haptic objects were affected by a complex interplay of different distortions. For example, the majority of previous research concerned with object distortion limited exploration to just one hand, imposed exploratory style limitations, and often used simple stimuli (e.g., Davidson & Cheng, 1964; Millar & al-Attar, 2000; Reid, 1954). Very few authors have tested for haptic object distortion when exploration is unconstrained, or when haptic objects were truly 3-dimensional (however, see Heller, Bracket, Salik, Scroggs, & Green, 2003). It is beyond the scope of the present study to attempt to isolate the factors affecting the distortion of an object's representation (that was speculated to affect the results in the previous chapter). So, while it is the aim of this chapter to identify if object representations are being differentially distorted dependent up their position in space, the mechanisms which would cause such distortions will only be speculated on. Here, it is also tested to see whether location-specific distortion of object representations is modulated by the distance between the object and the body. If object representations

are affected by distance from the body, this should be taken into consideration in future research.

There are a number of reasons why haptic object shape may be distorted in memory. For example, L-shaped objects, such as those used in Chapters 2 and 3, may be prone to the horizontal-vertical illusion (Davidson & Cheng, 1964; Heller et al., 2003; Reid, 1954) so it seems pertinent to discuss the factors which cause the illusion, and whether or not the illusory effect may be modulated by an object's location in space. When an L-shaped object is positioned in the horizontal plane, its y-axis arm aligned with the sagital midline of the observer, its x-axis arm aligned fronto-parallel to the observer (i.e., the object lies flat on the tabletop so that it is square to the observer), the object's y-axis arm is perceived as being overestimated in length relative to its x-axis arm (Davidson & Cheng, 1964; Day & Wong, 1971; Reid, 1954). This has been termed the radial-tangential illusion after the radial and tangential arm movements used to explore each arm of the object. Radial movements are typically used to explore an object's vertical arm and tangential arm movements are used to explore an objects horizontal arm (e.g., Day & Wong, 1971). It has also been called the haptic horizontal-vertical illusion, after its counterpart in vision (Coren & Girgus, 1978). In Chapters 2 and 3 however, objects were arranged in the fronto-parallel plane (both arms are explored with tangential arm movements, Day & Wong, 1971). Heller et al. (2003) interestingly found that the illusion reverses when a 3D L-shaped object is orientated in this plane, with the object's x-axis arm now perceived longer than its y-axis arm (see also Day & Avery, 1970). Exploratory strategies such as grasping using a pincher grip and free exploration were speculated to enhance the illusion, as illusory effects were not found for all exploratory strategies (i.e., finger tracing). The illusion was also speculated to depend on the awkward placement of the arm during exploration. For

example, in Heller et al.'s experiments, objects were always placed on the body-midline, and a reviewer suggested that at the midline an awkward arm posture must be adopted to explore objects, possibly giving rise to the illusion. Presumably objects placed at other locations are not affected by the illusion as the arm posture is not so awkward, although this was not directly tested. Thus, exploratory strategy and awkward arm posture do suggest that distortion of an object's representation may be location-specific.

Whole-arm movements may also be involved in location-specific distortion of object representations. Heller et al. (1997) were interested in how shoulder and arm movements contributed to the positive horizontal-vertical illusion for 2D stimuli in the horizontal plane. Unconstrained fingertip exploration led to an overestimation of a 2D object's y-axis arm relative to its x-axis arm (see also Martinez, 1971; Paillard, 1991), but when whole-arm movements were prevented by anchoring the elbow on the table or using small stimuli which only allowed exploration by finger movements, the illusion was much reduced. Crucially however, Heller et al. also found that the object's x-axis arm was also overestimated, albeit to a smaller extent. It seems likely then that whole-arm movements will also act to overestimate the perception of the x-axis arm of an object orientated in the fronto-parallel plane (the x-axis of an object remains in the same orientation irrespective of whether the object is aligned in the horizontal or fronto-parallel plane). Millar & al-Attar (2000) directly tested this whole-arm movement hypothesis for 2D stimuli arranged in the horizontal plane and predicted that if the hand had to cross the body-midline to reach an object, more whole-arm and shoulder movements would be required to explore the x-axis arm of an object, compared to when the object was on the same side of the body as the exploring hand. An object located across the body-midline should therefore be more affected by the predicted effects of

whole-arm movements than objects located elsewhere. Millar & al-Attar did indeed find that lateral location affected how object shape was perceived. Objects explored with the left hand were affected by the horizontal-vertical illusion when they were located to the left, but not to the right or when objects were located on the body-midline. For right-handed exploration on the other hand, illusory effects were found for objects located to both the left and the right, but not to the midline. The authors concluded that whole-arm movements did seem involved in the illusion, but the illusion also seemed dependent upon discrepancies between the reference frame used to encode object information during movement and body-centred reference frames. Although it is unclear whether these findings for 2D stimuli will generalise to 3D stimuli, shoulder movements are still involved in the exploration of 3D objects. It seems possible then that shoulder movements may also be a factor in location-specific distortion of object representations.

In the previous chapters, haptic exploration was always bi-manual. Hemisphere laterality effects (Hellige, 1993), such as the left hand being more accurate than the right in tactile form recognition (Fagot, Hopkins, & Vauclair, 1993; Flanery & Balling, 1979; c.f. Heller, Joyner, & Dan-Fodio, 1993) and handedness effects (speculated at by Bradshaw, Nettleton, Nathan, & Wilson, 1983; in the visual literature, see e.g., Luh, 1995; Marzi, Grabowska, Tressoldi, & Bisiacchi, 1988) would be present when objects are explored with just one hand. Such effects may however still play a role in bi-manual exploration as the hands, when working together, do not necessarily encode identical object information. For example, Fagot, Lacreuse & Vauclair (1994) demonstrated that the right hand prefers to explore the right hand side of an object, and the left hand, the left side of an object. It is possible then, that the left side of an object (the x-axis arm of the stimuli used previously) will be associated with the left hand, and

the right side (or the y-axis arm of the stimuli) associated with the right hand and consequent laterality effects may be involved. It also seems likely that at more extreme object positions from the body-midline, where it is hard to reach the object with a particular hand, one hand would likely dominate in the encoding of haptic object information. In this situation, hemisphere laterality effects may play more of a role in distortions affecting the representations of haptic objects depending upon an object's position in space.

The distance of an object from the body may also affect how that object is perceived. Objects explored at a distance from the body may appear smaller than the same object felt close to the body (Bartley, 1953; Cheng, 1968; Liddle & Foss, 1963). However, this effect with distance is not consistently reported (Armstrong & Marks, 1999), and even the opposite effect has been observed, such that objects felt at a distance from the observer are perceived as being larger than objects felt close to the body (Marchetti & Lederman, 1983).

A range of mechanisms such as exploratory strategy, whole-arm movements and hemisphere laterality effects may act to distort an object's representation in memory depending upon the position of the object in space relative to the body-midline. It also seems likely that these distortions may be modulated by the distance of the object from the body. In the following experiment an investigation was conducted to see whether haptic object representations are indeed distorted dependent upon their position in space, relative to the body-midline. Participants were required to recreate the shape of a haptic object as a visual stimulus. Deviations from the actual shape were calculated as a measure of perceptual distortion.

4.2 Experiment 9

In this experiment it was tested to see whether the representation of a haptic object could be distorted dependent upon the object's location in space. In Chapter 3 objects were most likely to be mismatched when haptic objects were located to the right of the body-midline. If haptic object distortion acted to make haptic percepts more different in shape to their visual counterparts (i.e., in the "different" trials), then matching performance should have been better to those locations where haptic objects were distorted. Conversely, haptic objects that were not perceived as distorted might be more likely to be integrated with their visual counterparts thus reducing matching performance. This would predict that haptic objects located to the right of the body-midline should not be perceived as distorted relative to objects in other locations.

To test for this, haptic objects were placed in a number of lateral positions relative to a visual display. Participants were required to reproduce the haptic shape as a visual object. Haptic objects were also placed at various distances from the body to test for effects on the representation of the object.

4.2.1 Method

4.2.1.1 Participants

Thirteen undergraduate students from Trinity College Dublin participated in the experiment for course credits (8 were female). Their ages ranged from 17 to 28 years, with a mean age of 22.4 years. All of the participants reported normal or corrected-to-normal vision and none reported any tactile impairment. All participants reported to be right-hand dominant.

4.2.1.2 Stimuli

The haptic stimuli consisted of 4 'L' shaped objects described previously (see Experiment 1). The smallest object measured 3cm along its x-axis arm and 3cm along its y-axis arm. The other objects measured 3cm x 5cm, 5cm x 3cm and the largest measured 5cm x 5cm.

A two-dimensional manipulatable 'L' shaped visual image was created as the visual stimulus. The lengths of the visual image's arms could be adjusted in steps of 0.5cm. The minimum possible length of both arms was 1.5cm, the maximum 6.5cm. In total, each arm could be set to 11 different sizes. Participants adjusted the length of the object's arms by using the cursor keys on a keyboard that was located to their far right.

4.2.1.3 Apparatus

The apparatus consisted of a computer monitor that displayed the visual stimulus and was placed 57cm away from the participant at eye level. Participants were required to look through a 57cm long card tube of radius 6cm that was attached to the centre of the monitor. The tube was made of card and was mounted in a circular hole cut in a piece of cardboard. The cardboard was mounted on the monitor and acted to both maintain the tube's shape and to anchor the tube to the monitor. The tube acted to hide the haptic stimuli from view and to prevent the monitor's square frame from being used as a reference frame. The visual object was visible when peering down the tube and was positioned in the centre of the monitor. Figure 13 illustrates the arrangement of the apparatus.

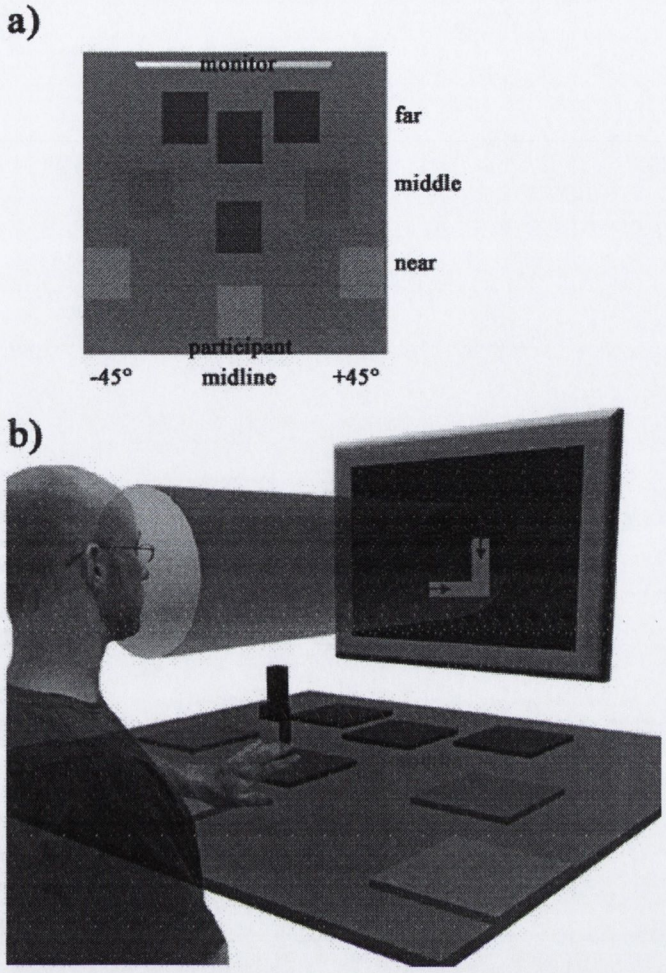


Figure 13. a) A schematic view of the apparatus showing the positions the haptic object could occupy. b) An illustration of the apparatus showing the monitor, supporting structure and card tube through which the participants viewed the visual object. The position markers in both figures are included for illustration purposes only and could not be discerned by touch. Participants felt haptic objects with 2 hands.

Haptic stimuli could occupy 9 different positions arranged in a semi-circular array radiating outwards from the monitor in the transverse plane (see Figure 13a). Haptic object positions fell along three different lines which radiated from centre of the monitor. The 3 lines were at 3 orientations to the monitor, -45 degrees (left), 0 degrees and +45 degrees (right). The centre line (0 degrees) was aligned with the participant's midline. Objects were placed at 3 distances from the participant along these lines (12cm, 27cm, or 42cm along the lines). At each position, haptic stimuli were held in a

vertical orientation 10cm above the table, facing towards the participant so they could be freely palpated.

4.2.1.4 Design

The experiment was based on a 3x3x2 factorial design using repeated measures. The main factors were DISTANCE of the object from the participant (near, middle and far), POSITION of the object in respect to the body-midline (-45°, 0° and +45°) and the AXIS of the objects (the x-axis and y-axis arms of the haptic objects). The dependent variable was the mean error in cm between the adjusted visual object's arm lengths, and its haptic object reference (ERROR). Mean error rates were calculated for both the objects' x-axis and y-axis arms.

There were a total of 72 trials in this experiment (9 positions x 4 objects x 2 visual object size repeats). At each of the 9 positions, four objects were presented, and each object was presented twice at each position. Trials were randomly presented across participants.

4.2.1.5 Procedure

The participant's task was to adjust the shape of the visual stimulus so that it exactly matched the shape of a haptic object. At the beginning of each trial, participants were required to locate the base of the supporting structure of the haptic object with both their hands. A pilot study determined that 3 seconds was sufficient to locate the supporting structure. Participants then rested their hands at the base of the structure until instructed to freely palpate the haptic stimulus. After 3 seconds of haptic exploration of the object they were visually informed to stop feeling the stimulus and to press a key on the keyboard.

The participant's task was then to adjust the length of the arms of a visual object until it matched exactly the shape of the haptic object. No time limit was imposed but participants were instructed to perform the task as fast and as accurately as possible. Once the visual object had been adjusted the participants pressed a key on the keyboard to continue to the next trial. An inter-trial interval (ITI) of 10 seconds followed, during which the experimenter changed the haptic stimulus and placed the supporting structure in the new position for the next trial.

During the course of the experiment each haptic stimulus was presented twice in each location. For one presentation of the haptic object in a particular location, the starting size of the visual object's arms were set to be 1.5cm in length for the horizontal and vertical arms respectively. For the other presentation of the haptic object in the same location, the arms of the visual object were both set to be 6.5cm in length. The reproduced visual arm lengths for a haptic object at a particular position on the haptic plane were calculated by taking the average of these two trials from the visual stimuli.

At the start of the experiment a block of 4 practice trials was given to ensure that the experimenter was satisfied that the participant was familiar with the procedure. The experiment lasted for approximately 45 minutes.

4.2.2 Results

The mean estimation error for each position and for both arms of the objects are listed in Table 4.

Table 7.
Mean error and standard error (in brackets) of measurements of the x-axis and y-axis of the haptic objects relative to the actual measurements for the objects in each position in Experiment 9. Measurements in size were always overestimated.

Distance from the body	Lateral position	x-axis (StdError)	y-axis (StdError)
Far	-45°	0.54 (0.12)	0.24 (0.15)
Middle	-45°	0.54 (0.14)	0.26 (0.14)
Near	-45°	0.32 (0.16)	0.41 (0.18)
Far	0°	0.46 (0.16)	0.14 (0.18)
Middle	0°	0.55 (0.15)	0.38 (0.15)
Near	0°	0.42 (0.14)	0.23 (0.15)
Far	+45°	0.61 (0.11)	0.35 (0.15)
Middle	+45°	0.52 (0.15)	0.29 (0.13)
Near	+45°	0.49 (0.19)	0.33 (0.20)

A 3-way within-subject ANOVA was conducted on object size error with object AXIS (horizontal or vertical), POSITION (-45°, 0°, +45°) and DISTANCE from the body (12cm, 27cm & 42cm) entered as factors. No main effects of POSITION [$F(2,24)=1.15$, p=n.s.] or DISTANCE [$F(2,24)<1$] were found. However, the main effect of AXIS approached significance [$F(1,12)=4.14$, $p=0.065$] with the objects' horizontal axes being non-significantly overestimated relative to the vertical axes.

Sign-tests were also conducted to see if horizontal and vertical estimation error across the different lateral positions were significantly greater than the veridical dimensions of the objects. The length of the x-axis error was estimated as significantly greater than the actual measurement ($p<0.03$), whilst vertical error did not statistically differ from the actual size of the haptic object.

There was no interaction between the AXIS and POSITION factors [$F(2,24)<1$] nor between the POSITION and DISTANCE factors [$F(4,48)=1.00$, n.s.]. However, the AXIS and DISTANCE factors did interact with each other [$F(2,24)=3.98$, $p<0.05$]. A simple effects analysis on the effect of DISTANCE across each level of AXIS revealed no effect of DISTANCE on estimates of the objects' y-axis [$F(2,24)<1$], and no effect on estimates of the objects' x-axis [$F(2,24)<1$].

The triple interaction between the AXIS, POSITION and DISTANCE factors was not significant [$F(4,48)=2.04$, $p=0.10$.].

4.2.3 Discussion

Haptic objects did differ from the visual objects in shape, as the estimates of the objects' horizontal arms were overestimated compared to their veridical measurements. This finding replicates Heller et al.'s (2003) finding that objects orientated in the fronto-parallel plane are subject to a negative horizontal-vertical illusion, and also extends to bi-manual exploration. However, the location of an object in space did not influence

the degree of object distortion as predicted, as object distortions were not affected by an objects lateral position or distance relative to the body (object percepts were however affected by how far they were felt from the body). There was an interaction between the distance of an object from the body and the object's axes but posthoc tests could not isolate any individual effects of this interaction, suggesting that the interaction was quite weak.

There are a number of other confounds which may have hidden a location-specific distortion of the representation of the haptic object. The objects were arranged along a radial grid centred on the visual object on the monitor. A consequence of this is that objects at far distances were laterally closer together compared to objects explored at close distances to the body. Consequently, the fact that objects were located close together at far distances may have reduced any effects of laterality at far distances. In a similar way, distance effects would also be reduced as objects at each specific depth are actually at a range of distances from the observer. In the following experiment, the same lateral positions of the objects were maintained at each distance from the observer.

4.3 Experiment 10

Here, Experiment 9 was repeated, but a square grid was used to position haptic objects. A square grid more closely approximates the experimental setup in Experiments 4, 5 and 7, which enhances the generalizability of any findings in the present experiment to the experiments reported in the previous chapter. To increase experimental power, more participants were tested. Also, participants were all female as it is possible that gender may have obscured experimental effects (although there is little support gender differences, see Summers & Lederman, 1990). Again, it was

predicted that the representations of haptic objects would be subject to a location-specific distortion specifically for objects located to the left of the body-midline and on the body-midline. The participant's task was the same as in the previous experiment, that is, to reproduce visually the shape of a haptic object just explored.

4.3.1 Method

4.3.1.1 Participants

Eighteen undergraduate students from Trinity College Dublin participated in the experiment for course credits (all female). Ages ranged from 19 to 24 years, with a mean age of 21.7 years. All of the participants reported normal or corrected-to-normal vision and none reported any tactile impairment. All participants reported to be right-hand dominant.

4.3.1.2 Stimuli and apparatus

Stimuli were identical to those used in Experiment 9. The apparatus was the same, except that the haptic stimuli occupied 9 positions that were presented as a square array along the horizontal plane in front of the participant (see Figure 13). These positions were 3 positions relative to the body-midline (-15cm, 0cm and +15cm) and 3 distances from the body (i.e., near, middle and far respectively) in steps of 15cm. The centre position of the array was aligned with the participant's midline. At each position, haptic stimuli were held in a vertical orientation 10cm above the table, facing towards the participant so they could be freely palpated.

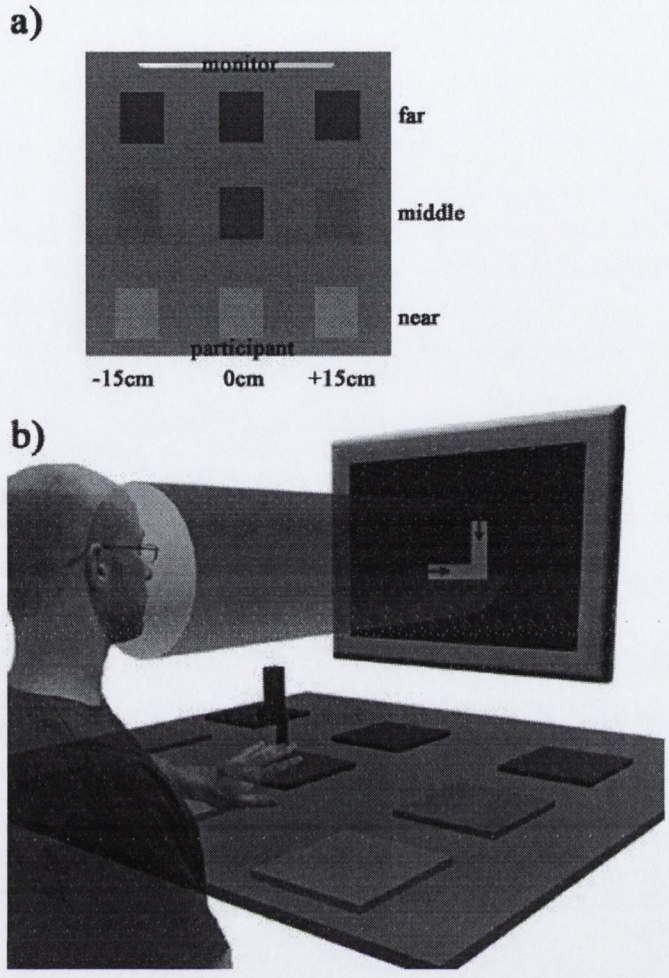


Figure 14. a) A schematic view of the apparatus showing the positions the haptic object could occupy. b) An illustration of the apparatus showing the monitor, supporting structure and card tube through which the participants viewed the visual object. The position markers in both figures are included for illustration purposes only and could not be discerned by touch. Participants felt haptic objects with 2 hands.

4.3.1.3 Design and Procedure

The experiment was based on a 3x3x2 factorial design using repeated measures. The main factors were DISTANCE of the object from the participant (near, middle and far), lateral POSITION of the object in respect to the participant (-15cm, 0cm and +15cm) and the AXES of the stimuli (horizontal and vertical). The dependent variable was the mean error in cm between the adjusted visual object's arm lengths, and the

haptic reference object (ERROR). Mean error rates were calculated for both the objects' x-axis and y-axis arms. Trials were randomly presented across participants.

The participant's task was to adjust the length of the arms of a visual stimulus until it matched exactly the shape of the haptic object. Haptic objects were always explored with two hands. The procedure was identical to Experiment 9 and the reader is directed to Experiment 9's methodology for further information.

4.3.2 Results

Mean error of the estimation of the object shape was calculated for each position, distance and for both arms of the objects (x-axis and y-axis arm estimation error are plotted in Figure 15).

A 3-way within-subject ANOVA was conducted on object shape error with object AXIS (horizontal or vertical), lateral POSITION (-15cm, 0cm, +15cm) and DISTANCE (near, middle and far) entered as factors. There was no main effect of POSITION [$F(2,34)=2.66$, $p=0.09$] or of DISTANCE [$F(2,34]<1$]. However, the main effect of AXIS was significant [$F(1,17)=10.67$, $p<.005$] with the objects' horizontal axes being significantly overestimated relative to the vertical axes.

There was a significant interaction between the AXIS and POSITION factors [$F(2,34)=4.53$, $p<0.02$]. A simple effects analysis was conducted on the effect of POSITION across each level of AXIS and revealed that there was no effect of laterality on the estimation of the objects' y-axis [$F(2,34)<1$]. Lateral POSITION did, however, influence the estimation of the length of the x-axis [$F(2,34)=4.21$, $p<0.05$]. Posthoc Newman-Keuls analysis revealed that objects placed on centre positions were overestimated relative to objects placed to the right (e.g., +15cm, $p<0.02$), but not to objects placed on the left.

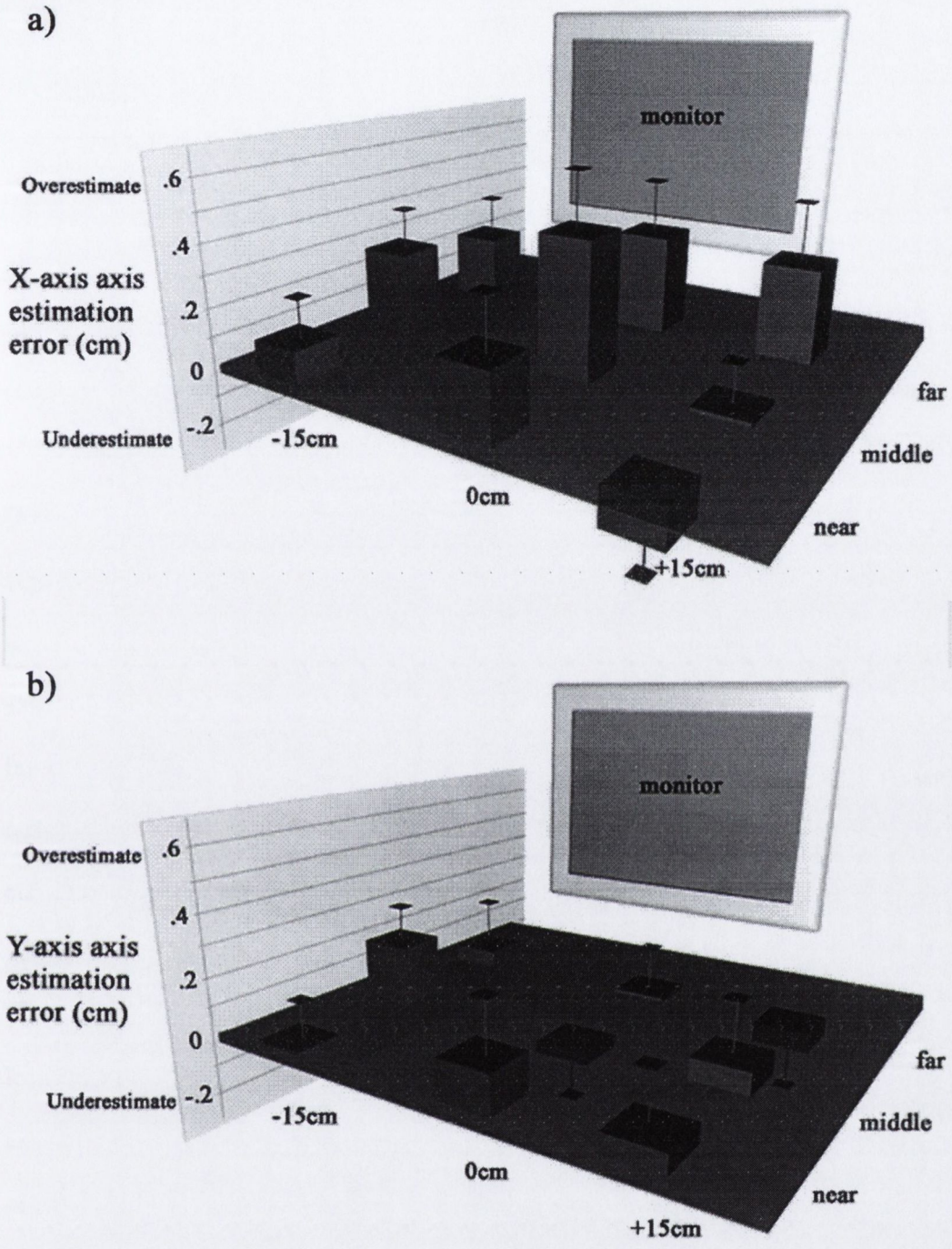


Figure 15. Plot depicting the haptic object shape estimation error at each location for the object's x-axis arms (a) and y-axis arms (b). Error bars depict the standard error of the mean. The scale on the ordinal axis depicts the magnitude of the haptic object shape estimation error and is measured in cm.

A significant interaction between the AXIS and DISTANCE factors [$F(2,34)=14.03$, $p<0.001$] was also found. A simple effects analysis on the effect of

DISTANCE across each level of AXIS revealed no effect of DISTANCE on estimates of the objects' y-axis [$F(2,34)=2.02$, $p=0.15$] but a DISTANCE effect was found on the estimates of the objects' x-axis [$F(2,34)=8.74$, $p<0.001$]. A subsequent Newman-Keuls posthoc analysis revealed that the mean lengths of the objects' vertical/horizontal axes placed at far and middle distances from the participant were overestimated relative to the mean lengths of the objects' axes placed at near distances ($p<0.01$ and $p<0.005$ respectively).

There was no interaction between the lateral POSITION and DISTANCE factors [$F(4,68)<1$]. The triple interaction of AXIS, POSITION and DISTANCE was significant [$F(4,68)=3.02$, $p<0.005$]. To tease this triple-interaction apart, horizontal and vertical estimation data were treated as if they had been collected in separate experiments, and were thus re-analysed with POSITION and DISTANCE entered as factors in both analyses. The results from these subsequent analyses were in concordance with those already discussed however, so they will not be detailed (i.e., effects of laterality and distance on the objects' horizontal axes, but not on the objects' vertical axes).

Summarising so far, the estimates of the length of an object's x-axis is subject to location-specific estimation error. The interaction between an object's lateral position and its distance from the observer was found to effect estimates of the length of the object's x-axis. A test of whether the error in the length of the x-axis estimate was actually significantly different from the actual size of the object was conducted. Sign tests found that the estimate of the x-axis of the object for objects located at the middle-left position was greater than the veridical length of the axis ($p<0.05$), whilst the estimates of the x-axis of objects located on the midline at far and middle distances

were not significantly different from veridical. Sign tests did not find that the estimates of the objects' vertical axes differed from that predicted from veridical measurements.

4.3.3 Discussion

Here, the position of an object in space was found to affect the percept of the estimate of the x-axis dimensions of that object. Specifically predicted was that the percepts of haptic objects located to the right of the body-midline should not be distorted whilst the percepts of objects at other locations would be distorted. However, objects were only distorted when placed in the centre-left position and only their horizontal axes were estimated incorrectly. Furthermore, an object's x-axis was overestimated with increasing distances from the body. This finding was unexpected.

It is interesting that it was when the right hand had to reach to the centre and across the body that haptic objects were distorted. In other words, estimates of object dimensions were not different from veridical when the objects were positioned on the right. Seemingly, this location-specific distortion supports Miller & al-Attar's (2000) hypothesis that a reach across the body by the hand will involve more whole-arm movements and lead to more distortion (see also Heller et al., 1997). However, it is impossible to say for certain if object distortion was caused by whole-arm movement. As exploration was bi-manual, it is likely that the left hand would have significantly contributed to the exploration of the x-axis arms of the objects, as the objects' x-axis arms were to the left of the objects (e.g., Fagot et al., 1994): objects should actually have been most distorted when located on the right if Millar & al-Attar's midline reach hypothesis was true. As the participants were all right-handed however, it is possible that the left hand played a less significant role in exploring object horizontals than it would have done for left-handed participants.

To attempt to identify the underlying causes of this location-specific distortion, future research should explicitly measure the contributions of each hand used to the explore the objects. It also would seem necessary to restrict haptic exploration to deduce how particular exploratory strategies influence a distortion of the objects representation. By using splints and rigid gloves (e.g., Heller et al., 1997; Lederman & Klatzky, 2004) it may also be possible to further deduce what processes are causing the distortion of 3D objects positioned in the fronto-parallel plane.

4.4 General Discussion

In Chapter 3 it was found that haptic objects located to the right were harder to match to visual objects than were haptic objects located in other positions. One explanation was that the percepts of haptic objects were distorted, depending upon the object's location in respect to the body-midline. Here this location-specific distortion was tested for. Experiment 9 found no evidence for a location-specific haptic object distortion, but did find that that the x-axis of objects were overestimated, replicating Heller et al.'s (2003) recent finding of a negative horizontal-vertical illusion for objects positioned in the fronto-parallel plane, and extending the finding to bi-manual exploration.

In Experiment 10 a square matrix was used for the placement of haptic objects. The location of a haptic object in space did seem to influence the perception of object shape. Haptic objects located to the centre-left were found to be the most distorted. This partially supports the hypothesis suggested in that rightward located objects were not perceived as being distorted whereas objects positioned on the left of the body-midline were perceived as distorted. As objects located to the right were not perceived to be distorted, they were thus more likely to be integrated with their visual counterparts (see the previous chapter). However the results here do not explain fully the results in

the previous chapter since objects positioned on the body-midline were not distorted and therefore should have been more likely to have been integrated in previous experiments but this was not found.

There are two potential explanations for this centre-leftward object distortion and the absence of object distortion for objects located on the midline. Firstly, inspection of Figure 14b does suggest that haptic objects located on the body-midline at middle distances were distorted along their x-axis arm, yet the experiment was not sufficiently powerful to detect this. Secondly, because of methodological differences, the findings here may not generalise completely[6] to those predicted to occur Chapter 3, but that the identification of a location-specific distortion may still, albeit tentatively, support the initial hypothesis that location-specific distortion may be responsible for the findings in Chapter 3.

4.4.1 Reasons why additional haptic object distortion was not detected

The visual assessment of the representation of a haptic object may have influenced any effects due to a distortion of the representation of the haptic object. For example, any effects of distortion may have occurred as a consequence of combining both haptic and visual distortion. L-shaped objects are subject to a positive horizontal-vertical illusion in vision: the y-axis arm of a visual object is generally perceived as longer than its equally sized x-axis arm in vision (Fick, 1851). Cohen & Girgus (1978) argue that when a figure is subject to two illusory phenomena, the reported distortion is in fact a combination of the effects of the individual illusions. The combination of opposite visual and haptic illusions would drastically reduce the strength of any haptic

---

[6] Haptic objects at middle distances from the observer in this experiment, and in Experiments 4, 5, and 7 in the previous chapter, were both approximately 30cm from the participant which does suggest that results here may generalise, at least in terms of object-body distance.

distortion as the illusory effects would interfere with each other. This may explain why only small location-specific distortions were found. A simple compensatory measure for future research would be to assess the strength of the visual object distortion for each object at the start of the experiment and subtract this effect from the combined visual-haptic object distortion (in the visual illusion, vertical extents are typically perceived as 10% longer than horizontal lengths: Avery & Day, 1969). Such interference, however, would not explain why objects located on the midline were not overestimated. As the visual object remained in the same location throughout the experiment, there is no reason to suspect that the visual illusion would differ over trials. Visual and haptic illusions would have to combine in a non-additive manner to produce location specific distortion and there is no reason to suggest that this is the case.

4.4.2 Experimental validity

There were some differences between these experiments and those in Chapter 3 assessing cross-modal integration, including the height at which haptic objects were presented at in respect to the participant, the location of the visual stimuli, and the spatial separation between different haptic object locations. These small differences between the experimental layout in this chapter and the layout in Experiments 4, 5 and 7 may have contributed to the difference in findings across these studies. A crucial difference between this chapter and the last was that haptic objects here were placed at table height, whereas haptic objects in the preceding chapter were placed at eye height. By placing objects at eye height, the participants were forced to raise their elbows off the table to explore objects. When the objects were at table height however, their elbows could rest on the table. Interestingly, Heller et al. (1997) tested this manipulation. The authors found that when the elbows were anchored on the table, object distortion was reduced. They argued that this was because whole-arm

movements were prevented. In other words, Heller et al. suggested that when the elbows were down on the table that this *prevented* haptic object distortion. The consequence for the findings is that the distortion of the representation of the haptic objects potentially may be much less here than that predicted in the previous chapter, and potentially the chapters may differ in a qualitative manner in regards to distortion. Actually finding a location-specific distortion is therefore encouraging, and does lend support that the previous chapter's findings may have been influenced by haptic object distortion.

I was also interested to see whether location-specific haptic object distortion would be modulated by the distance objects were from the body. In Experiment 10 it was found that objects, which were located 15cm from the body, were not affected by location-specific haptic object distortion but that a further increase in distance from the body did introduce more error. Ideally then, haptic objects in the fronto-parallel plane should be a maximum of 15cm away from the body to avoid issues due to location-specific distortion. However, it would be necessary to see if this distance was robust especially once the above issues have been taken into consideration.

CHAPTER 5: GENERAL DISCUSSION

Abstract

The findings from this thesis are discussed in terms of the initial hypotheses and in terms of their importance for the different models of integration. One suggestion is that haptic object percepts are automatically recoded into visual percepts allowing for efficient cross-modal integration. An alternate interpretation of the findings is also discussed, that object information from vision and haptics is always integrated despite being spatially and temporally separate. However, this interpretation is argued to be tentative. Finally, the contributions this thesis makes to the literature are highlighted and suggestions are made about potential avenues of future research.

5.1 Summary of findings

This thesis was concerned with how temporal and spatial factors influenced the cross-modal integration of objects. Specifically tested was whether temporal synchrony and spatial co-location are crucial for the integration of visual and haptic object information in the brain. It was hypothesised that objects would be integrated if they were presented in temporal synchrony and were spatially co-located (e.g., Ernst & Bülthoff, 2004; Lalanne & Lorenceau, 2004; Stein & Meredith, 1993; Wallace et al., 2004).

A necessary step before testing these predictions was to deduce how visual and haptic object information is maintained in memory, be it as a multisensory percept or in a modality-specific manner.

5.1.1 Type of memory representations

Memory decay rate across modalities was found to be equivalent (Experiments 1 & 2) suggesting that visual and haptic object representations are stored either in a multisensory manner or stored independently but in a functionally similar manner. An alternative suggestion was that haptic objects may be recoded into visual co-ordinates and held as visual representations according to the demands of the task (i.e., visual and haptic objects are thus both held within visual memory). Haptic memory was found to be less efficient than visual memory in the within-modal delayed matching task used in Experiment 2. However, in a similar cross-modal matching task used in Experiment 1, the order at which the modalities encoded objects was found not to be significant, as would be expected if haptic information was visually-recoded and stored in visual memory in the cross-modal matching task, but not recoded for the within-modal haptic task (i.e., visual memory is more capable at maintaining information than haptic memory, therefore, a percept held in visual memory will be better recalled over delays, irrespective of the modality at test). This idea that haptic information is recoded into visual co-ordinates has much support in the literature. For example, Zhang, Weisser, Stilla, Prather, & Sathian (2004) suggest that haptic objects are 'visualised'. Some neuroimaging studies have implicated the involvement of visual areas in haptic object processing, such as the lateral occipital cortex (LOC) area of the temporal lobe (Amedi et al., 2001; James et al., 2002), and even primary visual areas (Zangaladze et al., 1999). This shall be discussed in more detail later.

5.1.2 Temporal factors

Experiment 3 was designed to further elucidate how visual and haptic object percepts are represented in memory. In this experiment, objects were presented simultaneously or in a serial fashion. If temporal synchrony was necessary for

multisensory integration then it was predicted that objects presented simultaneously would be mandatorily integrated and thus be harder to discriminate, even if there were slight differences in shape. However, there was no evidence that simultaneously presented objects were more difficult to discriminate than serially presented objects. This suggests that visual and haptic object percepts are held independently (i.e., the percepts can be easily contrasted as they are not integrated together). Integration perhaps did not occur automatically as some criterion was not met. One such criterion is that objects may need to be co-located in space in order to be integrated (e.g., Ernst & Bülthoff, 2004; Stein & Meredith, 1993). As objects were not co-located in Experiment 3, a difference in the location of the objects may have prevented the objects from being integrated. The next hypothesis tested was whether temporal synchrony and spatial co-location were necessary for percepts to be visually-recoded or for multisensory representations of objects to be integrated[7].

5.1.3 Spatial factors

The hypothesis tested in the next experiments was that if objects must be temporally synchronous and spatially co-located to be integrated across modalities, two different objects presented simultaneously and at the same spatial location would be harder to discriminate as they would be combined as a multisensory percept, compared to objects which were spatially separated. However, synchronous co-located objects were not harder to discriminate (Experiment 4), suggesting that objects were not integrated when co-located, which goes against popular opinion in the literature that co-

[7] An alternative explanation not considered in chapter 2 was that visual and haptic objects were indeed stored in a multisensory manner; however they were not integrated together. This will be discussed in more detail later.

location is crucial for cross-modal integration (e.g., Ernst & Bülthoff, 2004; Lalanne & Lorenceau, 2004; Stein & Meredith, 1993; Wallace et al., 2004).

A possibility was that some top-down mechanism acted to inhibit the visual and haptic object percepts from being integrated. This top-down mechanism may have detected that the assumption of unity was violated (Welch and Warren, 1980) in that the visual and haptic stimuli were quite different in appearance from each other (visual stimuli were 2D whilst haptic stimuli were 3D). However, rendering the visual and haptic objects as more similar to each other did not increase the probability of integration occurring (Experiment 5). Similarly, there was a concern that visual and haptic objects may have been perceived differently according to their size. However in Experiment 6 it was established that visual and haptic objects were perceived to be identical in size. Lastly, a difference in visual and haptic encoding or sensorimotor patterns during encoding may have made the object percepts different in appearance such that they were not integrated. However, after haptic encoding was equated with visual encoding (haptic encoding was made more optimal), there was still no evidence that co-located objects were integrated (Experiment 7). Thus no evidence for these top-down factors (i.e., relating to the assumption of unity) which prevented object integration were found.

Conceivably visual and haptic objects may indeed have been integrated, but not at the spatial location at which they were initially perceived as co-located. Previously the horizontal position of the haptic object in space was found to influence matching performance (Experiments 4 and non-significantly in Experiments 5 and 7) and there was some evidence (a non-significant trend) that haptic objects placed to the right were harder to match in shape with their visual counterparts. It may have been that at these rightward haptic object locations, visual and haptic object percepts were being

integrated. This was supported by the finding that the point of co-location was slightly rightward of the veridical central location (Experiment 8).

However, the reduced matching performance for rightward-located haptic objects may also be because haptic objects underwent a distortion which was dependent upon their location in space (Chapter 4). Specifically, L-shaped objects explored with two hands were found to be affected by a negative horizontal-vertical illusion (Experiment 9, which extends the findings from a similar one-handed experiment, Heller et al., 2003). Importantly, this haptic object distortion was modulated by the haptic objects location in space (Experiment 10). Thus, it is not clear whether the objects located to the right in the previous experiments were harder to match due to location-specific object distortion, or due to a rightward shift in location of where integration between visual and haptic objects was taking place. In any case, evidence for the reduction in matching performance is tentative since this was not a consistent finding across the experiments. Also, as none of the Experiments 4, 5 and 7 found that haptic objects located to the right exhibited significantly worse matching performance than haptic objects at other locations, spatial co-location would seem not to be a necessary condition for cross-modal integration.

To summarise, there was no direct evidence that haptic objects were integrated with visual objects when objects were either temporally synchronous (Experiment 3), spatially co-incident (Experiment 4), highly similar in terms of number of dimensions (Experiment 5) and optimally encoded (Experiment 7). Although there was some partial evidence implicating these factors as criteria for integration (Experiment 8), due to the ambiguity of this evidence (Chapter 4 found a potential alternative explanation), it seems at this point that these factors are not necessary to promote cross-modal integration. However, there is an alternate interpretation of these findings.

5.2 A model of cross-modal integration

One suggestion for how cross-modal objects are matched is that haptic *object* percepts are automatically visually-recoded, allowing for direct comparison across modalities. That is, visual and haptic representations of objects in general may be indiscriminable in memory. Alternatively it is also argued that visual and haptic percepts may indeed be mandatorily integrated when spatially co-located and temporally synchronous as a multisensory percept. In the present experiments, this may not have occurred (as the objects were not perceived as co-located) because the hand could not be visually perceived exploring the haptic objects, which potentially explains why there was no evidence for mandatory integration (i.e., proprioception was not an available cue for whether visual and haptic objects were co-located).

5.2.1 Automatic visual-recoding of haptic object percepts

The results of Chapter 2 suggested that haptic objects may have been recoded into visual co-ordinates and held as visual representations (Experiments 1 and 2). It was also suggested that this recoding was not mandatory as there was no evidence that simultaneously presented visual and haptic objects were integrated together (Experiment 3). However, it may be that in a cross-modal context, haptic *spatial* information (i.e., object shape information) is mandatorily visually-recoded and processed using visual mechanisms, perhaps because mechanisms traditionally deemed visual are in fact dedicated to processing spatial information irrespective of encoding modality (Pascual-Leone & Hamilton, 2001). Non-spatial haptic information would then not be recoded as it is best processed by the somatosensory cortex (see Figure 2 in Chapter 1 for an illustration of other modality encoding biases). If it is considered that some information pertaining to a haptic object will be processed by visual mechanisms, and some will be processed by somatosensory mechanisms, then this suggests that

information is processed in a way akin to Millar's (2000) interrelated networks model. In this model Millar explains that information is processed and held across a number of interrelated networks, such that cross-modal information is sometimes integrated, but at other times independently held.

Haptic percepts may indeed be automatically recoded, yet in Experiments 3, 4, 5, and 7 objects were not found to be mandatorily integrated. However, it may be that there were still sufficient differences between the object percepts to allow discrimination to occur. A novel interpretation of cross-modal integration can now be entertained: cross-modal integration may be subject to the same laws governing the *within* vision integration of object features into a single percept. Kahneman, Treisman, & Gibbs (1992) suggest that the percepts of visual objects in the environment are held independently in memory as object files. Object features are grouped into object files in terms of "spatiotemporal contiguity of low-frequency information", a step which helps to ensure that it is only object features that belong to the same object which are grouped in the same object file. The authors termed this the correspondence process. Although it is beyond the scope of this thesis to speculate further on how within-modal object feature integration is achieved in vision, it is interesting that unimodal visual integration also seems dependent on temporal and spatial factors (Kahneman et al., 1992; see also Stark, Anes, & Liederman, 2004). Research here does suggest however that within-vision integration and cross-modal integration differ as it seems that in cross-modal integration some top-down process is also involved. This is discussed further in the next section.

5.2.2 Amodal or multisensory integration of visual and haptic inputs

An alternative suggestion is that visual and haptic information pertaining to the same object may be represented as multisensory or amodal information in memory but

that all information does not completely overlap (i.e., haptics may encode the back of an object, whilst vision encodes the objects front: Newell et al., 2001). Cross-modal information referring to the same object dimension encoded by both modalities will be integrated. However, cross-modal information that does not overlap may also contribute to the cross-modal object percept, making it more rich than that derived from a single modality and thus benefiting object recognition (Newell et al., submitted).

For both accounts, it is speculated that the Bayesian Maximum likelihood model (Ernst & Bülthoff, 2004) may account for cross-modal integration, but it may need to be modified to take into account potential top-down effects which were suggested earlier, which likely influence cross-modal integration (such as in Experiments 3, 4, 5 & 7). Consider that the site of integration is within visual cortical areas (i.e., this site is Ernst and Bülthoff's third multisensory population of neurons) and information here is integrated as suggested by the MLE model (i.e., in a multiplicative manner). Top-down factors may affect cross-modal integration if such factors act to modulate the information which feeds into these multisensory areas, perhaps directing modality specific information to independent object files (which is acceptable under the model). It is also likely that intermediate layers may be required to incorporate more complex top-down mechanisms within the neurological MLE model (consider also that primary cortical areas may interact with each other, implying horizontal transfer between unimodal neural populations, Zangaladze et al., 1999; Zhou & Fuster, 1997).

5.3 Top-down factors affecting cross-modal integration

The question remains, however, as to why objects in Experiments 4, 5, and 7 that were temporally synchronous and spatially co-located were not integrated together? One reason was that some top-down factor is also implicated in cross-modal integration. Integration may have been prohibited as the stimulus-display violated the

assumption of unity (Experiment 5, Welch & Warren, 1980) as logically objects would only be integrated if they were deemed similar enough in appearance. However, increasing the assumption of unity of the stimulus-display did not promote cross-modal integration. Some additional factor must therefore have acted to prevent information from being integrated.

5.3.1 The visual hand hypothesis

Chapter 3 discussed how participants could not see their hands exploring the haptic objects and that proprioception may have been a useful cue for integrating the object percepts. If visual and haptic integration relies on seeing the hand exploring the object (the "visual hand hypothesis"), then the absence of seeing your hand exploring the object may act as a powerful cue that two separated objects were encoded (i.e., "the hand is not present, therefore I am seeing and feeling different objects"). Welch and Warren (1980) also suggest that a lifetime of experience of seeing and feeling the hand in the same place would mean that occluding the hand from view whilst exploring an object in both modalities would strongly reduce the unity assumption. Consequently, in Experiments 4, 5, and 7, two spatially co-located objects would always be considered as separate objects since the hands were not seen exploring the haptic object. This might explain why there was no evidence that 'co-located' objects were mandatorily integrated, as the objects may have been perceived as distinct.

But what benefit would sight of the hand offer for cross-modal integration? Interestingly, Heller (1989) found that Braille recognition was hampered when the Braille was tilted, yet when the fingers could be seen, the effects were ameliorated. Heller suggests this is because the task would be easier in the prescience of vision as the Braille orientation in respect to the body would be more easy to determine, presumably allowing the perceptual system to compensate for Braille tilt (i.e., if you know Braille is

tilted to such an angle, you can compensate for the tilt; however if you do not know the tilt angle, it is much harder to compensate for). However, it has recently been discovered that non-informative vision may also be used to ground haptic information within an allocentric reference frame (Newell, Finucane, & Pasqualotto, 2004). It is possible that participants in the studies presented in this thesis used non-informative vision to anchor haptic information to allocentric cues (participants were afforded some peripheral vision and perhaps could make out their chest whilst looking into the head-mounted display).

5.3.2 Other evidence for the role of seeing the hand in multisensory integration

In support for the role of the visual hand hypothesis, Làdavas and colleagues found that the visual position of the hand in space is much more vital for certain cross-modal effects compared to the proprioceptive position of the hand (for reviews, see Làdavas, 2002; Làdavas & Farnè, 2004). In one study, Làdavas, Farnè, Zeloni, & di Pellegrino (2000) were interested in the effects of the sight of the hand on visual extinction. Visual extinction refers to the finding that one of two stimuli presented simultaneously to each hemispace will typically be ignored by a patient with unilateral neglect. Làdavas et al. found that the effects of extinction were reduced when the patient held their hands in close proximity to each stimulus in each hemifield (the activation of a spatial representation in the intact tactile modality acted to reduce the effects of the damaged visuo-spatial representation: Làdavas, di Pellegrino, Farnè, & Zeloni, 1998). Crucially, however, there was no reduction in extinction when the patient's hands were obscured from view but were nevertheless close to the stimuli. The authors concluded that proprioceptive location of the fingertip is not a crucial factor for the reduction in extinction effects compared to the visual sight of the hand. Another demonstration of the role of vision in hand localisation is the so called "rubber

arm illusion". That is, a rubber hand may be perceived as an observer's own hand, even though the rubber hand and visually occluded real hand are spatially separate (Botvinick & Cohen, 1998; and see also Pavani, Spence, & Driver, 2000). Neurophysiological studies have also suggested a role of vision in hand localisation. For example MacKay & Crammond (1987) found that some bimodal neurons in the posterior parietal cortex ceased firing when the limb was occluded from view, suggesting that neural mechanisms exist that are in correspondence with the above. Others have found however that both vision and proprioception contribute to the location of the hand (Graziano, 1999; Graziano, Cooke, & Taylor, 2000; Haggard, Newman, Blundell, & Andrew, 2000; Rossetti, Desmurget, & Prablanc, 1995; van Beers, Sittig, & Gon, 1999). Even so, it may be that both visual and proprioceptive information needs to be considered for objects to be integrated. Graziano (1999) found that neurons in the premotor cortex of monkeys responded to both the proprioceptive position of the hand in space and also to visual stimuli which moved in close proximity to the hand. However, when the hand was hidden from view, the activation of these neurons was reduced (see also Mattingley, Driver, Beschin, & Robertson, 1997). If it were considered that a similar population of neurons controlled whether or not cross-modal integration took place (or whether information should be directed to the same object file), hiding the hand from view in a cross-modal task may potentially have acted to reduce the activation of these neurons below a certain threshold preventing objects being integrated.

5.3.3 Suggestions for future work

This presents an interesting methodological problem for future studies, namely, that participants must be able to see their hands interacting with an object, yet see and feel different objects. This may be achievable using state of the art augmented reality

techniques. Moreover, a partial view of the hand may be sufficient for cross-modal effects (Welch & Warren, 1980), which would reduce the complexity of the problem of simulation. It is also worth considering that fake arms elicit the same reduction in extinction effects for patients with tactile extinction (Farne, Pavani, Meneghello, & Ladavas, 2000; in a monkey study: Graziano et al., 2000). Passive touch could be used rather than active touch, i.e., take pictures of the hand passively holding objects. Lastly it also seems that a pictorial representation of the hand may be sufficient to influence cross-modal effects (Igarashi, Kitagawa, & Ichihara, 2004).

5.4 Other potential factors affecting cross-modal integration.

It is reasonable to suggest that objects of radically different shapes should not be integrated from across the modalities. Indeed, it seems likely that there would be some mechanism in place to prevent such integration, the consequences of which perhaps is that the perceptual system would be more sensitive to object shape change when the criteria of cross-modal integration are met. For example, it has recently been demonstrated that the visual and haptic discrimination of curvature distortion in the Z-dimension is more sensitive when curves are co-located compared to when spatially separated (Jansson & Öström, 2004). However, when object shapes are very similar across modalities integration may be more likely. This is exemplified by the well studied phenomena of visual capture (e.g., Rock & Victor, 1964), where it is the participants' perception that they are seeing and feeling one object, yet the visual percept is a distraction of the true haptic shape.

Some studies suggest that participant's expectation of whether objects will be similar or different affect integration. Miller (1972) directly tested the effects of shape discrepancy awareness and found that visual capture only occurred when the participant expected they were dealing with a single object. When the participants *knew* they were

dealing with 2 objects, visual capture did not occur. Similarly, Warren (1979) also found that when participants were not explicitly told that differently shaped objects presented to vision and touch were the same (presumably allowing the participants to deduce that they were exploring two different objects), visual capture was not found. According to the results presented in the experiments in this thesis, participants' responses were not 100% accurate in the different trials suggesting that participants were not aware of shape discrepancy across all trials. This finding could imply that the studies reported here and those of Miller and Warren are not subject to the same effects of discrepancy awareness.

5.5 Limitations of the studies

5.5.1 The role of the position of co-location

An interesting possibility is that the predicted effects of integration were cancelled out in Experiments 4, 5 and 7, by some unforeseen confound. For example, the predicted effect of integration was that matching performance would be reduced as different objects would be harder to determine. If matching performance is facilitated at the body-midline (i.e., the position of co-location in most experiments), this effect and the predicted effect of integration may have cancelled each other out. There is some evidence for improved performance in a symmetry detection task when stimuli are positioned along the body-midline, and also evidence that stimuli are less susceptible to illusory effects when located on the body-midline as they are more easily related to a body-centred frame of reference (Ballesteros, Manga, & Reales, 1997; Ballesteros & Reales, 2004; Millar & al-Attar, 2000). However, it remains to be seen whether haptic objects placed on the midline in the experiments here were easier to discriminate with visual objects due to some facilitation effect. A compensative

measure in future studies would be to vary the location of both visual and haptic objects so that they may be co-located both at the body-midline and in other positions.

5.5.2 Visual-haptic ventriloquist effects

These results could also be interpreted to suggest that visual and haptic object percepts were *always* integrated, irrespective of simultaneous presentation or serial presentation (i.e., temporal visual-haptic ventriloquist effect), and also irrespective of the haptic object's position in space (i.e., a spatial ventriloquist effect). It is argued however that this interpretation is tentative at best.

Visual and haptic object percepts may have been integrated when they were simultaneously presented *and* when they were presented in a serial manner. This would explain why there was no difference in matching performance between the simultaneous and serial conditions in Experiment 3. The introduction discussed how complex objects may have larger temporal windows of integration compared to more simple objects. For example, visual and auditory information in the McGurk effect may be presented separately within a 180ms temporal window, whereas simple stimuli in a temporal order judgement (TOJ) task appear to have a temporal window of only 20ms outside of which TOJ judgements are easily achieved (Hirsh & Sherrick, 1961; Munhall et al., 1996). It is possible then that the relatively complex objects that were presented in a serial manner may have also been integrated as their temporal window of integration was greater than the serial presentation time used in Experiment 3.

But consider the argument presented previously that objects must be spatially and temporally co-incident for objects to be integrated. As objects in Experiment 3 were spatially separate, it follows that some form of spatial ventriloquism (i.e., that the objects were perceived as co-located even though they were not) must have occurred if this ventriloquism explanation is to be valid. Thus, for both serial and simultaneous

presentations, a visual-haptic temporal ventriloquism effect may have acted to bind the two objects together (Howard, 1966; an auditory-visual temporal ventriloquist has recently been identified by Morein-Zamir, Soto-Faraco, & Kingstone, 2003). It may be that the temporal window of integration may be more dynamic then previously thought (Spence & Squire, 2003).

5.5.3 Reference frames

Previously, proprioception was argued to be less important at determining the position of the hand in space relative to vision (M. S. Graziano, 1999; Mattingley et al., 1997) and may not even contribute to hand position for certain cross-modal tasks (Botvinick & Cohen, 1998; Làdavas, 2002; Làdavas & Farnè, 2004; 1987; Pavani et al., 2000). Information from proprioception about hand location in Experiments 4, 5, and 7 may not have been relevant to the task. Instead a hand-centred reference frame, that is invariant to the position of the hand in space, may have been used to encode the objects. As such, visual and haptic object percepts would *always* be co-located, and therefore were *always* integrated. The evidence for such a proposal is discussed below.

Kappers and colleagues recently found that haptic space is encoded in a reference frame that is intermediate to an allocentric and hand-centred frame[8] (for reviews see Kappers, 2004; Kappers & Koenderink, 2002). For example, in one study, Volčič, Kappers, & Koenderink (2005) found that when participants were required to rotate a bar in the frontoparallel plane rotated so that it was parallel to a reference bar, errors were consistently made in that the bar was rotated clockwise from parallel. Furthermore, the greater the spatial separation between the test bar and the reference bar, the greater the magnitude of this error. Kappers argued that the orientation at

---

[8] Kappers actually terms this as an egocentric reference frame. However, for purposes of clarification and consistency, this shall be referred to as a hand-centred reference frame.

which the hand lies at different locations in space may partially explain this error. The more rightward the hand is positioned from the body-midline, the more clockwise the hand is rotated during encoding due to the mechanics of the shoulder/elbow/hand joints. Consequently a greater degree of rotation error of the bar would occur if haptic perception of orientation relies solely on hand-centred reference cues. Kappers (2004) also suggested that allocentric cues do partially contribute to the perceived orientation of the bar as rotation errors were less than that predicted by hand-centred reference cues alone. She concluded that an intermediate reference frame is used to represent orientation in a haptic stimulus. If this was the case, then in the experiments it could be expected that objects which were presented at positions which were furthest away from the body-midline (irrespective of direction) would be distorted and thus more easily discriminable from their visual counterparts. This result was not found, however, in Experiments 4, 5, and 7.

5.5.5 An Alternative theory

Research here was based on an indirect theory of perception, where cognitive effort is required to encode and process information. It is important to also consider that there are other theories of perception which can also address how we make use of multisensory information in the environment. In the ecological theory of perception advocated by Gibson (e.g. 1966), perception is based on the recognition of useful information in the environment. It is not necessary that information be combined together as a percept as the information relating to an object already exists in the environment and as such, the importance of spatial and temporal congruency would seem unclear. However, it seems likely that temporal and spatial co-incidence are also crucial in direct theories of perception as they may provide additional 'invariances' that may be detected in the sensory stream arriving at the observer (Gibson, 1966). It would

seem then that irrespective of the type of theory, such factors play a crucial role in how we perceive the world.

5.5.5 Ecological validity of the theory

It may be argued that the research here is not applicable to everyday experience as in the natural environment, it rarely happens that a single object be experienced at different times and locations across the modalities. Here, two objects should be perceived. It was the intention of this thesis however to test whether it was such temporal and spatial disparity which led to the perception of two objects. The methodology in this thesis does differ somewhat from everyday experience as different objects that were temporally and spatially co-located were expected to be perceived as one object. This, coupled with the findings here that temporal and spatial congruency does not lead to the perception of a single object would seem to indicate that this research superficially is not applicable to the real world. The underlying theory however has been strongly argued to be directly applicable throughout this thesis. This research does however have direct applicability to artificial environments. Specifically, a number of measures were suggested to make artificial environments more realistic (e.g. the 'visual hand' hypothesis suggested a virtual hand may facilitate visual-haptic object integration).

5.6 Conclusions and contributions of this work to the literature

The results of this thesis suggest that depending upon both bottom-up factors (sight of the hand, temporal synchrony and spatial co-location) and top-down factors (assumption of unity), visual and haptic information within memory may be held

independently (i.e., two objects are perceived) or combined in a statistically optimum fashion according to the MLE model (Ernst & Banks, 2002).

It is interesting that visual-haptic cross-modal matching performance was so low. For example, participants in Experiment 5 on average only managed to identify that different objects were in fact different 71% of the time even though objects differed from each other by 2cm along one arm. The implications of such findings are far reaching. For example, it is worrying that a surgeon could operate on a patient who is hundreds of miles away through tele-surgery (e.g., Burdea, 1996a, 1996b) and not recognise potentially valuable discrepancies between what is seen and what is felt (e.g., haptics reveal unwanted inconsistencies in soft tissue, unseen by vision). However, one limitation of the studies presented here is that the participants could not see their hands. It seems unlikely that bi-modal exploration where the hands can be seen would be similarly bad at detecting differences in visual and haptic shape. This research thus acts to highlight the importance for incorporating the (virtual) hand into cross-modal studies or indeed into designs involving augmented reality.

There are a number of other contributions the studies reported here as part of this thesis could make to research. For example, in Chapter 2 it was discussed that stimulus complexity may explain the discrepancies in the literature on haptic memory decay. In Chapter 3 the rightward bias found in the perception of co-location between visual and haptic objects may have been influenced by proprioception, pseudoneglect, or even hand dominance. This finding may have consequences for research on augmented reality when the virtual hand must be exactly aligned with eye. In Chapter 4 it was discussed that haptic object representations are distorted depending upon their horizontal position in space, extending the finding from a similar study where objects were explored with one hand (Heller et al., 2003) to two-handed exploration. Also

identified was that haptic objects should be explored at a maximum of 15cm away from the body to avoid issues due to location-specific distortion. Together, these findings suggest improvements and caveats on the factors that affect the integration of object information across vision and haptics.

REFERENCES

Amedi, A., Jacobson, G., Hendler, T., Malach, R., & Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex. *Cerebral Cortex, 12*, 1202-1212.

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience, 4*, 324-330.

Armstrong, L., & Marks, L. E. (1999). Haptic perception of linear extent. *Perception and Psychophysics, 61*, 1211-1226.

Atkins, J. E., Fiser, J., & Jacobs, R. A. (2001). Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vision Research, 41*, 449-461.

Avery, G. C., & Day, R. H. (1969). Basis of the horizontal-vertical illusion. *Journal of Experimental Psychology, 81*, 376-380.

Baek, M. J., Lee, B. H., Kwon, J. C., Park, J. M., Kang, S. J., Chin, J., et al. (2002). Influence of final search direction on tactile line bisection in normal subjects. *Neurology, 58*, 1833-1838.

Ballesteros, S., Manga, D., & Reales, J. M. (1997). Haptic discrimination of bilateral symmetry in 2-dimensional and 3-dimensional unfamiliar displays. *Perception and Psychophysics, 59*, 37-50.

Ballesteros, S., & Reales, J. M. (2004). Visual and haptic discrimination of symmetry in unfamiliar displays extended in the z-axis. *Perception, 33*, 315-327.

Banati, R. B., Goerres, G. W., Tjoa, C., Aggleton, J. P., & Grasby, P. (2000). The functional anatomy of visual-tactile integration in man: a study using positron emission tomography. *Neuropsychologia, 38*, 115-124.

Bartley, S. H. (1953). The perception of size or distance based on tactile and kinesthetic data. *The journal of Psychology, 36*, 401-408.

Bertelson, P., & de Gelder, B. (2004). The psychology of multimodal perception. In C. S. a. J. Driver (Ed.), *Crossmodal space and crossmodal attention* (pp. 141-177). Oxford: Oxford university press.

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W., Jr. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology, 103*, 597-600.

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature, 391*, 756.

Bowers, J. L., & Heilman, K. M. (1980). Pseudoneglect - effects of hemispace on a tactile line bisection task. *Neuropsychologia, 18*, 491-498.

Bradshaw, J. L., Nettleton, N. C., Nathan, G., & Wilson, L. (1983). Head and body space to left and right, front and rear--II. Visuotactual and kinesthetic studies and left-side underestimation. *Neuropsychologia, 21*, 475-486.

Brown, L. E., Rosenbaum, D. A., & Sainburg, R. L. (2003). Limb position drift: Implications for control of posture and movement. *Journal of Neurophysiology, 90*, 3105-3118.

Burdea, G. (1996a). *Force and Touch Feedback for Virtual Reality*. New York: John Wiley & Sons.

Burdea, G. (1996b). *Virtual Reality and Robotics in Medicine.* Paper presented at the IEEE International Workshop on Robot and Human Communication (RoMAN'96), Tsukuba, Japan.

Calvert, G. A., Brammer, M.J. & Iversan, S.D. (1998). Crossmodal identification. *Trends in Cognitive Sciences, 2*, 247-252.

Chan, J. S., & Newell, F. N. (in prep.). What/Where pathways in haptics.

Cheng, M. F. (1968). Tactile-kinesthetic perception of length. *American Journal of Psychology, 81*, 74-82.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving: the psychology of visual Illusions*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Craig, J. C., & Rollman, G. B. (1999). Somesthesis. *Annual review of Psychology, 50*, 305-331.

Dassonville, P., Bridgeman, B., Kaur Bala, J., Thiem, P., & Sampanes, A. (2004). The induced Roelofs effect: two visual systems or the shift of a single reference frame? *Vision Research, 44*, 603-611.

Davidson, D. A., & Cheng, M. F. (1964). Apparent distance in a horizontal plane with tactile-kinesthetic stimuli. *Quarterly Journal of Experimental Psychology, 16*, 277-281.

Day, R. H., & Avery, G. C. (1970). Absence of the horizontal-vertical illusion in haptic space. *Journal of Experimental Psychology, 83*, 172-173.

Day, R. H., & Wong, T. S. (1971). Radial and tangential movement directions as determinants of the haptic illusion in an L figure. *Journal of Experimental Psychology, 87*, 19-22.

Desmurget, M., Vindras, P., Grea, H., Viviani, P., & Grafton, S. T. (2000). Proprioception does not quickly drift during visual occlusion. *Experimental Brain Research, 134*, 363-377.

Driver, J., & Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences, 2*, 254-262.

Easton, R. D., Srinivas, K., & Greene, A. J. (1997). Do vision and haptics share common representations? Implicit and explicit memory within and between modalities. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 153-163.

Edelman, S. (1995). Class Similarity and Viewpoint Invariance in the Recognition of 3d Objects. *Biological Cybernetics, 72*, 207-220.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research, 32*, 2385-2400.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*, 429-433.

Ernst, M. O., Banks, M. S., & Bülthoff, H. H. (2000). Touch can change visual slant perception. *Nature Neuroscience, 3*, 69-73.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Science, 8*, 162-169.

Ettlinger, G., & Wilson, W. A. (1990). Cross-modal performance: behavioural processes, phylogenetic considerations and neural mechanisms. *Behavioural Brain Research, 40*, 169-192.

Fagot, J., Hopkins, W. D., & Vauclair, J. (1993). Hand movements and hemispheric specialization in dichhaptic explorations. *Perception, 22*, 847-853.

Fagot, J., Lacreuse, A., & Vauclair, J. (1994). Hand-movement profiles in a tactual-tactual matching task: effects of spatial factors and laterality. *Perception and Psychophysics, 56*, 347-355.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition.* Cambridge, MA: MIT Press.

Farne, A., Pavani, F., Meneghello, F., & Ladavas, E. (2000). Left tactile extinction following visual stimulation of a rubber hand. *Brain, 123 (Pt 11)*, 2350-2360.

Felleman, D. J., & van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1*, 1-47.

Fick, A. (1851). *De errore quodam optico asymmetrica bulbi effecto*. Marburg: J. A. Kockin.

Flanery, R. C., & Balling, J. D. (1979). Developmental changes in hemispheric specialization for tactile spatial ability. *Developmental psychology, 15*, 364-372.

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA, USA: Bradford books, MIT Press.

Forster, K. L., & Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behavior Research Methods Instruments Computers, 35*, 116-124.

Franz, V. H., Gegenfurtner, K. R., Bulthoff, H. H., & Fahle, M. (2000). Grasping visual illusions: no evidence for a dissociation between perception and action. *Psychological Science, 11*, 20-25.

Freides, D. (1974). Human information processing and sensory modality: cross-modal functions, information complexity, memory, and deficit. *Psychological Bulletin, 81*, 284-310.

Gallace, A., & Spence, C. (2004). Examining the crossmodal consequences of viewing the Muller-Lyer illusion. *Experimental Brain Research.*

Garvill, J., & Molander, B. (1973). Effects of standard modality, comparison modality and retention interval on matching of form. *Scandinavian Journal of Psychology, 14*, 203-206.

Gentaz, E., & Hatwell, Y. (2004). Geometrical haptic illusions: the role of exploration in the Muller-Lyer, vertical-horizontal, and Delboeuf illusions. *Psychonomic Bulletin Review, 11*, 31-40.

Gibson, J. J. (1962). Observations on active touch. *Psychological Review, 69*, 477-491.

Gilson, E. Q., & Baddeley, A. D. (1969). Tactile short-term memory. *Quarterly Journal of Experimental Psychology, 21*, 180-184.

Gray, R., & Tan, H. Z. (2002). Dynamic and predictive links between touch and vision. *Experimental Brain Research, 145*, 50-55.

Graziano, M. S. (1999). Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences, 96*, 10418-10421.

Graziano, M. S., Cooke, D. F., & Taylor, C. S. (2000). Coding the location of the arm by sight. *Science, 290*, 1782-1786.

Graziano, M. S. A., & Gross, C. G. (1995). The representation of extrapersonal space: A possible role for bimodal, visual-tactile neurons. In Michael S. Gazzaniga (Ed.), *Cognitive neurosciences* (Vol. xiv, pp. 1021-1034). Cambridge, MA, US: The MIT Press.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research, 41*, 1409-1422.

Hadjikhani, N., & Roland, P. E. (1998). Cross-modal transfer of information between the tactile and the visual representations in the human brain: A positron emission tomographic study. *Journal of Neuroscience, 18*, 1072-1084.

Haggard, P., Newman, C., Blundell, J., & Andrew, H. (2000). The perceived position of the hand in space. *Perception and Psychophysics, 62*, 363-377.

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., et al. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences, 88*, 1621-1625.

Hay, J. C., Smith, C., Pick, H. L., & Ikeda, K. (1965). Visual capture produced by prism spectacles. *Psychonomic Science, 2*, 215-216.

Heller, M. A. (1984). Active and passive touch: the influence of exploration time on form recognition. *Journal of General Psychology, 110*, 243-249.

Heller, M. A., Bracket, D. D., Salik, S. S., Scroggs, E., & Green, S. (2003). Objects, raised lines, and the haptic horizontal-vertical illusion. *Quarterly Journal of Experimental Psychology, 56*, 891-907.

Heller, M. A., Calcaterra, J. A., Burson, L. L., & Green, S. L. (1997). The tactual horizontal-vertical illusion depends on radial motion of the entire arm. *Perception and Psychophysics, 59*, 1297-1311.

Heller, M. A., Calcaterra, J. A., Green, S. L., & Brown, L. (1999). Intersensory conflict between vision and touch: The response modality dominates when precise, attention-riveting judgments are required. *Perception and Psychophysics, 61*, 1384-1398.

Heller, M. A., Joyner, T. D., & Dan-Fodio, H. (1993). Laterality effects in the haptic horizontal-vertical illusion. *Bulletin of the Psychonomic Society, 31*, 440-442.

Hellige, J. B. (1993). *Hemispheric asymmetry: what's right and what's left*. London: Harvard University Press.

Hershberger, W. A., & Misceo, G. F. (1996). Touch dominates haptic estimates of discordant visual-haptic size. *Perception and Psychophysics, 58*, 1124-1132.

Hirsh, I. J., & Sherrick, C. E., Jr. (1961). Perceived order in different sense modalities. *Journal of Experimental Psychology, 62*, 423-432.

Hohmuth, A., Phillips, W. D., & VanRomer, H. (1976). A discrepancy between two modes of haptic length perception. *Journal of Psychology, 92*, 79-87.

Holmes, N. P., Calvert, G. A., & Spence, C. (2004). Extending or projecting peripersonal space with tools? Multisensory interactions highlight only the distal and proximal ends of tools. *Neuroscience Letters, 372*, 62-67.

Howard, I. P. (1966). *Human Spatial Orientation.* New York: Wiley.

Igarashi, Y., Kitagawa, N., & Ichihara, S. (2004). Vision of a pictorial hand modulates visual-tactile interactions. *Cognitive, Affective and Behavioral Neuroscience, 4*, 182-192.

Jack, C. E., & Thurlow, W. R. (1973). Effects of degree of visual association and angle of displacement on the "ventriloquism" effect. *Perceptual and Motor Skills, 37*, 967-979.

Jackson, C. V. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology, 5*, 52-65.

James, T. W., Humphrey, G. K., Gati, J. S., Servos, P., Menon, R. S., & Goodale, M. A. (2002). Haptic study of three-dimensional objects activates extrastriate visual areas. *Neuropsychologia, 40*, 1706-1714.

Jansson, G., & Monaci, L. (2005). *Improving haptic displays: providing differentiated information at the contact areas in more important than increasing the number of areas.* Paper presented at the Worldhaptics, Pisa, Italy.

Jansson, G., & Öström, M. (2004). *The effects of co-location of visual and haptic space on judgments of form.* Paper presented at the Eurohaptics, Munich, Germany.

Johansson, R. S., & LaMotte, R. H. (1983). Tactile detection thresholds for a single asperity on an otherwise smooth surface. *Somatosensory Research, 1*, 21-31.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology, 24*, 175-219.

Kappers, A. M. (2004). The contributions of egocentric and allocentric reference frames in haptic spatial tasks. *Acta Psychology, 117*, 333-340.

Kappers, A. M., & Koenderink, J. J. (2002). Continuum of haptic space. In L. Albertazzi (Ed.), *Unfolding Perceptual Continua* (pp. 163-180). Amsterdam: John Benjamins Publishing company.

Kelvin, R. P. (1954). Discrimination of size by sight and touch. *Quarterly Journal of Experimental Psychology, 6*, 23-34.

Kerzel, D. (2001). Visual short-term memory is influenced by haptic perception. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27*, 1101-1109.

Kiphart, M. J., Hughes, J. L., Simmons, J. P., & Cross, H. (1992). Short-term haptic memory for complex objects. *Bulletin of the Psychonomic Society, 30*, 212-214.

Klatzky, R. L., Lederman, S. J., & Matula, D. E. (1993). Haptic exploration in the presence of vision. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 726-743.

Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: an "expert system". *Perception and Psychophysics, 37*, 299-302.

Klatzky, R. L., Loomis, J. M., Lederman, S. J., Wake, H., & Fujita, N. (1993). Haptic identification of objects and their depictions. *Perception and Psychophysics, 54*, 170-178.

Làdavas, E. (2002). Functional and dynamic properties of visual peripersonal space. *Trends in Cognitive Sciences, 6*, 17-22.

Làdavas, E., di Pellegrino, G., Farnè, A., & Zeloni, G. (1998). Neuropsychological evidence of an integrated visuotactile representation of peripersonal space in humans. *Journal of Cognitive Neuroscience, 10*, 581-589.

Làdavas, E., & Farnè, A. (2004). Visuo-tactile representation of near-the-body space. *Journal of Physiology (Paris), 98*, 161-170.

Làdavas, E., Farnè, A., Zeloni, G., & di Pellegrino, G. (2000). Seeing or not seeing where your hands are. *Experimental Brain Research, 131*, 458-467.

Lakatos, S., & Marks, L. E. (1999). Haptic form perception: relative salience of local and global features. *Perception and Psychophysics, 61*, 895-908.

Lalanne, C., & Lorenceau, J. (2004). Crossmodal integration for perception and action. *Journal of Physiology (Paris), 98*, 265-279.

Lederman, S. J., & Klatzky, R. L. (1987). Hand movements: a window into haptic object recognition. *Cognitive Psychology, 19*, 342-368.

Lederman, S. J., & Klatzky, R. L. (1993). Extracting object properties through haptic exploration. *Acta Psychology (Amsterdam), 84*, 29-40.

Lederman, S. J., & Klatzky, R. L. (2004). Haptic identification of common objects: effects of constraining the manual exploration process. *Perception and Psychophysics, 66*, 618-628.

Lederman, S. J., Klatzky, R. L., Collins, A., & Wardell, J. (1987). Exploring environments by hand or foot: time-based heuristics for encoding distance in

movement space. *Journal of Experimental Psychology: Learning, Memory and Cognition, 13*, 606-614.

Lederman, S. J., Summers, C., & Klatzky, R. L. (1996). Cognitive salience of haptic object properties: role of modality-encoding bias. *Perception, 25*, 983-998.

Lewkowicz, D. J. (1999). The development of temporal and spatial intermodal perception. In G. Aschersleben, T. Bachmann & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 395-420). Amsterdam: Elsevier Science.

Liddle, D., & Foss, B. M. (1963). The Tactile Perception of Size - Some Relationships with Distance and Direction. *Quarterly Journal of Experimental Psychology, 15*, 217-219.

Logie, R. H., & Marchetti, C. (1991). Visuo-spatial working memory: Visual, spatial or central executive? In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (Vol. xvi, pp. 105-115). Oxford, England: North-Holland.

Luh, K. E. (1995). Line Bisection and Perceptual Asymmetries in Normal Individuals - What You See Is Not What You Get. *Neuropsychology, 9*, 435-448.

MacKay, W. A., & Crammond, D. J. (1987). Neuronal correlates in posterior parietal lobe of the expectation of events. *Behavioral Brain Research, 24*, 167-179.

Mahrer, P., & Miles, C. (2002). Recognition memory for tactile sequences. *Memory, 10*, 7-20.

Marchetti, F. M., & Lederman, S. J. (1983). The haptic radial-tangential effect: Two tests of Wong's "moments-of-inertia" hypothesis. *Bulletin of the Psychonomic Society, 21*, 43-46.

Martinez, F. (1971). Comparison of two types of tactile exploration in a task of mirror-image recognition. *Psychonomic Science, 22*, 124-125.

Marzi, C. A., Grabowska, A., Tressoldi, P., & Bisiacchi, P. S. (1988). Left hemisphere superiority for visuospatial functions in left-handers. *Behavioral Brain Research, 30*, 183-192.

Mattingley, J. B., Driver, J., Beschin, N., & Robertson, I. H. (1997). Attentional competition between modalities: extinction between touch and vision after right hemisphere damage. *Neuropsychologia, 35*, 867-880.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*, 746-748.

McKeon, R. (1941). *The basic works of Aristotle*. New York: Random House.

Meredith, M. A. (2002). On the neuronal basis for multisensory convergence: a brief overview. *Brain Research: Cognitive Brain Research, 14*, 31-40.

Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience, 7*, 3215-3229.

Meredith, M. A., & Stein, B. E. (1996). Spatial determinants of multisensory integration in cat superior colliculus neurons. *Journal of Neurophysiology, 75*, 1843-1857.

Mesulam, M. M. (1998). From sensation to cognition. *Brain, 121*, 1013-1052.

Millar, S. (1972). Effects of interpolated tasks on latency and accuracy of intramodal and cross-modal shape recognition by children. *Journal of Experimental Psychology, 96*, 170-175.

Millar, S. (1974). Tactile short-term memory by blind and sighted children. *British Journal of Psychology, 65*, 253-263.

Millar, S. (1975). Effects of tactual and phonological similarity on the recall of Braille letters by blind children. *British Journal of Psychology, 66*, 193-201.

Millar, S. (1981). Crossmodal and intersensory perception and the blind. In R. D. Walk & H. L. Pick (Eds.), *Intersensory perception and sensory integration* (pp. 281-314). New York: Plenum.

Millar, S. (2000). Modality and mind: convergent active processing in interrelated networks as a model of development and perception by touch. In M. A. Heller (Ed.), *Touch, representation and blindness* (pp. 99-142). Oxford: Oxford university press.

Millar, S., & al-Attar, Z. (2000). Vertical and bisection bias in active touch. *Perception, 29*, 481-500.

Millar, S., & Al-Attar, Z. (2004). External and body-centered frames of reference in spatial memory: evidence from touch. *Perception and Psychophysics, 66*, 51-59.

Miller, E. A. (1972). Interaction of vision and touch in conflict and nonconflict form perception tasks. *Journal of Experimental Psychology, 96*, 114-123.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Brain research: Cognitive brain research, 17*, 154-163.

Munhall, K. G., Gribble, P., Sacco, L., & Ward, M. (1996). Temporal constraints on the McGurk effect. *Perception and Psychophysics, 58*, 351-362.

Newell, F. N. (1998). Stimulus context and view dependence in object recognition. *Perception, 27*, 47-68.

Newell, F. N. (2004). Cross-modal object recognition. In G. Calvert, C. Spence & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 123-139). London: The MIT press.

Newell, F. N., Bülthoff, H. H., & Ernst, M. O. (submitted). Multisensory enhancement in the recognition of actively explored objects.

Newell, F. N., Ernst, M. O., Tjan, B. S., & Bülthoff, H. H. (2001). Viewpoint dependence in visual and haptic object recognition. *Psychological Science, 12*, 37-42.

Newell, F. N., Finucane, C., & Pasqualotto, A. (2004). Noninformative vision improves haptic recognition of a scene. *Perception, 33*, 142-142.

Newell, F. N., Woods, A. T., Mernagh, M., & Bülthoff, H. H. (2004). Visual, haptic and crossmodal recognition of scenes. *Experimental Brain Research.*

Newport, R., Rabb, B., & Jackson, S. R. (2002). Noninformative vision improves haptic spatial perception. *Current Biology, 12*, 1661-1664.

Paillard, J. (1991). Knowing where and knowing how to get there. In J. Paillard (Ed.), *Brain and space* (pp. 461-481). London: Oxford University Press.

Pascual-Leone, A., & Hamilton, R. (2001). The metamodal organization of the brain. *Progressive Brain Research, 134*, 427-445.

Patton, P., Belkacem-Boussaid, K., & Anastasio, T. J. (2002). Multimodality in the superior colliculus: an information theoretic analysis. *Brain Research: Cognitive Brain Research, 14*, 10-19.

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

Radeau, M., & Bertelson, P. (1977). Adaptation to Auditory-Visual Discordance and Ventriloquism in Semi-Realistic Situations. *Perception & Psychophysics, 22*, 137-146.

Ramachandran, V. S., Rogers-Ramachandran, D., & Stewart, M. (1992). Perceptual correlates of massive cortical reorganization. *Science, 258*, 1159-1160.

Reales, J. M., & Ballesteros, S. (1999). Implicit and explicit memory for visual and haptic objects: Cross-modal priming depends on structural descriptions. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*, 644-663.

Reed, C. L., Caselli, R. J., & Farah, M. J. (1996). Tactile agnosia. Underlying impairment and implications for normal tactile object recognition. *Brain, 119*, 875-888.

Reed, C. L., Shoham, S., & Halgren, E. (2004). Neural substrates of tactile object recognition: an fMRI study. *Human Brain Mapping, 21*, 236-246.

Reed, E. S. (1988). *James J. Gibson and the psychology of perception*. New Haven: Yale University Press.

Reid, R. L. (1954). An illusion of movement complimentary to the horizontal-vertical illusion. *Quarterly Journal of Experimental Psychology, 6*, 107-111.

Rock, I., & Victor, J. (1964). Vision and Touch: An Experimentally Created Conflict between the Two Senses. *Science, 143*, 594-596.

Rossetti, Y., Desmurget, M., & Prablanc, C. (1995). Vectorial coding of movement: vision, proprioception, or both? *Journal of Neurophysiology, 74*, 457-463.

Sampaio, E., & Chokron, S. (1992). Pseudoneglect and reversed pseudoneglect among left-handers and right-handers. *Neuropsychologia, 30*, 797-805.

Sampaio, E., & Philip, J. (1991). Sensory and motor aspects of pseudoneglect, hemifield, and hemispace in the tactile modality. *Brain and Cognition, 16*, 74-82.

Schultz, L. M., & Petersik, J. T. (1994). Visual-haptic relations in a two-dimensional size-matching task. *Perceptual and Motor Skills, 78*, 395-402.

Seizova-Cajic, T. (1998). Size perception by vision and kinesthesia. *Perception and Psychophysics, 60*, 705-718.

Soto-Faraco, S., & Kingstone, A. (2004). Multisensory integration of dynamic information. In G. Calvert, C. Spence & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 49-67). London: The MIT press.

Soto-Faraco, S., Lyons, J., Gazzaniga, M., Spence, C., & Kingstone, A. (2002). The ventriloquist in motion: illusory capture of dynamic information across sensory modalities. *Brain Research: Cognitive Brain Research, 14*, 139-146.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General, 130*, 799-832.

Spence, C., & Squire, S. (2003). Multisensory integration: maintaining the perception of synchrony. *Current Biology, 13*, R519-521.

Stark, J. W., Anes, M. D., & Liederman, J. (2004). Spatiotemporal continuity implicated in the binding of features and object identity. *International Journal of Neuroscience, 114*, 751-775.

Stein, B. E., Jiang, W., & Stanford, T. R. (2004). Multisensory integration in single neurons in the midbrain. In G. Calvert, C. Spence & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 243-264). London: The MIT press.

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. London: The MIT Press.

Stein, B. E., & Wallace, M. T. (1996). Comparisons of cross-modality integration in midbrain and cortex. *Progressive Brain Research, 112*, 289-299.

Sullivan, E. V., & Turvey, M. T. (1972). Short-term retention of tactile stimulation. *Quarterly Journal of Experimental Psychology, 24*, 253-261.

Summers, D. C., & Lederman, S. J. (1990). Perceptual asymmetries in the somatosensory system: a dichhaptic experiment and critical review of the literature from 1929 to 1986. *Cortex, 26*, 201-226.

Suzuki, K., & Arashida, R. (1992). Geometrical haptic illusions revisited: haptic illusions compared with visual illusions. *Perception and Psychophysics, 52*, 329-335.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *Journal of Experimental Psychology: Human Perception and Performance, 21*, 1494-1505.

Teghtsoonian, R., & Teghtsoonian, M. (1970). Two varieties of perceived length. *Perception & Psychophysics, 8*, 389-392.

Thurlow, W. R., & Jack, C. E. (1973). Certain determinants of the "ventriloquism effect". *Perceptual and Motor Skills, 36*, 1171-1184.

Tillery, S. I. H., Flanders, M., & Soechting, J. F. (1994). Errors in Kinesthetic Transformations for Hand Apposition. *Neuroreport, 6*, 177-181.

van Beers, R. J., Sittig, A. C., & Denier van der Gon, J. J. (1998). The precision of proprioceptive position sense. *Experimental Brain Research, 122*, 367-377.

van Beers, R. J., Sittig, A. C., & Gon, J. J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology, 81*, 1355-1364.

Vindras, P., Desmurget, M., Prablanc, C., & Viviani, P. (1998). Pointing errors reflect biases in the perception of the initial hand position. *Journal of Neurophysiology, 79*, 3290-3294.

Volcic, R., Kappers, A. M., & Koenderink, J. J. (2005). *Bimanual parallelity task on the frontoparallel plane.* Paper presented at the Worldhaptics, Pisa, Italy.

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., & Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Experimental Brain Research, 158*, 252-258.

Warren, D. H. (1979). Spatial Localization under Conflict Conditions - Is There a Single Explanation. *Perception, 8*, 323-337.

Welch, R. B. (1999). Meaning, attention, and the "Unity Assumption" in the Intersensory bias of spatial and temporal perceptions. In G. Aschersleben, T. Bachmann & J. Müsseler (Eds.), *Cognitive contributions to the perception of spatial and temporal events* (pp. 371-387). Amsterdam: Elsevier Science.

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin, 88*, 638-667.

Wuillemin, D., van Doorn, G., Richardson, B., & Symmons, M. (2005). *Haptic and visual size judgements in virtual and real environments.* Paper presented at the Worldhaptics, Pisa, Italy.

Zampini, M., Brown, T., Shore, D. I., Maravita, A., Roder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychology (Amsterdam), 118*, 277-291.

Zampini, M., Shore, D. I., & Spence, C. (2003). Audiovisual temporal order judgments. *Experimental Brain Research, 152*, 198-210.

Zangaladze, A., Epstein, C. M., Grafton, S. T., & Sathian, K. (1999). Involvement of visual cortex in tactile discrimination of orientation. *Nature, 401*, 587-590.

Zhang, M., Weisser, V. D., Stilla, R., Prather, S. C., & Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cognitive, Affective and Behavioral Neuroscience, 4*, 251-259.

Zhou, Y. D., & Fuster, J. M. (1997). Neuronal activity of somatosensory cortex in a cross-modal (visuo-haptic) memory task. *Experimental Brain Research, 116*, 551-555.