8-1-2004

The Effect of Experience Upon the Visual and Haptic Discrimination of 3-D Object Shape

Anna Clayton
Western Kentucky University

Recommended Citation

Clayton, Anna, "The Effect of Experience Upon the Visual and Haptic Discrimination of 3-D Object Shape" (2004). *Masters Theses & Specialist Projects.* Paper 544.
http://digitalcommons.wku.edu/theses/544

THE EFFECT OF EXPERIENCE UPON THE VISUAL AND HAPTIC DISCRIMINATION OF 3-D OBJECT SHAPE

A Thesis

Presented to

The Faculty of the Department of Psychology

Western Kentucky University

Bowling Green, Kentucky

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Anna Marie Clayton

August 2004

THE EFFECT OF EXPERIENCE UPON THE VISUAL AND HAPTIC DISCRIMINATION OF 3-D OBJECT SHAPE

Date Recommended 6/29/2004

J. Farley Norman
Director of Thesis

Elmer Gray 8/11/04
Dean, Graduate Studies and Research Date

Acknowledgments

First and foremost, I would like to thank my graduate advisor, Dr. J. Farley Norman, for his guidance and support throughout my years of study at Western Kentucky University. He has taught me many skills to aid me in my future in academics. I would also like to thank Dr. Hideko Norman for her patience, understanding, and help in completing my thesis, as well as Chuck Crabtree for the many hours he contributed in collecting data. I could not have done it without the help of these three people. I would like to thank the other members of my committee, Dr. Joseph Bilotta for his support and suggestions, and Dr. Dan Roenker for his patience and help with tweaking my statistics and SPSS skills. I wish to thank my parents for all of their love and encouragement. And last but not least, I want to thank all of my fellow graduate students, especially Josh and Allison, for all of the laughs we have shared. Thank you for not letting me take any of this too seriously.

Table of Contents

Page

List of Figures

THE EFFECT OF EXPERIENCE UPON THE VISUAL AND HAPTIC DISCRIMINATION OF 3-D OBJECT SHAPE

Anna Marie Clayton August 2004 32 Pages

Directed by: J. Farley Norman, Ph.D., Joseph Bilotta, Ph.D., and Dan Roenker, Ph.D.

Department of Psychology Western Kentucky University

Both our sense of touch and our sense of vision allow us to perceive common object properties such as size, shape, and texture. The extent of this functional overlap has been studied in relation to infant perception (Bushnell & Weinberger, 1987; Gibson & Walker, 1984; Streri, 1987; Streri & Gentaz, 2003), overlap in brain regions (Amedi, Malach, Hendler, Peled, & Zohary, 2001; Deibert, Kraut, Kermen, & Hart, 1999; James, Humphrey, Gati, Menon, & Goodale, 2002), and adult perception (Gibson, 1962, 1963, 1966; Klatzky, Lederman, & Reed, 1987; Lakatos & Marks, 1999; Norman, Norman, Clayton, Lianekhammy, & Zielke, 2004). The current experiment extended the findings of Norman et al. (2004) by examining the effect of experience upon the visual and haptic discrimination of 3-D object shape, as well as examining for differences in how long visual and haptic shape representations can be held in short-term memory. Participants were asked to compare the shapes of two objects either within a single sensory modality (both objects presented visually or haptically) or across the sensory modalities (one object presented visually, the other presented haptically) for 120 trials. Their task was to compare whether the objects possessed the "same" or "different" 3-D shapes. The objects were presented for a duration of 3 seconds each, with a 3-, 9-, or 15-second inter-stimulus interval (ISI) between them. Both the unimodal (visual-visual and haptic-haptic) and cross-modal (visual-haptic and haptic-visual) conditions exhibited a linear pattern of learning, and were unaffected by the various ISI's used. However, different levels of discrimination accuracies were observed for the various groups with the highest level of accuracy occurring for the visual-visual group (*M* = 78.65 % correct) and the lowest level of accuracy occurring for the haptic-visual group (*M* = 65.31 % correct). Different patterns of errors for "same" versus "different" trials were observed for the unimodal and cross-modal conditions. Taken together, the results of the current experiment give us a better understanding of the similarities and differences that exist between the visual and haptic sensory modalities representations of 3-D object shape.

Chapter 1

Introduction

Imagine for a moment that you are a child coming home after a long day at school. You walk into your house to an enticing aroma. It is warm and sweet, and makes your mouth immediately start watering. Your nose lures you into the kitchen and over to the oven. You open the oven door to reveal a pan of light tan treats with dark, gooey chips scattered about. They are round with a diameter of approximately 2 inches. They are relatively thin, with a thickness that appears to be ¼ to ½ inch. You can hear a slight sizzling noise, and the dark, gooey chips are starting to bubble. You give in to temptation and reach in to grab one of these mysterious treats. But, ouch! They are hot, and pain shoots through your fingers. A mixture of crumbs and the warm, gooey substance is left on your fingers. You lick it off, and to your delight the taste is sweet, and satisfying, with just the right hint of chocolate. You are in heaven.

As you have probably figured out, the treats described in this scenario are not mysterious at all, but are chocolate chip cookies. The purpose of this description was to demonstrate how our senses work together to allow us to perceive and make sense of the world. All our senses have very specific functions and perceive distinct properties about the environment. For the most part they do not overlap in their functions. For example, we cannot taste a sound, or smell a flavor (although there are exceptions, see Cytowic, 1993). However, we do see overlap in the environmental properties we perceive from our senses of touch and vision. The same information about an object's size, shape, and texture can be perceived by both vision and touch. This overlap is referred to as amodal perception, and can only occur if the modalities are analogous in their functioning, or equivalent in the information transferred "both from vision to touch and from touch to vision" (Streri, 1993, p. 10). While these senses do have independent functions (i.e., visual color perception, haptic temperature perception), the extent of their overlap has intrigued philosophers and scientists for centuries (Berkeley, 1963/1709; Lucretius Carus, 1950/~58 B.C.). Common questions asked include: Are these senses equivalent in ability? Are there conditions in which one is better than the other? What is the extent of their overlap? Is the information perceived interchangeable? How easily can information be transferred and compared across the senses?

Much of the research that has been conducted on the relationship between the perception of object shape through touch and vision has been with infants. In order to operationally study the intermodal transfer of shape information across the visual and haptic systems in infants, there must be two phases to the situation (Gottfried, Rose, & Bridger, 1977, as cited in Streri, 1993). First the infants must be haptically familiarized with an object, and second, they must be visually presented with two objects, one familiar and the other novel. If a difference is shown in the amount of time the infant looks at each object, it is interpreted to mean the transfer of information from the haptic to the visual modality has taken place. Though there are some discrepancies in the literature, it is generally accepted that a preference for the novel object (longer looking time) is evidence that the baby visually recognized the haptically familiarized object.

In a study conducted by E. J. Gibson and Walker (1984) it was demonstrated that babies as young as 1-month old could recognize a substance visually with which they had been oral haptically familiarized. The infants were given either a hard or sponge-like substance to mouth for 60 seconds. They then were presented with both the familiar and the novel substances rotating in depth on vertical axes at approximately the same speed. The mean proportion of looking time for the novel substance (.605) was significantly greater than for the familiar substance (.395), $p = .0066$. Gibson and Walker (1984) concluded that the substance information was being transferred across the visual and haptical modalities, but they considered the movement of the visual stimuli to be an important factor in this transfer being able to occur.

A study conducted by Streri (1987) gave further evidence for the transfer of information from touch to vision. In this study, 2-month-olds were familiarized tactually with a circular object that either had a hole in it (doughnut shaped), or did not. As with Gibson and Walker's (1984) findings, Streri found that shape information could be transferred from touch to vision. However, the transfer was not demonstrated when the objects were familiarized visually and then to be discriminated between haptically. In a 2003 study, Streri and Gentaz demonstrated that the ability to transfer simple shape information from touch to vision was present even in newborns (*M* age = 62 hours), indicating that the ability may be inherent.

In a study conducted by Bushnell and Weinberger (1987) a sort of "trick" method was used to assess the ability of 11-month-old infants to transfer object shape and texture information across the visual and haptic modalities. The infants were introduced visually to the reflection of a 3-D object through a "reaching box device" (p. 602; see also Bushnell, 1982). They then felt an object in the location where the reflected object appeared to be. They could not see the actual object they were feeling, though they could view the reflected object and feel the actual object simultaneously. In discrepancy trials the object felt was different in shape, texture, or both from the object seen. In control trials, the reflected and actual objects were identical. The infants' facial reactions and hand movements were recorded as they came in contact with the actual objects. A panel of judges then reviewed the tapes and were forced to decide whether the infant was perceiving a discrepant pair of objects or an identical pair. In other words, did what the infant was feeling match what they were seeing? "The panel of judges together was considered to have scored a 'hit' (i.e., to have been correct) for a trial if either three or four of them individually were correct, and the panel was considered to have scored a 'miss' if none, one, or two of them were correct" (p. 603). The frequency of hits and misses was then compared to expected frequencies using a one-sample χ^2 test. When the objects were different in both shape and texture the infants appeared to notice, by searching for the object they were viewing and by appearing puzzled. They were also able to detect the discrepancy when the objects differed drastically in shape only (i.e., seeing an egg shaped object, but feeling a square one). The panel of judges accurately discriminated between the discrepancy and control trials for these pairs. However, the discrepancy did not appear to be detected by the infants when the object viewed was simpler in shape or texture than the object felt. The judges could not accurately differentiate between the discrepancy and control trials for these pairs. A second study revealed that when the more complex object was viewed (i.e., the one with texture, or more edges), and the simpler one was felt, the discrepancy was detected by the infants. The authors concluded that the "pattern of results is consistent with the idea that visual information defines the parameters for manual exploration to focus on; it sets the agenda for the hands" (p. 607).

There has also been neurophysiological evidence that the haptic and visual modalities share common functions. Brain regions typically thought to be involved solely in visual processing have been shown to be equally active during haptic processes. Specifically, using fMRI brain scans one study demonstrated that Area MO (middle occipital) in the extrastriate visual cortex was equally activated during the haptic and visual exploration of 3-D objects (James, Humphrey, Gati, Menon, & Goodale, 2002). The same study also found that both Areas MO and LO (lateral occipital) in the extrastriate visual cortex became as activated during a haptic-to-visual priming task as during a visual-to-visual priming task. A separate study was able to demonstrate that the area termed the lateral occipital complex (LOC) within the occipito-temporal region, typically known to respond to visual object shape, is equally activated by object shape presented haptically (Amedi, Malach, Hendler, Peled, & Zohary, 2001). These findings are congruent, as Area MO is located in the lateral occipital complex. Similarly, Deibert, Kraut, Kermen, and Hart (1999) found activation in the calcarine and extrastriatal cortices of the occipital lobe during both grasping and index finger/thumb pinching haptic exploration. However, it should be noted that none of these studies examined the activation of haptic areas in the cortex during visual object exploration.

The implications of the infant and neurophysiological research are limited in that they only provide evidence that the transfer of 3-D shape information across the visual and haptic systems is possible and that it occurs in specific regions of the brain. They do not provide any information about the extent to which this transfer is possible nor about the accuracy with which it occurs.

When comparing the visual and haptic modalities, an obvious question is whether or not they have equivalent short-term memory capacities. While a large amount of research has been conducted on visual short-term memory (STM) span (for a review see Hurt & Ellis, 2004, chap. 4), much less has examined tactile STM or compared the visual and haptic modalities directly on any one perceptual dimension (e.g., 3-D shape). In a study conducted by Bliss, Crane, Mansfield, and Townsend (1966), tactile memory span was examined by stimulating various points on the fingers of both hands (24 total points) simultaneously and having the participants report the number of points they felt. Stimulation occurred by way of air pressure pulses, and only the fingers of both hands were tested (thumbs excluded). When corrected for guessing, it was found that participants could report an average of 3.5 to 7.5 (out of 12) stimulated positions, a span similar to what has been observed for visual studies.

In 1971, Goodnow conducted a study to see if cross-modal matching from vision-to-touch was differential affected by number of comparison objects than matching from touch-to-vision. Participants were presented with a standard 3-D object for 4 seconds and then presented with 1, 3, or 5 comparison objects one at a time. Overall, participants' performance, when required to transfer the object shape information from touch-to-vision, was more severely affected by increasing the number of comparison objects than the requirement to transfer from vision-to-touch. From this Goodnow (1971) concluded that "memory for information gathered by hand appears to be less stable than for information gathered by eye…" (p. 89).

According to leading researchers in haptic perception Klatzky, Lederman, and Reed, there are two theoretical frameworks for haptic perception (Katzky & Lederman, 2002; Klatzky, Lederman, & Reed, 1987; see also Lakatos & Marks, 1999). What they refer to as the "image-mediated model" treats "haptics as an inferior form of vision" (Klatzky, Lederman, & Reed, 1987, p. 357). According to this model, haptic information of object contour is translated into a visual image and is then "'reperceived' (cf. Kerst & Howard, 1978) by visual processors" (p. 357). In contrast, the second framework "assumes that the haptic system has its own encoding processes and pathways, which may or may not be shared with vision" (p. 357). Even when an object is perceived simultaneously by vision and touch, this model suggests that the modalities are likely to give different weights to specific features. What is perceptually important to the haptic system may differ based on the presence or absence of visual information. Specifically this model is "concerned with the 'salience' of object attributes under haptic exploration" (p. 357).

To test this second model, Klazky, Lederman, and Reed (1987) conducted a series of experiments in which participants had to haptically sort a series of objects based on similarity. There were a total of 81 objects varying across three levels for the dimensions of shape, size, hardness, and texture. In one study, participants were divided into "four instructional conditions: unbiased haptics, haptically biased haptics, visual-imagery biased haptics, and haptics with vision" (p. 361). In the unbiased haptic condition, participants were blindfolded and told to sort the objects based on similarity. In the haptically biased haptic condition, participants were blindfolded and told that objects go together that feel similar. In the visual-imagery biased haptic condition, participants were blindfolded and told that objects should be grouped together if their visual images are similar. In the haptics with vision group, participants were not blindfolded and, like the unbiased haptic group, were not given specific sorting instructions. The results showed that participants in both the haptically biased and unbiased conditions sorted primarily based on object hardness. When instructed to sort based on visual image, however, participants relied almost exclusively on object shape. For the group in which vision and haptics were both used, all dimensions were used as a basis for sorting, though shape was the most emphasized. Texture was used in all groups, though less so in the with-vision and visually imagery groups. Size was used only in the with-vision condition, and even then the usage was minimal. These findings support the authors' theory that when an object is naturally encoded by haptics alone, the features that are salient differ substantially from when it is encoded by vision or visual imagery.

Further support was found for this model of haptic perception, the "direct apprehension model," in a study that examined the weighting of local versus global features of objects in a visual or haptic sorting task (Lakatos & Marks, 1999, p. 907). It was found that in the haptic conditions participants weighted the local features of objects more heavily than in the visual conditions. It was also found that "differences in local features tended to exert their greatest influence during early exploration, and progressively less so given longer time periods" (p. 907). From this finding, the authors concluded that when an object is perceived haptically, a global image of the object is not developed until after the local features have been perceived.

Few studies have compared the performance of congenitally blind persons to that of visually normal persons on visual or haptic tasks. However, a study conducted by Aleman, van Lee, Mantione, Verkoijen, and de Haan (2001) compared the performance of these two groups on a pictorial imagery task and a spatial imagery task. In the pictorial imagery task subjects were asked to compare the outline shape of three objects and indicate which one was the most deviant. For example, in one triad subjects were asked to indicate the "odd-one-out" of a saw, a hammer, and an axe (p. 2602). In the spatial imagery task, subjects were orally given directions of a pathway to follow with their finger on a 2-D or 3-D wood matrix. Some subjects were also required to simultaneously do a finger tapping distracter task while conducting the imagery task. The authors found that both the blind and normal sighted groups were able to perform these tasks at a level well above chance. The finger tapping distracter task also equally negatively affected them. However, for both the pictorial imagery and spatial imagery tasks the normal sighted group performed significantly better than the congenitally blind group. "This suggests that visual experience contributes to visual imagery performance" (p. 2603). This study did not examine performance for the haptic exploration of unknown objects nor the ability of the two groups to transfer shape information across the visual and haptic sensory modalities. More research is needed for conclusive evidence that differences in object shape discrimination exist between blind and sighted persons.

During the 1960's, James Gibson published a series of papers discussing his beliefs about the relationship between the visual and haptic systems (1962, 1963, 1966). Gibson believed that the haptic and visual perceptions of 3-D object shape were essentially equivalent, and that matching objects across the two modalities could be done easily with little or no practice (1962). Gibson based these conclusions on a series of experiments he conducted, but never published. While he did briefly describe the procedure for these experiments in his 1963 paper, he did not provide any data.

As Gibson made important observations regarding the equivalence of the visual and haptic perceptions of 3-D object shape, Norman, Norman, Clayton, Lianekhammy, and Zielke (2004) decided to replicate his studies using natural 3-D objects. To do this further replication, they obtained James Caviness's (Gibson's student at the time) Master's Thesis (1962, as cited in Norman et al., 2004) and Dissertation (1964, as cited in Norman et al., 2004), which referred to and replicated Gibson's original studies.

Norman et al. (2004) replicated two of Gibson's original experiments but used natural 3-D objects as stimuli in place of the manmade sculptures ("feelies") that Gibson used. The natural stimuli were 2 plastic copies of 12 bell peppers (*Capsicum annuum*). In the first experiment observers participated in a cross modal matching task. They were presented with one of the twelve objects behind a curtain to haptically explore, while simultaneously being able to view replicas of objects 1-12 on a table. Their task was to match the object they were feeling with its replica on the table based on its 3-D shape. Observers were allotted 3, 5, 7, 9, or 15 seconds to actively touch the object. A total of 120 trials were conducted (10 per object), broken into 5 sessions. The authors found that observers were able to recognize all twelve objects with a fairly high level of accuracy. A significant, though modest, effect of time given to haptically explore the objects was obtained, with the highest level of performance occurring at 7 seconds. There was also a small effect of experience with the objects, with the largest improvement occurring between sessions 1 and 2.

The second experiment of Norman et al. (2004) consisted of a same/different task. In this experiment, observers were presented with one object to explore (visually or haptically). After 3 seconds it was taken away and observers were presented with a second object (visually or haptically). Their task was to state whether the second object possessed the "same" or "different" shape as the first object. Observers were randomly assigned to one of four conditions based on how the objects were presented and in what order. The conditions were visual-visual (VV), haptic-haptic (HH), visual-haptic (VH), and haptic-visual (HV). A total of 120 trials were conducted, 50 percent in which the objects were the same in shape, and 50 percent in which they were different in shape. Significant differences in performance occurred between the vision-vision condition and both cross-modal conditions (vision-haptic and haptic-vision). There was also a significant difference in the percentage of errors made on "same" trials versus "different" trials, which was dependent on the modality condition (VV, HH, VH, HV). Based on these results Norman et al. concluded that Gibson (1962) was correct in that important similarities do exist between the visual and haptic systems. While cross-modal comparisons can be done at a fairly high level of accuracy, errors still occur, suggesting that the modalities may not necessarily share a single common representation of 3-D shape.

The purpose of the current experiment was to extend the findings of Norman et al. (2004), as well as to fill a void in the existing literature on the equivalence (or nonequivalence) of visual and haptic perception. It was designed to examine the effect of experience upon the visual and haptic discrimination of 3-D shape, and to see if differences exist in how long visual and haptic shape representations can be held in short-term memory. A further examination of the error differences ("same" vs. "different" trials) that occurred in Norman et al. when shape information was transferred across the modalities was also conducted.

Chapter 2

Method

Observers

The observers were 96 undergraduate and graduate students at Western Kentucky University (8 for each of the 12 between-subjects conditions). All observers received $5.00 and/or class extra credit in exchange for their participation.

Stimulus Displays

The stimuli used in the experiment were selected from the twelve stimuli used in the cross-modal matching task conducted in Experiment 1 of Norman et al. (2004) on the basis that they were 'confusable' with at least one of the other stimuli. All stimuli were selected in pairs in which the confusion was reciprocal. A pair was labeled 'confusable' only if each was incorrectly identified as the other in at least 10 percent of the trials (10 out of 100 trials per object). For example, objects 1 and 3 were selected as a pair because object 1 was misidentified as object 3 30 percent of the time, and object 3 was misidentified as object 1 15 percent of the time. The pairs were selected based on the data from the condition in Experiment 1 in which the participants were allowed 3 seconds to haptically explore the object. The confusion matrix for this condition is shown in Table 1. The selected pairs are underlined. A total of six "different" pairs were selected, consisting of eight individual objects. The pairs used were (1,3), (1,7), (2,11), (3,7), (3,8), and (5,12). Photographs of the pairs are shown in Figure 1. The eight individual objects were also paired with their replicas to make eight "same" pairs. In total, 14 distinct pairs (8 same, 6 different) were used as stimuli in the experiment. The actual stimuli consisted of two copies of eight bell peppers (*Capsicum annuum*) made from C-

Stimulus Object

Response Object:	1	2	3	4	5	6	7	8	9	10	11	12
1	33	0	15	4	2	0	10	6	0	1	0	4
2	6	60	7	3	0	5	17	8	6	1	13	0
3	30	5	46	20	0	0	23	26	0	7	0	1
4	3	0	2	28	2	0	1	4	0	1	0	7
5	0	0	0	3	70	0	1	0	0	7	0	12
6	0	12	0	1	0	88	1	0	8	0	11	0
7	15	2	12	12	4	0	31	7	0	1	0	0
8	11	0	17	23	2	0	12	42	1	1	0	9
9	1	7	0	1	0	3	1	5	59	4	1	0
10	1	0	1	2	8	0	1	1	2	77	0	16
11	0	14	0	1	0	4	2	1	24	0	75	1
12	0	0	0	2	12	0	0	0	0	0	0	50

Table 1. Confusion matrix from Norman et al. (2004)—Experiment 1; 3-second condition. Objects selected for current experiment are underlined.

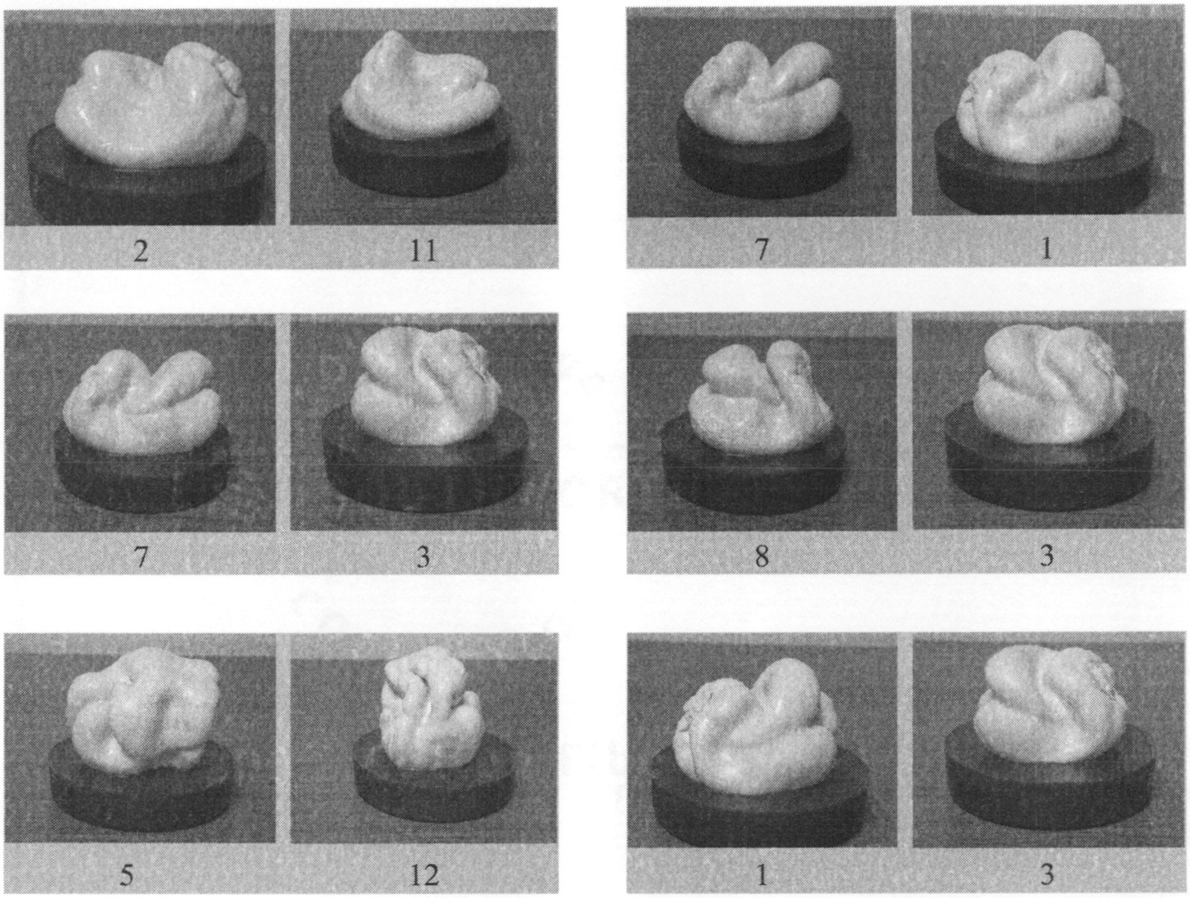

Figure 1. Objects used in current experiment. Shown here are the six “different” pairs. The eight “same” pairs were the objects paired with their exact replica.

1506 rigid urethane casting compound (Smooth-on, Inc.).

Procedure

On any given trial, participants were asked to compare the shapes of two objects presented to them sequentially. Participants were assigned to one of four groups in which both objects were presented either within a single sensory modality (both objects compared visually or haptically) or across the sensory modalities (one object presented visually, the other haptically). There were two cross-modal groups, one in which the first object was presented visually and the second was presented haptically, and another in which the first object was presented haptically and the second visually. The participants were asked to indicate whether the objects possessed the "same" or "different" 3-D shape. The objects were presented for a duration of 3 seconds each, with a 3-, 9-, or 15-second inter-stimulus interval between them. Thus there were four combinations of modalities (vision-vision, haptic-haptic, vision-haptic, haptic-vision) and three durations of the inter-stimulus interval (3, 9, & 15 seconds) separating the presentation of the two objects on any given trial, thus creating a total of 12 between-subjects conditions.

The objects presented in the visual conditions were placed on a turntable that rotated at a constant speed of 30 rpm. This rotation speed allowed for 1.5 complete revolutions in the 3 seconds allotted for viewing the object. The objects were viewed at the observers' eye height. In the haptic conditions, participants reached under a table and behind a black curtain to touch the object. They were allowed to use both hands to pick up the object. Timing did not begin until the participant had touched the object. In all conditions (vision and/or haptic) the object was presented to the participant in a randomly chosen orientation.

Each participant judged 120 trials in a single session lasting approximately 1.5 hours. In 50 percent of the trials the object pairs presented were the same in shape (exact replicas), and in 50 percent they were different in shape (see Figure 1). The session was divided into 5 subsessions of 24 trials each. Each of the 6 "different" pairs was observed twice per subsession. The 12 "same" trials in each subsession were determined by random selection from the 8 "same" object pairs (replicas). The order of the presentation of object pairs (same vs. different) was randomly determined. Short breaks were allowed, but not required between subsessions.

Chapter 3

Results

The participants' overall discrimination accuracies were examined and plotted in terms of percent correct and/or d'. The use of d', derived from Signal Detection Theory, was to control for possible response biases within the individual observers (Macmillan & Creelman, 1991). Two 3-way mixed Analyses of Variance (ANOVA's) were conducted, one using percent correct as the dependent variable and one using d' as the dependent variable, to examine the effects of two between-subjects factors [4 (modality) x 3 (inter-stimulus interval)], and one within-subjects factor [5 (learning/subsession)]. In both the percent correct and d' ANOVAs a significant main effect was found for learning with F $(4, 336) = 14.05, p < .001$ and $F\ (4, 336) = 11.46, p < .001$, respectively. Tests of within-subject contrasts revealed the learning trend to be linear in both cases (percent correct, F $(1, 84) = 43.38, p < .001$; d', $F\ (1, 84) = 36.35, p < .001$), indicating an overall gradual improvement in discrimination accuracy from subsession 1 to subsession 5. This overall learning trend is shown in Figure 2. Though the interaction of modality by learning was not found to be significant, the various patterns of learning for each modality group (HH, VV, HV, VH) are plotted in Figure 3. An overall significant main effect was also found for modality (percent correct, $F\ (3, 84) = 12.54, p < .001$; d', $F\ (3, 84) = 10.47, p < .001$), indicating that the ability to accurately discriminate between the objects was affected by whether they were presented unimodally (HH or VV) or cross-modally (HV or VH). In terms of percent correct, the lowest discrimination accuracy ($M = 65.31$) occurred in the haptic-vision condition, while the highest discrimination accuracy ($M = 78.65$) occurred in the vision-vision condition. The results of a Fisher LSD post-hoc analysis revealed

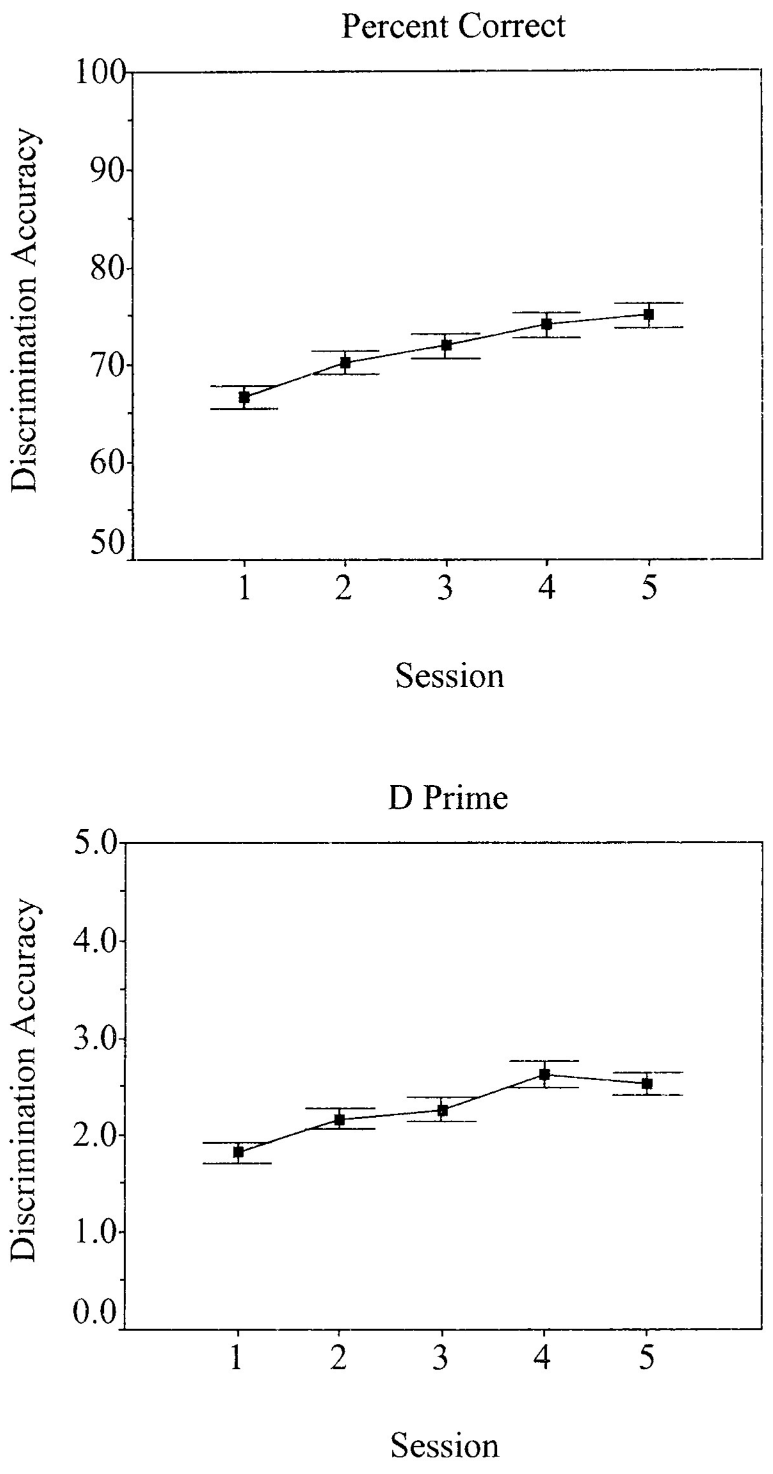

Figure 2. Change in discrimination performance over subsessions plotted both in terms of percent correct (top figure) and d' (bottom figure). Error bars mark +/- one standard error of the mean.

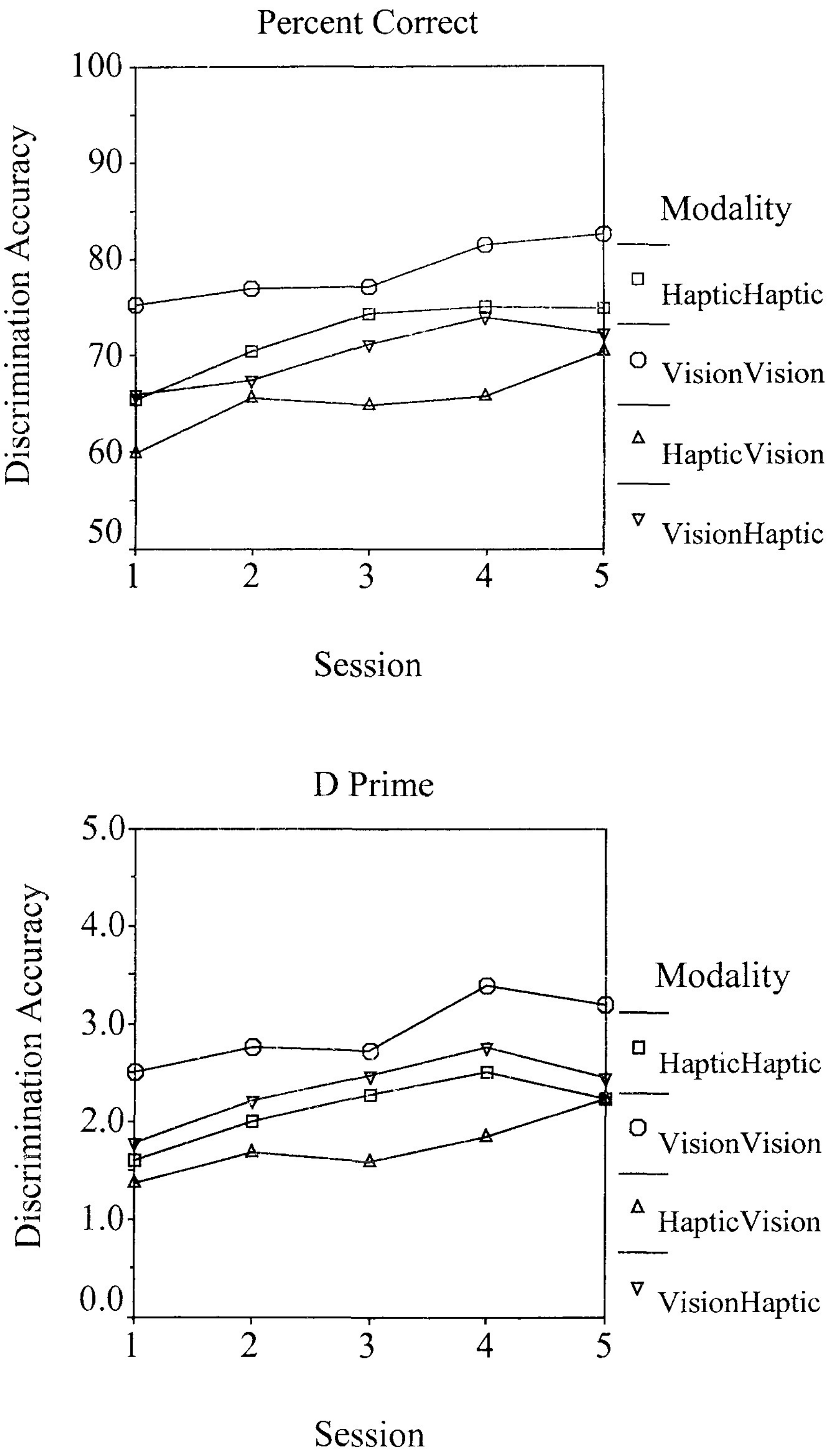

Figure 3. Change in discrimination performance over subsessions for each modality condition. Top graph is plotted in terms of percent correct, bottom graph is in terms of d'.

that discrimination accuracies for the vision-vision, haptic-haptic, and vision-haptic conditions were all significantly higher than for the haptic-vision condition ($p < .001$, $p = .003$, $p = .031$, respectively). The performance in the vision-vision condition was also significantly better than in the haptic-haptic ($p = .003$) or the vision-haptic ($p < .001$) conditions. When analyzed in terms of d', the Fisher post-hoc test revealed similar results. Once again, the performance for the vision-vision condition was found to be significantly better than in any of the other conditions ($p < .001$). However the only other difference found was between the two cross-modal conditions ($p = .007$). For d' the highest discrimination accuracy ($M = 2.9$) was again observed in the vision-vision condition, while the lowest ($M = 1.74$) was in the haptic-vision condition. The observers overall discrimination performances for the four modality conditions (HH, VV, HV, and VH) are shown in Figure 4, both in terms of d' and percent correct. No main effect for inter-stimulus interval or any interactions between the factors were observed.

To examine for possible performance differences on "same" and "different" trials a 3-way mixed ANOVA was conducted with two between-subjects factors (modality and inter-stimulus interval) and one within-subjects factor (trial type: "same" vs. "different") trials. It was only possible to examine for these differences in terms of percent correct, as the essence of d' is to account for biases to respond "same" or "different." Two main effects, trial type and modality, were found to be significant. The modality effect was also observed in the previous ANOVAs (see Figure 4). For trial type, performance on "same" trials ($M = 77.34$) was significantly higher than performance on "different" trials ($M = 65.63$), $F(1, 84) = 28.99$, $p < .001$. In other words, participants were significantly better at identifying the pairs of objects as being the same in shape than being different in

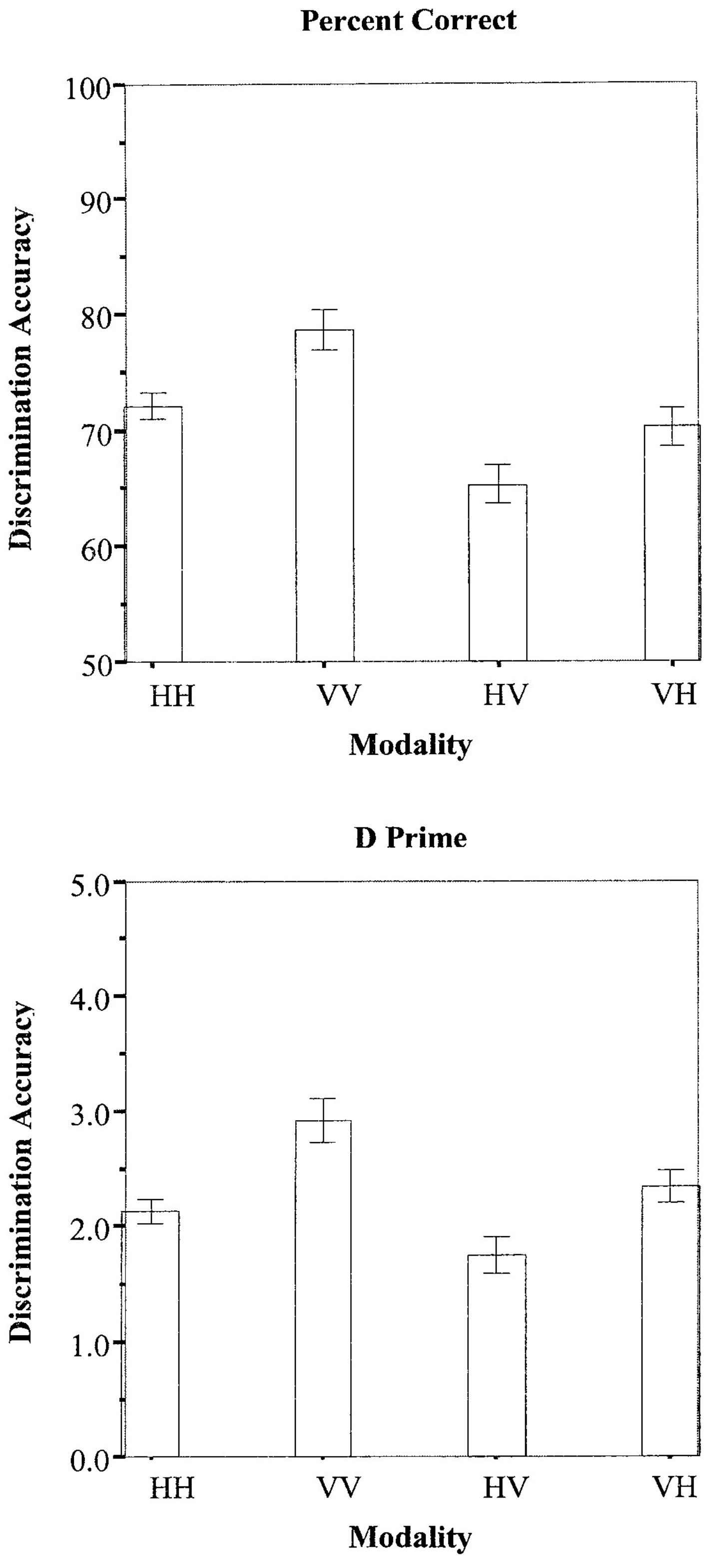

Figure 4. Discrimination performance for each modality condition, plotted both in terms of percent correct (top graph) and d' (bottom graph).

shape. It was more common for them to incorrectly perceive the objects as having the same shape when they were actually different than to incorrectly perceive them as being different in shape when they were actually the same. An interaction between trial type and inter-stimulus interval (ISI) was observed (F (2, 84) = 4.12, $p = .022$) and a Fisher LSD post-hoc analysis revealed that for "different" trials only, performance was significantly better when there was a 15-second ISI (M = 70.58) than when there was a 3-second ISI (M = 60.42), $p = .006$. For this interaction, there were no performance differences between ISIs within "same" trials. An interaction was also observed between trial type and modality, F (3, 84) = 12.62, $p < .001$. An LSD post-hoc analysis revealed that for "same" trials performance in the vision-haptic group (M = 82.21) was significantly better ($p = .003$) than for the haptic-haptic group (M = 71.74). For "different" trials, both the vision-vision (M = 79.79) and haptic-haptic (M = 72.29) groups were significantly better ($p < .001$) than both the haptic-vision (M = 52.29) and vision-haptic groups (M = 58.13). In other words, when the objects presented were different in shape, participants were significantly better at discriminating between them if both objects were perceived within the same modality (VV or HH) rather than across the modalities (VH or HV). A graph of this interaction is shown in Figure 5.

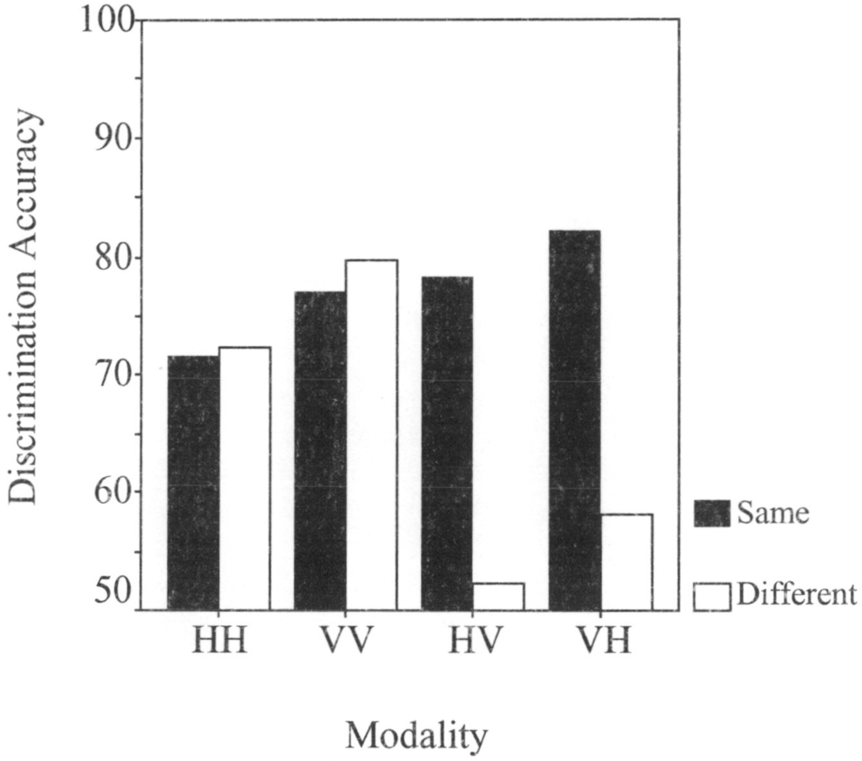

Figure 5. Graph of interaction between trial type (“same” v. “different”) and modality condition (HH, VV, HV, VH). For the unimodal (HH and VV) conditions performance is slightly worse for “same” trials as compared to “different” trials. For the cross-modal conditions (HV and VH) performance is much worse on “different” trials as compared to “same” trials. Discrimination accuracy is plotted in terms of percent correct.

Chapter 4

Discussion

The results from the current experiment give us a better understanding of the similarities and differences between the visual and haptic sensory modalities' representations of 3-D shape. It extends the findings of previous research on the cross-modal transfer of shape information, including the works of Gibson (1962, 1963, 1966) and Norman et al. (2004). Based on a series of experiments, Gibson concluded that "seeing and touching are two ways of getting much the same information about the world" (1979, p. 258). The results of Norman et al. (2004) confirmed this statement with the finding that participants could compare object shapes equally well when both objects were presented either visually or haptically. However, when shape information had to be transferred across the modalities, performance significantly decreased. This decrease in performance suggested that a possible recoding of shape information was necessary for the transfer to take place. While haptics and vision are both capable of perceiving object shape, they may give different weights to different features. They "code" features differently.

Like Norman et al. (2004) the current experiment found both similarities and differences between the visual and haptic modalities. A major similarity was in the pattern of learning for both the unimodal (visual-visual and haptic-haptic) and cross-modal (visual-haptic and haptic-visual) conditions (see Figures 2 and 3). The overall trend of improvement from subsession 1 to subsession 5 was linear, or in other words, it increased gradually. While this pattern is different from the one observed by Norman et al. in their cross-modal matching task (significant improvement from session 1 to session

2 and then a leveling off in performance), it not surprising. In a sense, the current experiment employed a more difficult discrimination task by using only object pairs that were highly confusable (based on Norman et al., 2004, Exp 1.). Even when a presented pair contained objects that differed in local shape they were still globally similar. This increased level of difficulty resulted in a more gradual improvement, as the participants gained increasing experience with the objects. It can be assumed that had more subsessions been added on to the experiment, (6, 7, 8, etc.) performance would have eventually leveled off in a manner similar to what was observed in Norman et al. (2004).

The significant main effect observed for modality (HH, VV, HV, and VH) gave evidence that different levels of discrimination accuracy exist for unimodal and cross-modal shape judgments (see Figure 4). In contrast to what was observed in Norman et al. (2004), in the current experiment discrimination performance for the vision-vision group was significantly better than for the haptic-haptic group. This finding again suggests perhaps coding differences and/or differences in feature salience for the two modalities. For the specific task given, the visual modality does appear to be superior to the haptic modality. It is important, though, not to over generalize this superiority of vision or the nonequivalence of the two modalities. While vision-vision appeared to be the "best" modality condition for performance on the discrimination task, haptic-vision was the worst. When analyzed both in terms of percent correct and d', it was worse than the other cross-modal condition (vision-haptic). It can be concluded that comparing an object presented visually to an object initially encoded haptically is a more difficult task than comparing an object haptically to an object that was initially encoded visually. While the studies of Gibson and Walker (1984), Steri (1987), and Steri and Gentaz

(2003) found the opposite to be true (touch-to-vision was superior to vision-to touch) with infants up to 2-months of age, it is important to remember that visual acuity is not fully developed until around 8 months of age. At the infant stage in development touch (both oral and haptic) appears to be the superior modality for object exploration (for infant perception review see Siegler, DeLoache, & Eisenberg, 2003). However the current finding, when interpreted in conjunction with the findings of Bushnell and Weinberger (1987), suggests that vision allows for the encoding of more information about an object's shape than touch from approximately 11-months of age on through adulthood.

Like learning, all of the modality conditions seemed to be equally affected (or unaffected) by the various interval lengths between the presentation of the first and second objects. Increasing the inter-stimulus interval (ISI) from 3 to 9 to 15 seconds had no effect on discrimination performance. As no distracter task was employed during the ISI, it can be hypothesized that the participants actively rehearsed the shape of the first object until the presentation of the second. It is difficult though to image how one mentally rehearses tactual mental images. One possibility is that participants rehearsed specific salient features on the objects (e.g., three bumps and a wide, shallow trough), as opposed to compiling the features into a whole, global mental image. In a similar study conducted by Abravanel (1972), participants were allowed to manipulate an object (one of Gibson's original "feelies") for three seconds. After a 15-second ISI, they were then presented with two objects successively for 3 seconds each and required to decide which of the two objects was identical in shape to the first object that had been presented to them. The same four modality conditions were used as in the current experiment. Like the current experiment, the participants did not seem to be affected by the 15-second ISI. Though only ten trials were conducted for each participant, the mean correct was well above chance for all groups.

An important goal of the current study was to get a clearer picture of the error differences that were observed in Norman et al. (2004). As a result, discrimination performance for "same" trials (both objects same in shape) and discrimination performance for "different" trials (objects have different shapes) were analyzed separately. It was found that, overall, participants were better at identifying "same" pairs than "different" pairs. Due to the fact that the presented pairs were highly similar even when they had different shapes it is not surprising that participants had more difficulty discriminating between the "different" pairs than the "same" pairs. In the only significant result related to ISI, it was found that on "different" trials participants performed significantly better in the 15-second ISI condition than in the 3-second ISI condition. While at first this looks puzzling, if taken into account with the fact that overall a 15-second ISI was no different than a 3-second ISI, this interaction can be explained by rehearsal. Perhaps having a longer period to rehearse the shape of the first object allowed for easier discrimination from the second object. However, this extra rehearsal was helpful only in that it decreased the likelihood of incorrectly perceiving a "different" pair as being the "same." The extra rehearsal was not necessary for identifying "same" pairs.

Not only were the different modality conditions found to have different levels of discrimination accuracy but they also affected how observers performed on "same" and "different" trials. For "different" trials a profound split surfaced between the unimodal (HH and VV) and cross-modal (HV and VH) conditions (see Figure 5). While observers in the unimodal groups were actually a little better (though not significantly) at discriminating between "different" pairs than "same" pairs, for observers in the cross-modal groups discrimination accuracy plummeted on "different" pairs. In fact, for the cross-modal conditions performance was barely even above chance (HV—*M* = 52.29 percent correct; VH—*M* = 58.13 percent correct). However for "same" trials, the cross-modal groups outperformed the unimodal groups, though the only significant difference was between the vision-haptic and haptic-haptic conditions. Again these differences can be explained by the confusability of the "different" object pairs. When the shape information had to be transferred across the modalities, it became more difficult to distinguish among the different local features of the objects. Either during transfer or perhaps recoding, important information was lost. As a result, the participants predominately perceived the second object to be the same in shape as the first object, whether it was or not. However, when the shape information stayed within the same modality, the information concerning the local shape characteristics remained intact. As a result, the participants made approximately the same number of errors on "same" and "different" trials.

Overall the current study found important similarities and differences between the haptic and visual modalities. When it comes to the perception of object shape, both modalities appear to learn at the same rate and have approximately equal capacities for short-term memory. They are not, however, equal in their ability. In the current study, vision appeared to be superior to haptics, and the transfer of information across the modalities made the cross-modal tasks significantly more difficult to complete. Taken together, these findings suggest a possible difference in how the modalities code shape information. Some interesting future studies would be to examine visual and haptic discrimination accuracies for texture and size, as opposed to shape. Though vision and touch are both capable of perceiving texture and size, it is difficult to say which modality would be superior at discriminating between objects based solely on those dimensions. If haptic perception truly is an "inferior form of vision" (Klatzky, Lederman, & Reed, 1987, p. 357), then vision should always come out on top. However if the difference lies in encoding processes, what is salient for the visual modality may not be salient for the haptic modality, and there is still room for haptic superiority on some dimensions.

References

Abravanel, E. (1972). Short-term memory for shape information processed intra- and intermodally at three ages. *Perceptual and Motor Skills, 35,* 419-425.

Aleman, A., van Lee, L., Mantione, M. H. M., Verkoijen, I. G., & de Haan, E. H. F (2001). Visual imagery without visual experience: Evidence from congenitally totally blind people. *Neuroreport, 12(11),* 2601-2604.

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience, 4(3),* 324-330.

Berkeley, G. (1963). *A new theory of vision and other writings.* New York: Dutton. (Original work published 1709).

Bliss, J. C., Crane, H. D., Mansfield, P. K., & Townsend, J. T. (1966). Information available in brief tactile presentations. *Perception and Psychophysics, 1,* 273-283.

Bushnell, E. W. (1982). Visual-tactual knowledge in 8-, 9 ½-, and 11-month-old infants. *Infant Behavior and Development, 5,* 63-75.

Bushnell, E. W. & Weinberger, N. (1987). Infants' detection of visual-tactual discrepancies: Asymmetries that indicate a directive role of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 13(4),* 601-608.

Cytowic, R. E. (1993). *The man who tasted shapes.* New York, NY: Warner Books.

Deibert, E., Kraut, M., Kremen, S., & Hart, J., Jr. (1999). Neural pathways in tactile object recognition. *Neurology, 52(7),* 1413-1417.

Gibson, E. J. & Walker, A. S. (1984). Development of knowledge of visual-tactual affordances of substance. *Child Development, 55,* 453-460.

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

Gibson, J. J. (1963). The useful dimensions of sensitivity. *American Psychologist, 18,* 1-15.

Gibson, J. J. (1966). *The senses considered as perceptual systems.* Boston, MA: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin.

Goodnow, J. J. (1971). Eye and hand: Differential memory and its effect on matching. *Neuropsychologia, 9,* 89-95.

Hunt, R. R. & Ellis, H. C. (2004). Short-term, working memory. In *Fundamentals of Cognitive Psychology* (7th ed.) (pp. 111-138). New York: McGraw Hill.

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.

Klatzky, R. L. & Lederman, S. J. (2002). Tactile object perception and the perceptual stream. In L. Albertazzi (Ed.), *Unfolding perceptual continua: Vol. 41. Advances in consciousness research.* (pp. 147-162). Amsterdam, Netherlands: John Benjamins Publishing Company.

Klatzky, R. L., Lederman, S. J., & Reed, C. (1987). There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *Journal of Experimental Psychology: General, 116(4),* 356-369.

Lakatos, S. & Marks, L. E. (1999). Haptic form perception: Relative salience of local and global features. *Perception & Psychophysics, 61(5),* 895-908.

Lucretius Carus, T. (1950). *Of the nature of things* (W. E. Leonard, Trans.) New York: E. P. Dutton. (Original work published ~58 B.C.E.).

Macmillan, N. A. & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Zielke, G. (2004). The visual and haptic perception of natural object shape. *Perception & Psychophysics, 66,* 342-351.

Siegler, R., DeLoache, J., & Eisenberg, N. (2003). Infancy: Perception, motor development, learning, and cognition. In *How Children Develop* (pp. 168-207). New York: Worth Publishers.

Streri, A. (1987). Tactile discrimination of shape and intermodal transfer in 2- to 3-month-old infants. *The British Psychological Society, 5,* 213-220.

Streri, A. (1993). *Seeing, reaching, touching: The relations between vision and touch in infancy.* (T. Pownall & S. Kingerlee, Trans.) Cambridge, MA: The MIT Press. (Original work published 1991).

Streri, A. & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory & Motor Research, 20(1),* 13-18.