Isomorphic Categories

by

Ryan Ferguson

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Arts

Approved June 2011 by the
Graduate Supervisory Committee:

Donald Homa, Chair
Arthur Glenberg
Stephen Goldinger

ARIZONA STATE UNIVERSITY

August 2011

ABSTRACT

Learning and transfer were investigated for a categorical structure in which relevant stimulus information could be mapped without loss from one modality to another. The category space was composed of three non-overlapping, linearly-separable categories. Each stimulus was composed of a sequence of on-off events that varied in duration and number of sub-events (complexity). Categories were learned visually, haptically, or auditorily, and transferred to the same or an alternate modality. The transfer set contained old, new, and prototype stimuli, and subjects made both classification and recognition judgments. The results showed an early learning advantage in the visual modality, with transfer performance varying among the conditions in both classification and recognition. In general, classification accuracy was highest for the category prototype, with false recognition of the category prototype higher in the cross-modality conditions. The results are discussed in terms of current theories in modality transfer, and shed preliminary light on categorical transfer of temporal stimuli.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

Human's amazing ability to categorize different objects, feelings, or experiences has been a central focus of cognitive psychology since its inception. Initial studies of categorization focused on paradigm development (Fisher, 1916; Hull, 1920; Smoke,1932), hypothesis-testing(e.g.,Bruner,Goodnow, &Austin,1956; Bourne,1966) and learning variables (e.g., Homa,1984). The development of mathematical models for categorization is more recent and has occurred primarily in the past 20 years (Nosofsky, 1984; Ashby, Alfonso-Reese, & Turken, 1998). The majority of formal and quantitative models of categorization fall within three classes. Prototype models (Reed, 1972; Minda & Smith, 2001) assume that subjects categorize based on the similarity of the stimuli to the prototypical stimuli of each category. The prototypes which represent each category are based upon a central tendency for that category built up by an integration of the observed examples within that category. Exemplar models (Medin & Schaffer, 1997; Nosofsky, 1998) suggest that subjects compute the similarity of any given stimulus to every exemplar in memory of each possible category. Decision bound models (Ashby & Gott, 1988; Ashby & Maddox, 1993) propose that subjects associate category responses to different regions of perceptual space and accept stimuli as category members depending on if they fall within the defined perceptual region. A myriad of connectionist models also exist that are often derivations from the previously discussed models; Kruschke's ALCOVE (1996) reduces to an exemplar-based model and Metcalfe's

Holographic model (1982) is a prototype model in which pattern features are combined via convolution. Knapp & Anderson (1984) has formulated an early connectionist model that can act either as a prototype or exemplar model depending on the number of exemplars within a category. Although there is empirical evidence which supports any of the classes of categorization models (and their endless variations), arriving at the conclusion that any one class of models appropriately describes all of human categorization would be built upon the erroneous assumption that subjects process category information similarly in every situation (Homa, 1984).

Modern categorization experiments attempt to capture the natural and commonplace experience of category exposure by randomly presenting variable instances from multiple domains. These experiments are usually deficient in one, critical way – virtually all studies explore stimuli presented through a single modal input. .Most sensory events we experience in the real world deal with some level of sensory integration that combines multiple inputs even if the information they provide is redundant (Stein & Meredith, 1993). However, the vast majority of research attempting to model categorization has dealt with category learning of stimuli which are constrained to a single modality. Visual stimuli have dominated category research with stimuli ranging from random dot patterns (Posner & Keele,1968) to rocket ships (Nosofsky, Palmeri, & McKinley, 1994). Auditory (Pitt, 1994), and to a much lesser extent, haptic stimuli (Homa et al., 2009) have been used infrequently to explore category learning. Multi-modal stimuli have rarely been used in categorization tasks, even though perception is typically multisensory and naturally occurring categories almost always involve multimodal integration in real life.

The integration of modalities into a common precept serves two functions: to maximize the information delivered from the different sensory modalities and to reduce the variance in the sensory estimate in order to increase its reliability (Ernst & Bulthoff, 2004). The mechanisms, both physically and psychologically, that work to achieve these goals are not as straight forward. Ernst and Bulthoff (2004) propose a model in which information from each modality is weighted and summed to form a robust percept. This weighting is hypothesized to be done on the fly by averaging the fluctuations of a signal over time. One study by Ernst and Banks (2002) illustrates this dynamic weighting principal. In this experiment, a subject viewed a reflected display in a mirror. Behind the mirror, the subject's hands were fitted into a haptic feedback device that could provide feedback based on the position of the subject's fingers. The subjects were asked to make size estimations of a line segment that was presented both on the mirror and represented spatially between their fingers via the force-feedback device. They first looked at the subject's estimations using either the visual or haptic modality, and then looked at how the visual and haptic modality interacted when information was available from both. The visual modality was weighted heavily when there was no manipulation of the display present. However, when visual noise was added to the display, subjects began to rely more heavily on the simultaneously available haptic information. They concluded that this is evidence that the nervous system has access to sensory reliabilities and adjusts accordingly to maximize performance on a given task.

A body of research does exist on the processing of stimuli which could be represented in more than one modality. Glenberg and Jona (1991) presented a sequence of rhythms consisting of both long and short events in either the auditory or visual modality. Subjects were then asked to recreate the sequences of long and short events. They found that modality differences did occur, namely with an auditory advantage for the reproduction of the temporal rhythms in general. This auditory advantage decreased as the inter-stimulus interval between the elements of the rhythms increased. Collier and Logan (2000) investigated short term memory performance for similar auditory and visually presented rhythms. They asked subjects to make same/different judgments about two rhythms that were presented sequentially either visually or auditorily. They too found an auditory superiority effect which decayed as presentation rate slowed. Watkins et al. (1992) had subjects to recreate sequences of flashes and beeps which had variable ISIs by tapping out the sequences on a computer. They found no differences between auditory and visually presented stimuli unless subjects were asked to mouth an irrelevant syllable during stimuli presentation, suggesting that sub-vocalization may have been occurring during stimulus presentation. Bresciani, Dammeier and Ernst (2008) also examined the perception of sequences of events, but included haptics along with vision and audition. Subjects were presented a number of beeps, tones, and taps and asked to focus on the number of events occurring in one modality (the target) and ignore the rest (the background).

They found that the visual modality was most susceptible to background-evoked bias and the least effective in biasing the other two modalities; audition was the least susceptible to background-evoked bias and the most susceptible at biasing the other modalities. In all cases, the background did bias the target response, leading the authors to conclude that the three modalities were automatically integrated. Notably lacking, however, are studies investigating non-modal specific stimuli using a category learning paradigm.

The primary focus of the present study is to investigate how categories are learned where the stimuli can be presented visually, auditorily, haptically with an underlying isomorphic structure. By isomorphic, we mean that the underlying abstract structure permits a mapping from one modality to another without apparent loss of information. This can be contrasted with studies that use stimuli that vary along dimensions that are more amenable to one modality than others, e.g., texture can be processed in greater detail in the haptic modality than in the visual modality (e.g., Pensky, Johnson, Haag, & Homa, 2008). The conditions explore category learning where the categories are structured identically but where only the modality of input differs. We are also interested in transfer performance following learning, both classification of novel patterns ("Which category does this stimulus belong to?") and recognition ("Is this stimulus old or new?"). Of major interest in the later is the performance costs involved in learning a category in one modality and being tested on the category in another modality. If novel stimuli are as easily recognized in a separate modality than that which they are learned, this could provide evidence that the category structure is amodal (Barsalou, 1999) or the abstraction process strips modality information.

The basic category structure is shown in Figure 1, where the duration of the stimulus is shown on the X-axis, and the stimulus complexity (defined as the number of activations within the duration) is shown on the Y-axis. This structure was inspired by Shepard's multidimensal scaling (1963) of Rothkopf's Morse Code confusability matrix (1957), in which he found he could represent the similarity of Morse Code letters and digits on a two dimensional plane with the number of components on one axis (here termed 'complexity') and the ratio of dots to dashes on the other (which is analogous to length).

Chapter 2

PREDICTED RESULTS

Given the paucity of data (and theory) on the learning of haptic and auditory categories that have a prototype structure, strong predictions are premature. Nonetheless, if modality of input is less critical that is the available stimulus information, then similar learning rates for each modality of input is predicted. That is, we should expect error rates to systematically decline across learning blocks, regardless of modality of input. In a similar vein, we might then predict that classification and recognition scores from the transfer test will be similar in all within-modal conditions (Auditory to Auditory, Visual to Visual, etc.). This would be consistent with the assumption that regardless of the modality of learning, the category structure is perceived and processed by each modality with equal facility. However, if processing differences exist among the various modalities, even with an isomorphic structure, then an advantage for one modality of input might occur (e.g., Glenberg and Jona (1991)). It is possible, for example, that even isomorphic structures might be retained differently in the different modalities or perhaps the integration of patterns might be superior in one modality than another. Regardless, we expect that learning and transfer would be similar for categories learned visually, haptically, or auditorily.

Another set of predictions exists for the cross-modality conditions at the time of transfer, with particular interest on the category prototype. To the extent that information is fully transferable between modalities for isomorphic categories, then classification of old and novel patterns is expected. That is, classification accuracy of transfer stimuli that were ‘learned’ in a different modality should be classified with comparable accuracy to conditions involving same-modality transfer. Consistent with previous category literature, we would predict that the category prototype might be classified best of all (e.g., Posner & Keele, 1968; Homa, 1984).

Results for the recognition test in the same and cross-modality conditions are less clear but potentially most intriguing. A common finding is that the category prototype is often falsely recognized as old (e.g., Omohundro, 1981), an outcome that has been found as consistent with both prototype models (Omohundro, 1981) but also claimed by advocates of exemplar models (Nosofsky & Zaki, 1998). However, the recognition test used in the cross-modality conditions of the present study is unique – subjects will be asked to identify a pattern as ‘old’ if it appeared in its isomorphic form in an alternate modality. If the subject stores only particulars and preserves the modality of input during learning, then the category prototype in the cross-modality conditions should be called ‘new’. That is, there is, currently, no mechanism in exemplar theory which permits information transfer from one modality to another. Without additional assumptions, the summed familiarity to stored instances in one modality should be low when patterns in an alternate modality are presented.

In contrast, prototype theory, in an expanded form, could accommodate this result. Typically, prototype theory has assumed that the training patterns are integrated in memory, with the summary representation functioning as the category prototype. However, no current theory of prototype abstraction assumes that the summary representation is also modality-free. Modality-free representation has been argued in other domains of memory. The foremost advocate of modality-free memory storage is Pylyshyn (1973) who has asserted that all information is ultimately coded in terms of modality-free assertions or properties. Should the prototype be falsely recognized at high rates in the cross-modality conditions – perhaps even at higher rates than in the same-modality conditions - then preliminary support would be provided for the view that the abstracted prototype is modality-free as well.

Chapter 3

METHOD

Participants

A total of 216 participants were used in the current study. Five participants were excluded for having learning errors above chance by the final learning phase.

Stimuli and Design

Stimuli were designed to encompass two non-modal specific dimensions: duration and complexity. Duration ranges from one to three seconds. Complexity was defined as the number of times an even occurred within the duration. For instance, a visual stimulus with a complexity level of two and a duration level of one second might be a light turning on for 300ms, off for 400ms, and then back on for 300ms. Complexity ranges from 2 to 10 activations within the given duration. Complexity and duration were correlated so that the longer the duration of the stimulus, more activations would occur within that time (see Figure 1).

Three combinations of duration and complexity were chosen to act as prototypes. Eight category members were generated from each of the three prototypes by moving up or down a level in duration and/or complexity. Levels in duration defined as +/-500ms of total stimulus duration and levels of complexity were +/- 1 activation. A bin-sorting algorithm which was constrained by the level of complexity and duration was then used to distribute the activations among the stimuli randomly. For instance, the simplest stimulus has a length of one second and a complexity level of two, meaning two activations. The algorithm would sort

20 units of duration (each being 50ms) into three ‘bins’ which represented the onset and offset of the stimulus (the two activations) and the ‘off bin’ which separated them in which there was no activity. Figure 2 illustrates the prototypes from each category, with each gray square representing a 50ms ‘on’ period (e.g., an LED light or the buzzer is activated) and each white rectangle representing a 50ms ‘off’ where none of the elements are active.

Each category consists of 9 stimuli total- 1 prototype and 8 distortions. Of these nine stimuli, four were chosen as learning trial stimuli that would be presented in the first phase of the experiment, and the remaining five stimuli (including the prototype) would be presented in the transfer phase of the experiment. The categories were linearly discriminable in each dimension. The duration and complexity of each stimulus were used as parameters and a computer generated and arranged the positions and length of the activations within each stimuli. The stimuli were programmed into an Arduino Diecimila microcontroller where they were controlled through the serial console of a Windows XP computer. Either a light emitting diode (LED), a 700Hz buzzer, or a 14k RPM vibrating motor were connected to the Arduino in order to deliver the stimuli to the subject.

Procedure

Subjects were randomly assigned to one of nine learning/testing conditions: visual-visual, visual-auditory, visual-haptic, auditory-auditory, auditory-visual, auditory-haptic, haptic-haptic, haptic-visual, and haptic-auditory. All subjects in each condition were read the same instructions informing them that they would be charged with the task of categorizing sequences of events into three categories. Subjects were told that there would be two stages to the experiment, a learning phase and a final testing phase. They initially received instructions about the learning phase and were told that instructions for the final testing phase would be given at that time. Each learning phase consisted of three study/test blocks containing each of the categories' 4 learning trial stimuli. The twelve stimuli were presented in random order using either the light, buzzer, or vibrating motor depending on the modal condition. Following the presentation of each individual stimulus, the experimenter informed the subject that the stimuli belonged to either category A, B, or C. If the subject was in the haptic condition, they were asked to wear headphones through which white noise was played so they could not get auditory feedback from the vibrating motor. After the experimenter presented all 12 learning stimuli and their respective categories, the same twelve stimuli were presented again in random order in the same modality and the subject was asked to place them in the correct category. Feedback was given as to whether or not the subject classified each stimulus correctly. This learning block was repeated three times. After the final study/test block, instructions for the transfer testing phase were read. Depending on the condition the modality of the final transfer phase was either the same as they experienced in the learning phase or switched to one of the other modalities. The transfer procedure uses that of previous studies (e.g., Omohundro, 1981) - the subject will be presented a stimulus and asked to render and old/new judgment. Following that, a category judgment (A,B,C) is required, followed by the next stimulus. All 9 stimuli in each of the three categories were presented resulting in a total of 27 recognition judgments and 27 category judgments. No feedback was given throughout the final testing phase.

Chapter 4

RESULTS

A repeated measures ANOVA was performed comparing the three learning trials from the three learning modalities. Significant learning occurred across learning trials, and pairwise comparisons confirmed that each learning trial had significantly less error than the trial preceding it, $F(2,426)=52, p<.001$. There was a significant main effect of modality $F(2,213)=3.31, p<.05$, and pairwise comparisons indicated that there was significantly fewer errors for the visual learning condition when compared to the haptic learning condition over all three trials (p<.05). There was a significant learning trial by learning modality interaction, $F(4,426)=3.006, p<.05$. Follow-up pairwise comparisons showed that the haptic condition was only significantly worse than the visual condition in learning trial one ($p<.01$), and haptic was significantly worse than both visual and auditory in trial two ($p<.05$), but by trial three there were no significant differences among the modalities as illustrated in Figure 3. Even though visual learning had an advantage in earlier learning trials, auditory learning ended up with the least amount of errors, however it was nonsignificant. Regardless, this does lend support for the auditory superiority in rhythm processing (Glenberg & Jona, 1991).

We analyzed the recognition data with the modalities collapsed into unimodal (audio to audio, etc.) and crossmodal (audio to visual, haptic to audio, etc.). We compared the recognition performance of the three types of items in the transfer test: old items from the learning trials, new items, and prototype items.

The results show that there was a significant main effect of item type $F(2,428)=12.05, p<.001$. There was also a significant main effect of modalities, with all items in the cross modal condition being called “old” more often , $F(1,214)=9.7, p<.01$. There was not a significant interaction. Post hoc tests showed all item types significantly different than one another with the prototype being called old the most ($M=.65$, $SE=.01$), followed by old items ($M=.60, SE=.01$), followed by new items ($M=.55, SE=.01$). The same analysis was performed without collapsed modal conditions, comparing each condition to one another. There was a significant main effect of item type, $F(2,207)=4.61, p<.05$, and of modal condition, $F(8,207)=3.95, p<.001$. There was no significant interaction. Pairwise comparisons showed that the unimodal haptic to haptic condition called items old significantly less than the haptic to visual ($p<.05$) or haptic to auditory ($p<.01$) group. The audio to visual group was also significantly less biased than the haptic to audio group ($p<.05$). Figure 4 shows the likelihood of calling an item old for each of the nine conditions.

Turning to transfer classification errors, we once again began with the modal conditions collapsed into unimodal and crossmodal groups. There was a significant difference in the classification performance across the different item types, $F(2,428)=12.14, p<.001$. There was no main effect of modality, nor was there a significant interaction. Pairwise comparisons showed that new items ($M=.23, SE=.01$) were classified with significantly more error than both old ($M=.18, SE=.01$) and prototype ($M=.16, SE=.02$) items. With the expanded modal category, there was of course still no main effect of modality. Item was still significant, $F(2,414)=15.574$, $p<.05$. Figure 5 shows the probability of an error on the transfer test for each item type and condition.

Another interesting finding was looking at the cost of crossmodal transfer in classification, as shown in Figure 6. Relative to the unimodal conditions, the crossmodal conditions classify items in the transfer test worse only in the auditory and haptic learning conditions. When subjects learned the structure visually, they actually tended to perform better in the crossmodal conditions.

Chapter 5

DISCUSSION

The present study investigated the learning and later transfer of categorical patterns where the stimuli were defined by a temporal sequence of events. Unique to the present study is the use of a categorical structure that was isomorphic across modality of presentation. By isomorphic, we mean that a mapping existed between the stimuli in different domains (Shepard & Chipman, 1970). In the present study, each temporal stimulus presented in one modality (e.g., a visual sequence of events) could be reproduced, without loss of information, into an alternate modality (e.g., a haptic sequence). Of critical concern was whether categorical information, acquired in one modality, could be transferred without significant loss into an alternate modality.

In the present study, nine separate conditions were run, three unimodal and six crossmodal. Common to each condition was the initial learning of three categories that were linearly separable in two dimensions. In the unimodal conditions, learning and transfer always occurred in the same modality; in the crossmodal conditions, transfer occurred to an alternate modality. Although considerable research has explored cross modal transfer in stimulus identification paradigms (Glenberg & Jona, 1991; Collier & Logan, 2000, Klatzky & Lederman, 1985; Guttman, Gilroy & Blake, 2005.), almost nothing is known about crossmodal transfer involving categorical knowledge. To my knowledge, no one has explored crossmodal transfer of categorical information for an isomorphic structure. Normally, the various modalities provide both shared (common) information about a stimulus, as well as unique properties. Thus, shape can be acquired both by vision and touch, whereas texture and malleability is gained primarily from touch. Transfer between modalities is often quite good but clear differences also can be obtained (e.g., Pensky, Johnson, Haag, & Homa, 2009). The question that arises in the present study is this - if the stimulus events are, in some sense, equivalent across modalities, will transfer differences be minimized? If there is a differential cost among the modalities, which are easily translated, and which are not?

In the present study, five major results were obtained: (a) Categories with an isomorphic structure were learned faster when apprehended visually, although all conditions asymtoped to a common terminal level; (b) Calling a pattern 'old' was highest for the prototype, regardless of learning or transfer modality; (c) Transfer accuracy, regardless of condition, followed by order prototype > old > new; (d) the haptic modality was associated with the greatest loss of information when transferred to an alternate modality; vision was the least affected; and (e) Calling the prototype 'old' was often higher in an alternate modality.

Each of these results is discussed briefly in turn. These results are then discussed in terms of a descriptive model which suggests that features directly experienced in one modality may activate corresponding features in an alternate modality.

The most basic result of the current study was that our isomorphic temporal categories are in fact learnable. Although there is a visual advantage in early learning, it tends to plateau and by the third learning trial all modalities see equal error rates. There was observed auditory superiority on the terminal learning trial, but it was not significantly different than the other modalities. Glenberg and Jona (1991) were able to eliminate the auditory advantage with rhythmic stimuli when the stimuli were made more complex by removing any relationship between their internal components. The stimuli in the present experiment were devoid of any rhythmic considerations or repetitions, and as such could be considered sufficiently complex to lack an auditory advantage.

The prototype of each category was false alarmed to in recognition testing during transfer more than both new and learning items. This is consistent with a number of findings (Posner & Keele, 1968; Homa, 1984). This suggests that subjects may be forming an abstracted prototype internally when they are learning the category structures, and this abstracted prototype is referenced during categorical decisions, resulting in a higher false alarm rate. When looking at unimodal versus crossmodal conditions, items received higher oldness ratings when the transfer modality differed from the learning modality, absent of any item type interaction. With the modal specific features of a stimulus no longer relevant, as when the modality is switched, recognition can no longer rely as much on source specific information and may be left to rely more on familiarity (McElree, Dolan & Jacoby, 1999). If a similar explanation is used to explain why subjects who learned the categories haptically had a significantly higher bias to call items old when modality was switched in transfer, this could be seen as evidence that the haptic modalities is less adept at encoding temporal stimuli versus the hearing and vision.

As opposed to recognition, classification performance in transfer did not reflect any significant differences between the modalities. However, a similar prototype effect was found with the prototype being classified with significantly less error than both the new items and the learning trial items. This type of finding (Homa,1984; Minda & Smith 2011) is once again consistent with the possibility that subjects are abstracting an internal prototype during category learning, and this prototype is sufficiently familiar by transfer that it is categorized with less error than even stimuli that subjects have had repeated exposure to. The lack of significant modality effects along with the relatively low overall classification error rates could suggest that subjects are learning the categories at a more amodal level of processing, where switching the modality on transfer would have a negligible effect. However, we cannot completely discount modalities effect in classification. As seen in Figure 6, classification in transfer did seem to be modulated by learning modality. Specifically, subjects who learned the stimuli visually seemed to have no deficit when transferring to other modalities. On the contrary, these subjects (in the visual-haptic and visual- auditory) performed better on classification than in the unimodal (visual to visual) condition. It is difficult to say what particular quality unique to vision would facilitate transfer. Freides (1974) suggested that when information is delivered to modality that is not ideal, that information is translated into the code of the more ideal modality. Visual learning is best when it deals with spatial rather than temporal information, and that auditory and haptic are similar in that they both possess temporal learning well (Mahar, Mackenzie & McNicol, 1994). If temporal learning is presented visually, then the process of translation may demand a deeper processing of the information.

One way of conceptualizing our results is represented by a descriptive model that includes the various modalities. Each encounter of a physical stimulus is represented by a feature vector containing several dimensions (e.g. Hintzman, 1986). Among these dimensions would be feature items that correspond with specific modal inputs, such as below:

$$Experience_n = \{v_1, v_2 \ldots v_n; a_1, a_2 \ldots a_n; h_1, h_2 \ldots h_n\}$$

where any encounter with any stimulus could be represented by a vector containing features that are experienced in a specific modality (v_1 being a visual feature, a_1 being an auditory feature, and h_1 being a haptic feature).

Because a great deal of human interaction with the world involves items which can deliver robust and diverse modal information, these modality specific vectors could contain both redundant (such as visual and proprioceptive estimates of size) and unique (such as the smell and taste of a vanilla bean) information for multimodal stimuli. What then would happen during an experience where not all of the modalities are involved, such as observing a basketball sitting on a shelf across the room? The simplest (computationally) output would assume that a vector that was only activated in the relevant modality would occur (so in this case vision) and remained unaffected in all other modalities :

$$E_{ball} = \{v_1, v_2 \ldots v_n; 0,0 \ldots a_n; 0,0 \ldots h_n\}$$

Such representations are problematic. If, while your gaze was not on the basketball on the shelf, you heard a boomy, rubbery bouncing noise, you would assume that it came from that same ball. Just as if the lights went out and you were fumbling through the dark trying to find a switch and your hand brushed the ball. As illustrated in the current experiment, stored experiences cannot exist statically in the same modality because cross modal transfer is possible with relatively little loss.

One solution to this problem would be to assume that the stored experiences are not static and that they could be manipulated when a situation called for it, such as the modality switch in our current experiment. This computed vector would be produced on demand and would need some sort of modality emulation system that was both quick and was based on past experience. Barsalou (1999) proposed such a system that employs internal simulations which determine if a novel input belongs to an existing concept. Under Barsalou's theory, any experience with a stimulus produces a rich multimodal framework in which manipulations of that stimulus can be carried out internally. Further experiences with this stimulus are integrated within this simulation, and experience with similar stimuli can be integrated into a more general simulation that can be used to test for category membership. For instance, extensive experience with one particular dog, Fido, would amass a simulator that is rich with multimodal information, including how he smells, sounds, feels, etc. Fido's simulation is a subset of the simulator for the more general concept of dog, which is an integrated account of all past dog experiences. When a new dog-like figure is heard barking in the back yard, this person first would see if the simulator for the general dog concept was able to recreate the dog-like figure and the barking without any glaring inconsistencies. If they were successful and the dog-like figure was indeed a dog, then the Fido simulator would attempt to simulate the figure and bark in turn. It is through this process of simulation that Barsalou is able to account for categorization. The fact that these simulators operate in a framework that is multimodal could account for the type of crossmodal transfer that we observed in the current experiment. While a subject was learning the categories in one modality, they were building a simulation of that category in which membership of learning stimuli was determined by the ability of that category's simulator to simulate that learning stimuli. When the crossmodal transfer was introduced, subjects were then able to use the existing category simulations to attempt to simulate the novel stimuli to determine if it did or did not belong to that category. Such a model would predict good classification in alternate modalities, just as we found. However, if simulators of each category were built up of experienced stimuli within that category, and novel stimuli (both in the same and different modality) required additional simulation and a comparison process to determine category membership, then you would expect good recognition of old stimuli just as a function of how difficult the additional simulation was. The recognition rates in our current experiment did not reflect this, as we had high false alarm rates and generally poor recognition performance.

A third possibility for storage of unimodal modality vectors exists in which the vector is stored with the experienced modality activated along with a partial resonant activation in the other modalities:

$$E_{ball} = \{v_1, v_2 \dots v_n; 0, a_2 \dots a_n; 0,0, h_3 \dots h_n\}$$

In this example, the object was experienced visually and, therefore, each visual feature has a value. However, some (but perhaps not all as indicated by zeros in some slots) auditory and haptic features would also be activated. Neurological research has found audio-visual (Calvert et. al, 1997; Giard & Peronnet, 1999), visual-haptic (Amedi, Malach, Hendler, Peled & Zohary, 2001), and auditory-somatosensory (Sperdin, Cappe, Foxe, & Murray, 2009) interactions in the brain given unimodal stimuli. These findings suggest that the neocortex as a whole is multisensory (Ghazanfar & Schroeder, 2006). James et. al (2002) found that haptic exploration of novel clay objects produced activation not only in the somatosensory cortex, but also in areas of the brain that are primarily associated with visual processing. Sathian and Zangaladze (2002) used transcranial magnetic stimulation (TMS) to disrupt processing in the same extrastriate visual cortex while subjects attempted to discriminate between different tactile gradients. When the TMS was activated, tactile discrimination was hindered, indicating that visual processing was necessary during the actual tactile discrimination task, not just used in simulations afterwards. So in our basketball illustration, visually experiencing the ball would not only activate the visual features in the vector, but would at the same time activate non-visual corresponding features that may represent features such as texture. Because one to one mappings from one modality to another do not exist, the activation in the ‘phantom modality’ would be weaker and possibly less reliable than in the experienced modality. In the context of our current experiment, this theory could allow for crossmodal transfer while at the same time account for poor recognition because of the decreased reliability of the crossmodal information. This could also account for the asymmetrical transfer shown in Figure 6 by suggesting that the visual modality has more resonant features in the auditory and haptic modalities than the other way around.

Overall, the current experiment demonstrated that categorical information can be reliably transferred from one modality to another. Classification error rates during transfer were low in both unimodal and crossmodal conditions, indicating that whatever internal representation was used for this category structure was robust for changes in modality. In every condition, the prototype had the fewest classification errors False alarms in recognition were higher in crossmodal conditions across all item types, and the prototype of each category was false alarmed to most of all. Further research into cross-modal categorization could employ a multiple modalities in the learning phase and a source identification task in the transfer phase, perhaps imbedded with delayed tests, to assess remaining questions of how accessible and resilient modality information of stored concepts really is.

REFERENCES

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. doi:10.1038/85201

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*(3), 442-481. doi:10.1037/0033-295X.105.3.442

Ashby, F. G., & Gott, R. E. (1988). Decision rules in the perception and categorization of multidimensional stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(1), 33-53. doi:10.1037/0278-7393.14.1.33

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-660. doi:10.1017/S0140525X99002149

Bourne, L.E. (1966). *Human conceptual behavior*. Boston: Allyn and Bacon.

Bresciani, J., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin, 75*(6), 753-760. doi:10.1016/j.brainresbull.2008.01.009

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. Oxford, England: John Wiley and Sons.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., & Campbell, R. (1997). Activation of auditory cortex during silent lipreading. *Science, 276*(5312), 593-596. doi:10.1126/science.276.5312.593

Collier, G. L., & Logan, G. (2000). Modality differences in short-term memory for rhythms. *Memory & Cognition, 28*(4), 529-538.

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

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*(4), 162-169. doi:10.1016/j.tics.2004.02.002

Fisher, S. C. (1916). The process of generalizing abstraction; and its product, the general concept. Psychological Monographs, *21*(2.90), 1–209.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278-285. doi:10.1016/j.tics.2006.04.008

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*(5), 473-490. doi:10.1162/089892999563544

Glenberg, A. M., & Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects. *Memory & Cognition*, *19*(5), 514-522.

Guttman, S. E., Gilroy, L. A., & Blake, R. (2005). Hearing what the eyes see: Auditory encoding of visual temporal sequences. *Psychological Science*, *16*(3), 228-235. doi:10.1111/j.0956-7976.2005.00808.x

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298*(5598), 1627-1630. doi:10.1126/science.1075396

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*(4), 411-428. doi:10.1037/0033-295X.93.4.411

Homa, D., Kahol, K., Tripathi, P., Bratton, L., & Panchanathan, S. (2009). Haptic concepts in the blind.*Attention, Perception, & Psychophysics, 71*(4), 690-698. doi:10.3758/APP.71.4.690

Homa, D., & Cultice, J. C. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 83-94. doi:10.1037/0278-7393.10.1.83

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs, 28*(1), 1-86.

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

Knapp, A, & Anderson, J. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(4), 616-637.

Klatzky, R. L., Lederman, S. J., & Metzger, V. A. (1985). Identifying objects by touch: An "expert system.".*Perception & Psychophysics, 37*(4), 299-302.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*(1), 22-44. doi:10.1037/0033-295X.99.1.22

Mahar, D., Mackenzie, B., & McNicol, D. (1994). Modality-specific differences in the processing of spatially, temporally, and spatiotemporally distributed information. *Perception, 23*(11), 1369-1386. doi:10.1068/p231369

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53*(1), 49-70.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(3), 563-582. doi:10.1037/0278-7393.25.3.563

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207-238. doi:10.1037/0033-295X.85.3.207

Metcalfe, J. A. (1982). A composite holographic associative recall model. ProQuest Information & Learning). (Electronic; Print). (1983-53209-001)

Minda, J. P., & Smith, J. D. (2001). Prototypes in category learning: The effects of category size, category structure, and stimulus complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(3), 775-799. doi:10.1037/0278-7393.27.3.775

Minda, J.P., & Smith, J.D. (2011). Prototype models of categorization:basic formulation, predictions, and limitations. In E. Pothos (Ed.*), Formal Approaches in Categorization* (pp. 40-64). New York, New York: Cambridge University Press.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review-New York-, 101*, 53-53.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(1), 104-114. doi:10.1037/0278-7393.10.1.104

Nosofsky, R. M., & Zaki, S. R. (1998). Dissociations between categorization and recognition in amnesic and normal individuals: An exemplar-based interpretation. *Psychological Science, 9*(4), 247-255. doi:10.1111/1467-9280.00051

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory & Cognition, 9*(3), 324-331.

Pensky, A. E. C., Johnson, K. A., Haag, S., & Homa, D. (2008). Delayed memory for visual-haptic exploration of familiar objects. *Psychonomic Bulletin & Review, 15*(3), 574-580. doi:10.3758/PBR.15.3.574

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology-Human Perception and Performance, 20*(5), 976-985.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology,77*(3), 353-363. doi:10.1037/h0025953

Pylyshyn, Z. (1973). What the Mind's Eye Tells the Mind's Brain. *Psychological Bulletin*, 1973.

Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*(3), 382-407. doi:10.1016/0010-0285(72)90014-X

Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology, 53*(2), 94-101.

Sathian, K., & Zangaladze, A. (2002). Feeling with the mind's eye: Contribution of visual cortex to tactile perception. *Behavioral Brain Research, 135*(1-2), 127-132. doi:10.1016/S0166-4328(02)00141-9

Shepard, R. N. (1963). Analysis of proximities as a technique for the study of information processing in man. *Human Factors, 5,* 33-48.

Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states.*Cognitive Psychology, 1*(1), 1-17. doi:10.1016/0010-0285(70)90002-2

Smoke, K. L. (1932). An objective study of concept formation. *Psychological Monographs.42, no, 4,* 46.

Sperdin, H. F., Cappe, C., Foxe, J. J., & Murray, M. M. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Frontiers in Integrative Neuroscience, 3.* doi:10.3389/neuro.07.002.2009

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

Watkins, M. J., LeCompte, D. C., Elliott, M. N., & Fish, S. B. (1992). Short-term memory for the timing of auditory and visual signals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 931-937. doi:10.1037/0278-7393.18.5.931

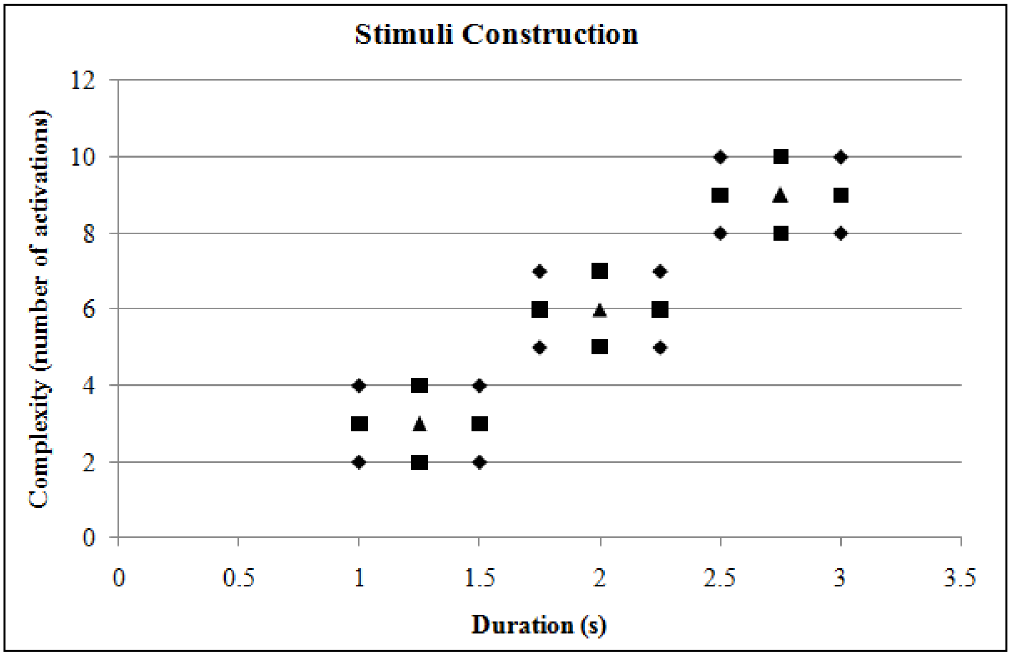

Figure 1. Stimuli Construction. The stimuli represented by a triangle are the category prototpyes. The square stimuli represent the learning trial stimuli. The diamond stimuli are category members that are not presented until the transfer test.

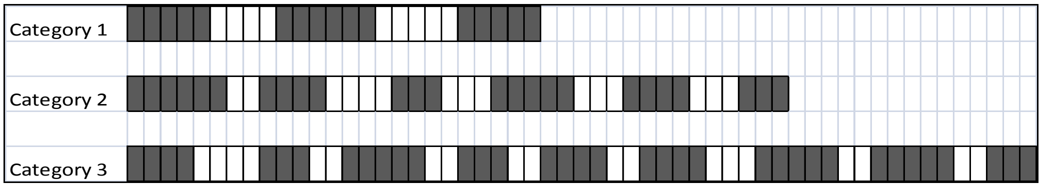

Figure 2. Temporal structure of the category prototypes. Each block represents 50ms. Gray squares are activated, white are inactive.

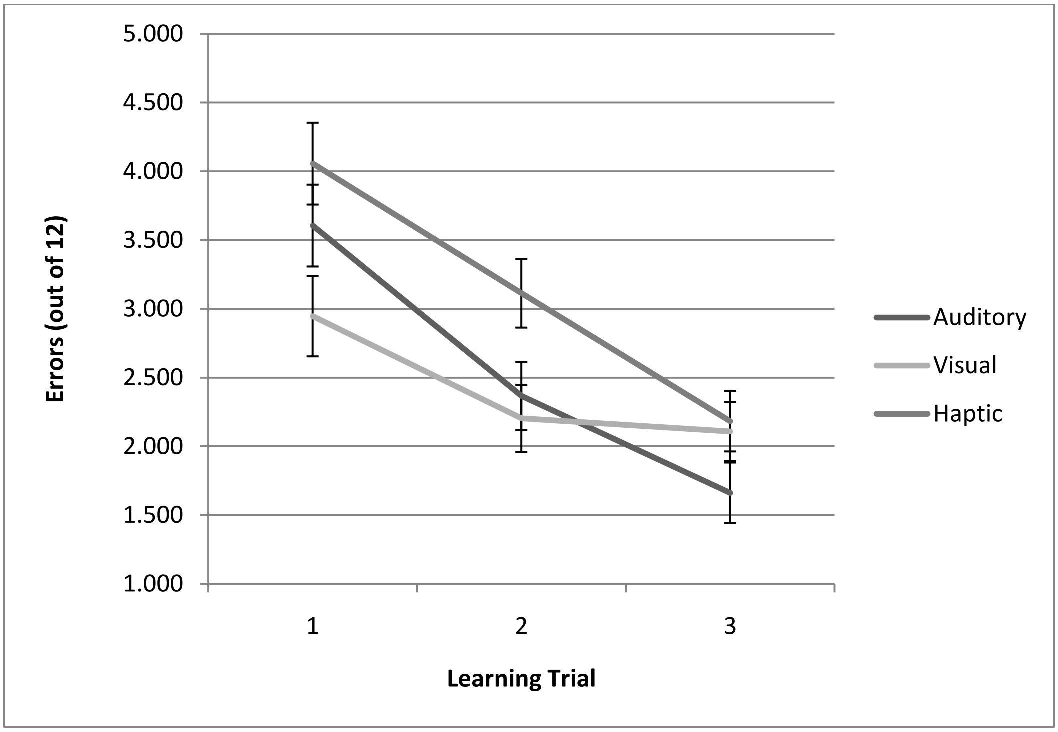

Figure 3. Learning trial errors. Learning trial errors out of 12 over three learning trials. There was no significant difference between modalities by the third trial.

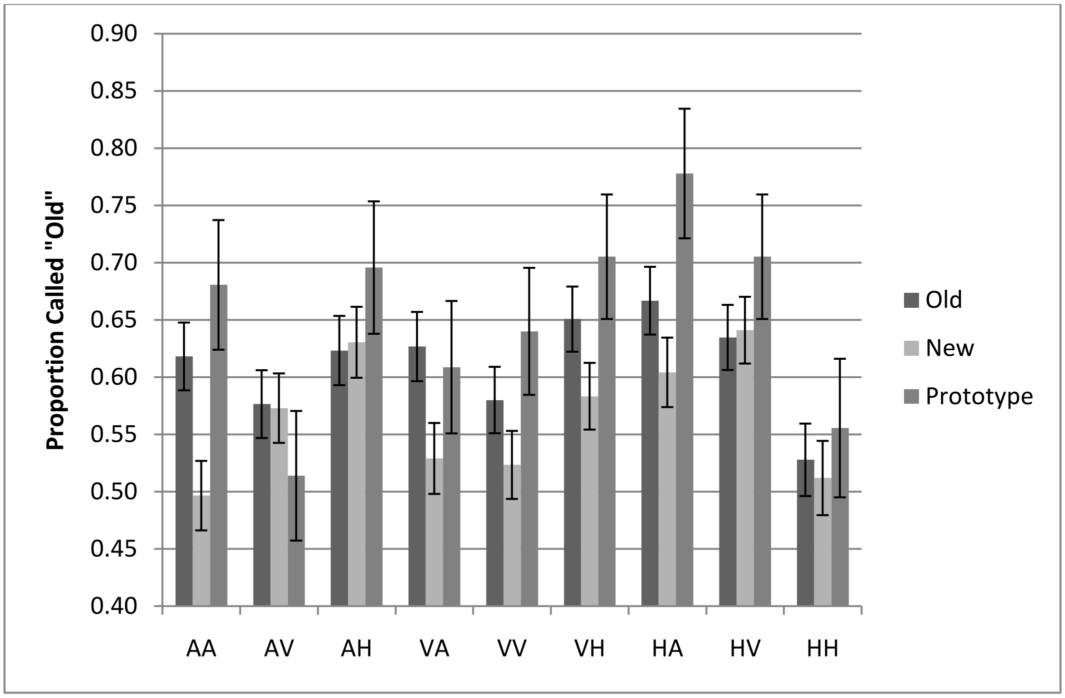

Figure 4. Recognition response for each condition. For each condition, the percent of stimuli in transfer called 'old.' The prototype was called old in all conditions but visual-audio, even though the subjects had never experienced it.

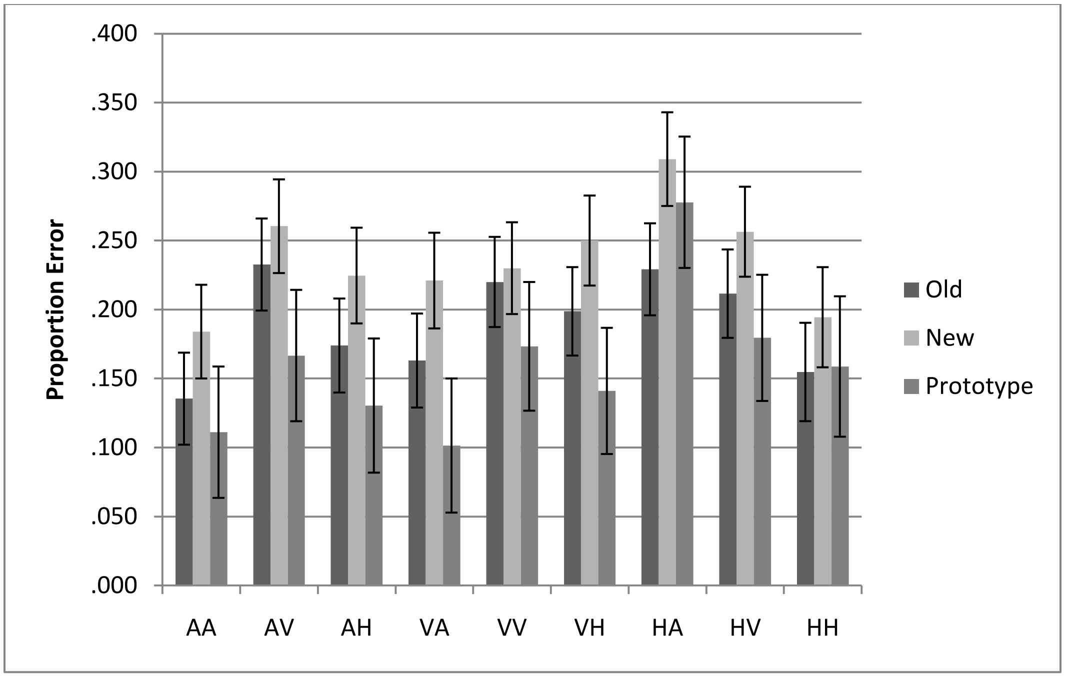

Figure 5. Classification error for each condition. The percent of error for each object type in each condition. For most cases, the prototype experienced the least amount of classification error.

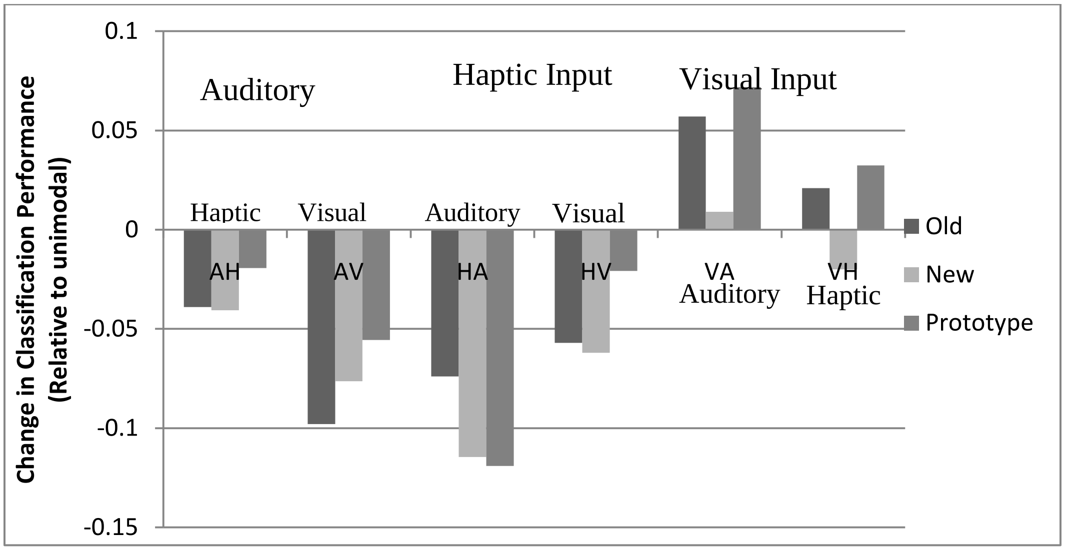

Figure 6. Classification cost of crossmodal transfer. This chart indicates the change in classification error performance in crossmodal conditions relative to the unimodal condition (shown here as the baseline