

2008

Integrating Conceptual Knowledge Within and Across Representational Modalities

Christopher P. McNorgan

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

McNorgan, Christopher P., "Integrating Conceptual Knowledge Within and Across Representational Modalities" (2008). *Digitized Theses*. 4852.
<https://ir.lib.uwo.ca/digitizedtheses/4852>

This Thesis is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Integrating Conceptual Knowledge Within and Across Representational Modalities

(Spine title: Multimodal Conceptual Integration)

(Thesis format: Monograph)

by

Christopher P. McNorgan

Graduate Program

in

Psychology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Christopher P. McNorgan 2008

Abstract

Research suggests that concepts are distributed across brain regions specialized for processing information from different sensorimotor modalities. Multimodal semantic models fall into one of two broad classes differentiated by the assumed hierarchy of convergence zones over which information is integrated. In shallow models, communication within- and between-modality is accomplished using either direct connectivity, or a central semantic hub. In deep models, modalities are connected by cascading integration sites with successively wider receptive fields. Four studies provide the first direct test of these models using speeded behavioural tasks involving feature inference and pattern completion. Shallow models predict no within- versus cross-modal difference in either task, whereas deep models predict a within-modal advantage for feature inference, but a cross-modal advantage for pattern completion. Study 1 investigated the prevalence of within- and cross-modal feature correlations in a large database of feature production norms. Studies 2 and 3 used relatedness judgments to tap participants' knowledge of relations for within- and cross-modal feature pairs. Study 4 was a dual feature verification task. The pattern of decision latencies across Studies 2 to 4 is consistent with a deep integration hierarchy.

Keywords: semantic memory, multimodal representations, binding problem, embodied cognition

Acknowledgements

Naturally, I should begin by thanking my parents, Richard and Donna McNorgan, whence *I* began, for I attribute my general desire to be self-sufficient, and by extension, my eagerness to learn new approaches to solving problems, to my upbringing. Ken McRae deserves much gratitude for his support and encouragement these last 67 years, and for calling me back to the academic life even after I had initially abandoned it. The members of my advisory committee and departmental examining committee: Marc Joannis, Stefan Köhler and Paul Minda are to be thanked for their useful feedback on earlier revisions of this document. I would also like to thank Larry Barsalou and Kyle Simmons for providing input when the ideas behind this project were just forming in my mind. Current and former lab mates Chris, Kaz, Saman and Ray have also been quite helpful in providing feedback and insight, most especially so at the Grad Club over a Guinness. Finally, the significance of my wife and best friend, Rebecca, cannot be understated, for she has made and will continue to make many sacrifices on my behalf, and for this I am eternally grateful.

TABLE OF CONTENTS

CERTIFICATE OF EXAMINATION	ii
Abstract.....	iii
Acknowledgements.....	iv
TABLE OF CONTENTS	v
List of Tables	vii
List of Figures.....	viii
List of Appendices.....	ix
Chapter One	1
Introduction.....	1
Multimodal versus Amodal Representations.....	2
Convergence Zones	4
Convergence Zone Hierarchies.....	5
Arguments for Deep and Shallow Convergence Zone Hierarchies.....	9
Proximity-Sensitive Integration.....	11
Chapter Two: Functional Connectivity and Feature Norms Analysis.....	20
Study 1: Feature Production Norms Analysis	20
Materials and Procedure	21
Results and Discussion	24
Chapter Three: Feature Integration in Feature Inference	30
Study 2	30
Method.....	30

Design	35
Results and Discussion	35
Study 3	37
Method.....	37
Results and Discussion	43
Chapter Four: Feature Integration in Pattern Completion	46
Study 4	46
Method.....	48
Design	55
Results and Discussion	55
General Discussion	60
Relation to Previous Research	63
Insights into Functional Connectivity.....	68
Perceptual and Conceptual Integration.....	70
Alternative Experimental Paradigms and Future Directions	72
Conclusion	76
References.....	77
Curriculum Vitae	90

List of Tables

Table 1	<i>Numerosity and proportion of total correlations by correlated pair type</i>	25
Table 2	<i>Mean magnitude of shared variance by correlated pair type</i>	26
Table 3	<i>Summed shared variance by correlated pair type</i>	29
Table 4	<i>Equated variables in Study 2</i>	31
Table 5	<i>Equated variables in Study 3</i>	39
Table 6	<i>Equated variables in Study 4 by density condition</i>	50
Table 7	<i>Equated variables in Study 4 by modality and density condition</i>	51
Table 8	<i>Feature verification latencies and error rates for Study 4</i>	56

List of Figures

<i>Figure 1.</i> A hierarchically shallow model with two directly interconnected modalities....	7
<i>Figure 2.</i> A hierarchically shallow model with two modalities connected via a single convergence zone.....	7
<i>Figure 3.</i> A hierarchically deep model containing three modalities, each possessing a unimodal convergence zone that feeds forward to bimodal and trimodal convergence zones.	8
<i>Figure 4.</i> Within- and cross-modal feature-feature activation in a shallow hierarchy takes the same number of processing steps.....	14
<i>Figure 5.</i> Feature-feature activation in a deep hierarchy takes fewer processing steps within-modally than cross-modally.	14
<i>Figure 6.</i> Feature-concept activation in a shallow hierarchy takes the same number of processing steps from within-modal pairs as from cross-modal pairs.....	15
<i>Figure 7.</i> Feature-concept activation in a deep hierarchy takes more processing steps from within-modal pairs than from cross-modal pairs.	16
<i>Figure 8.</i> Presentation sequence for stem and content components of stimuli in Study 3.	42

List of Appendices

Appendix A Within- and cross-modal pairs for Study 2	84
Appendix B Yoked within- and cross-modal pairs used in Study 3	86
Appendix C Yoked within- and cross-modal pairs and target concepts used in Study 4.	87
Appendix D Ethics Approval Notice	89

Chapter One

Introduction

Semantic memory contains a great deal of knowledge regarding lexical concepts such as *dog* and *banana*, and as such is important for language processing, perception, reasoning, and action. Concepts referring to living and nonliving things include information such as how something looks, tastes, feels, and sounds, and how it is used. The manner in which this knowledge is represented and organized therefore greatly impacts behavior. It is intuitive to think of lexical concepts in terms of features. For example, how a typical *dog* looks or sounds can be described by features such as *<has legs>*, *<has a tail>*, *<has a nose>*, *<has ears>*, *<barks>*, and so on. Although some models of semantic memory are not based on feature representations – for example, Latent Semantic Analysis (Landauer & Dumais, 1997) – feature-based models, which describe concepts as collections of features at some level of abstraction, dominate the literature. The manner in which types of featural knowledge are neurally organized and interact is one aspect of theories of semantic memory that differentiates among them. The goal of the present research is to provide the first direct test of two broad assumptions that have been made concerning how the brain organizes and uses different types of knowledge of concrete object concepts.

To gain insight into the kinds of features that may comprise people's concepts in a feature-based framework, researchers have used tasks in which participants list features such as *<has four legs>*, *<has fur>*, *<has a tail>*, and *<barks>* for concepts like *dog*. Features listed in these tasks have been useful in accounting for a range of behaviors, from similarity judgments (Tversky, 1977) to theory generation (Ahn, Marsh, Luhmann,

& Lee, 2002; McNorgan, Kotack, Meehan & McRae, 2007). Although some features listed for concepts like *dog* (<*man's best friend*>) reflect encyclopaedic-like knowledge, perhaps acquired linguistically, many features are learned by directly experiencing the referents of these concepts through the senses. For example, one can see that a dog has four legs, hear that it barks, or feel that it is covered in coarse fur. Many features are thus strongly associated with particular senses. Feature production norms, therefore, provide insight into the salience and amount of knowledge that people possess for each sensory modality with respect to individual concepts, categories, or in general.

Multimodal versus Amodal Representations

It has long been known that different brain regions are specialized for perception across different sensory modalities. The question of representational modality concerns the extent to which conceptual organization is tied to perceptual organization; that is, given that perception across the senses is distributed, at least in part, across specialized brain regions, it is possible that people's conceptual representations are similarly organized. One way of contrasting feature-based representational models is to partition them into *amodal* and *multimodal* theories of representation. Although various amodal theories make different assumptions with respect to *what* is stored in semantic memory, all assume that objects or their properties are represented in a single homogenous store. For these models, the sensory modality through which knowledge is gained is irrelevant to the representation of that knowledge because this information is lost when it is stored using mental symbol systems. In contrast, multimodal theories posit that concepts are distributed across a wide network of brain areas, and that a concept's features are tied to sensory modalities.

The issue of whether the human representational system is multimodal or amodal is still under debate. However the bulk of recent evidence from a number of lines of research favours the multimodal account. The literature regarding patients with category-specific semantic deficits has long been used to support the argument for multimodal representations. Warrington and McCarthy's (1987) sensory/functional hypothesis accounts for patterns of category specific impairments of knowledge in patients that have suffered focal or diffuse brain damage, under the assumption that living things and artifacts differentially depend on visual and functional information – an assumption that is supported by analyses of feature production norms (Cree & McRae, 2003; Garrard, Lambon Ralph, Patterson, & Hodges, 2001), and by a number of functional neuroimaging experiments (see Martin, 2007, for a review).

Citing the prevailing opinion that the N400 ERP waveform reflects the brain's processing of meaning, Sitnikova, West, Kuperberg, and Holcomb (2006) investigated the topographical distribution of the N400 for animals and tools in a picture naming task. Though it only provides coarse spatial resolution, the temporal resolution of the technique allowed the authors to discriminate between multiple late onset sub-components at anterior and posterior sites. Animals and tools differed with respect to the distribution of the ERP waveforms, with animals more negative than tools at anterior-inferior sites, and the opposite pattern at posterior sites. This followed predictions based on previous literature that suggested that if categories such as animals depend more on a visual feature semantic system, they might be expected to evoke greater anterior N400 effects than tools. Similarly, if tools depend more on functional representations, such as

that which might be encoded by motor sequences, they would be expected to evoke relatively greater posterior activity.

The imaging literature has also provided a wealth of evidence outside of the context of sensory/functional theory that supports a distributed multimodal representational system. Goldberg, Perfetti and Schneider (2006a) used fMRI to tie together a number of previous findings that were suggestive of a multimodal semantic system. Several researchers had previously reported neuroimaging evidence individually supporting modally bound colour (Martin, Haxby, Lalonde, Wiggs, et al., 1995; Mummery, Patterson, Hodges, & Price, 1998), auditory (Kellenbach, Brett, & Patterson, 2001), olfactory and gustatory representations (Goldberg, Perfetti, & Schneider, 2006b). Goldberg et al. (2006a) found that sensory brain areas for each of the four tested modalities were recruited during a feature verification task for features of the corresponding modality, indicating that the semantic representations used in these decisions were modally distributed across brain regions. Taken together, a number of complimentary techniques have provided converging evidence that supports a distributed multimodal representational system.

Convergence Zones

Though concepts may be distributed across a wide network of brain regions, our mental experiences of them are not a jumble of features, disjointed across space and time, but instead appear to be represented as coherent unified objects. Any model using distributed feature representations must account for what is sometimes called the binding problem: How are representational elements integrated into conceptual wholes? Similarly, how are we able to infer one feature from the presence of another, such as the

likelihood that something flies if it has feathers? If one makes the additional assumption that semantic representations are modally-distributed, the binding problem becomes further complicated because it raises the question of whether within-modal binding is accomplished differently than cross-modal binding, or differs by modality.

Understanding how distributed representations are integrated into conceptual wholes is therefore of central importance to generating and evaluating semantic memory models and understanding brain function.

A number of multimodal theories of semantic representations have been proposed in the last two decades, and each makes slightly different assumptions about the modalities that are represented and the relationships among them. These models can be broadly grouped into two classes, deep and shallow, on the basis of the assumed hierarchy of convergence zones, which are neural sites that integrate information (Damasio, 1989). The differences in assumed connectivity between these two classes of models lead to different and untested predictions for how modally distributed information is integrated. Thus, tasks that should be sensitive to the time course of integration of featural information either across or within modality boundaries may shed some light on which assumptions are most likely to be accurate, and thus constrain models of semantic representation.

Convergence Zone Hierarchies

Hierarchically Shallow Models. In this dissertation, I divide modality-specific models into two classes. I use the term “hierarchically shallow models” to refer to those in which modally segregated representational stores pass information to one another

either through direct connections (Figure 1), or through a single convergence zone that integrates and routs information from all representational modalities (Figure 2).

In these models, all units of representation are equidistant from one another. Examples of semantic models using direct connectivity include Farah and McClelland's (1991) implementation of Warrington and McCarthy's (1987) Sensory/Functional theory, depicted in Figure 1, and the attractor network used in Cree, McNorgan, and McRae's (2006) investigation of the roles played by distinguishing and shared features. Examples of models employing a single convergence zone include the attractor network described in Cree et al's (1999) simulation of semantic priming effects and Humphreys and Forde's (2001) Hierarchical Interactive Theory (HIT). In a recent review of the patient literature describing the pattern of deficits and brain atrophy associated with semantic dementia, Patterson, Nestor, and Rogers (2007) also presented a model with a single semantic convergence zone.

Hierarchically Deep Models. Hierarchically deep models are those for which connective distance differs. In these models, some convergence zones integrate information from nearby representational units, whereas others with successively larger receptive fields integrate information from more distant brain areas, passed forward from earlier convergence zones (Figure 3). Damasio's (1989) convergence zone proposal envisioned such an organization, and Simmons and Barsalou (2003) further elaborated on this idea in their Conceptual Topography Theory.

Amodal Models. Amodal models span a range of structural hierarchies. In amodal models, because information is not functionally segregated by sensorimotor modality, convergence zones are not strictly required. Hierarchically flat models encode

Figure 1. A hierarchically shallow model with two directly interconnected modalities.

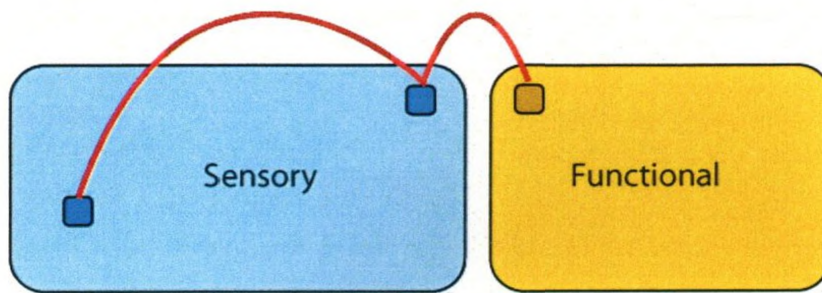


Figure 2. A hierarchically shallow model with two modalities connected via a single convergence zone.

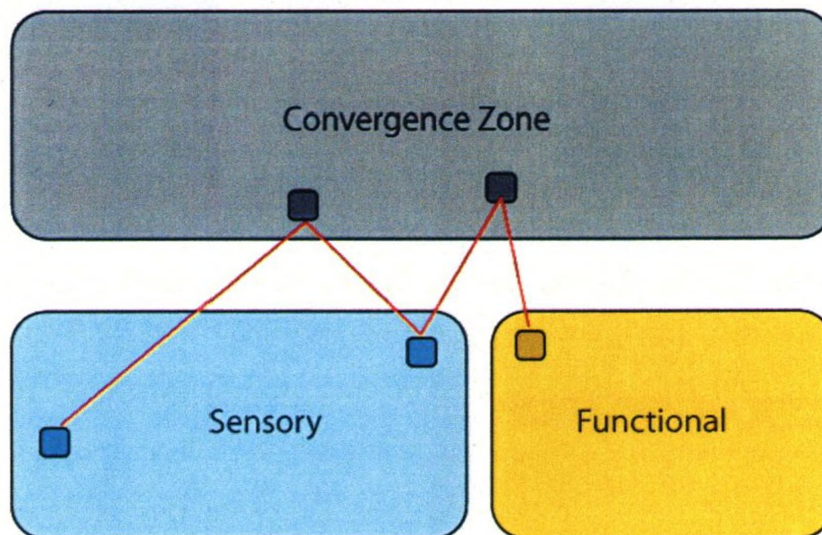
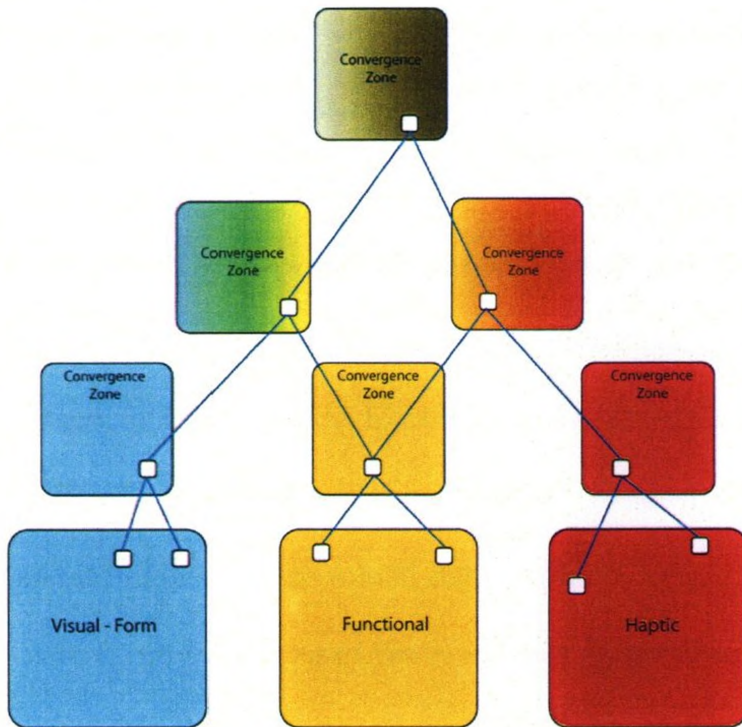


Figure 3. A hierarchically deep model containing three modalities, each possessing a unimodal convergence zone that feeds forward to bimodal and trimodal convergence zones.



associations among an object's features via direct connections reinforced through statistical learning (Moss & Tyler, 2001). Neither, however, does the absence of modality information in these models preclude convergence zones. For example, in Collins and Quillian's (1969) semantic network model, concept names encapsulate features of the concepts they represent, and could therefore serve to integrating featural information. Moreover, concepts in this model are themselves organized in a deep hierarchical taxonomy (a sparrow is a bird, which is an animal, which is a living thing). It is important, however, to distinguish multimodal semantic models that are the focus of this dissertation from amodal hierarchical models. In the former, sensorimotor modality is the primary organizing principle, whereas in the latter, this information is not represented, and therefore plays no role in cognitive processing.

Arguments for Deep and Shallow Convergence Zone Hierarchies

Theoretical considerations seem to favour a shallow integration hierarchy. Multimodal semantic models have been criticised by those arguing for a single amodal semantic system as lacking parsimony (Riddoch et al., 1998), therefore models specifying multiple hierarchical convergence zones would seem to be even more so. Furthermore, many semantic phenomena have been simulated using networks lacking convergence zones (Cree et al., 2006; Farah & McClelland, 1991), implying that a deep hierarchy of convergence zones may not be necessary. Indeed, Patterson et al. (2007) contend that a single semantic hub best explains the generalized impairments that accompany semantic dementia.

There are, however, anatomical constraints that seem to suggest a hierarchically deep organization. First, candidate brain regions for a single convergence zone multimodal

model should have reciprocal projections to all modalities, and ablation of such an area should preclude any sort of multimodal conceptualization. Damasio (1989) argues that the only such region is the hippocampus, and because bilateral ablation of this structure does not lead to a catastrophic loss of the ability to conceptualize, it is unlikely that semantic integration occurs within a single convergence zone. On the other hand, one could argue that the sort of generalized impairments that accompany semantic dementia constitute a progressive breakdown of the conceptual system. Because this disease is invariably accompanied by degeneration of anterior temporal lobes, Patterson et al. (2007) suggested this region as the locus for a single semantic hub. Finally, the arrangement of cells, such as in visual cortex, into functionally distinct layers with progressively larger receptive fields that respond to a greater array of stimuli may occur elsewhere in the brain, including those supporting conceptual processing, and would implement the sort of deep hierarchy of convergence zones suggested by Damasio (1989) and Simmons and Barsalou (2003).

The preceding discussion highlights a number of arguments favouring each of the two major assumptions regarding hierarchical organization, both of which being incorporated in models that have been used to explain a number of behavioural phenomena. Moreover, there is little empirical support for either assumption in the literature. A number of brain regions, including perirhinal cortex (Bussey, Saksida & Murray, 2002), anterior temporal cortex (Patterson et al., 2007), frontal and prefrontal cortex (Fuster et al., 2000; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006), and left inferotemporal cortex (Damasio, 2004) have been put forward in the neuroimaging, animal, and patient literatures as critical structures for learning abstract relationships

among features from multiple modalities. However, whether these areas represent a network of regions that act as a single convergence zone in a shallow hierarchy or convergence zones at different levels of a deep hierarchy is unclear. The behavioural literature is similarly ambiguous. Pecher, Zeelenberg and Barsalou (2003) contrasted feature verification latencies for consecutive trials involving features from either the same or different modalities and found a modality switching penalty. As is explained below, however, these results do not adjudicate between deep and shallow integration hierarchies. The goal of this dissertation, therefore, is to provide the first direct test of the predictions made by the two classes of models.

Proximity-Sensitive Integration

The physical relationships among modality-specific representational areas and their convergence zones are assumed to influence the time course over which, and extent to which, information from these areas is integrated. From a neural proximity perspective, proximal areas should generally communicate with one another in less time than distal areas. One way to view this information relay is in terms of processing steps, where each step represents the amount of time required to pass information from one processing unit to another. The assumed integration hierarchy within a model therefore influences the time course of semantic processing, and moreover this influence may differ by task.

The capacity to integrate featural knowledge is expected to facilitate performance in timed tasks in the following manner. The imaging literature (Simmons, Martin & Barsalou, 2005; Simmons, Ramjee, Beauchamp, McRae, Martin, & Barsalou, 2007) has demonstrated that, given a single written feature (<*has a blade*>), the underlying neural representation for this feature becomes active. For example, the same brain regions that

respond strongly to visually presented colour stimuli are also activated when reading colour words, such as YELLOW (Simmons et al., 2007). This activity spreads outwards via neural connections to begin to activate other representations, which in turn spread their activation outwards, and so forth. In this way, if a convergence zone forms a path between representational stores, then verbally presented features can effectively prime subsequently presented ones, either individually as in feature-to-feature inferences (inferring that a gold brick would feel heavy if it was picked up), or entire clusters of features as in feature-to-concept activation (classifying a small animal as a skunk on the basis of its size, colouration, and waddling gait). The manner in which features are integrated should thus be reflected in the speed with which decisions based on this integration are made, and therefore influence decision latencies in speeded tasks. Two such tasks are described below, and understanding how performance is influenced by neural organization provides insight into how they can be used to test how the brain represents and uses multimodal information.

Feature Inference. Feature inference involves determining whether some feature B exists for an object, given that the object is known to possess feature A, and may be accomplished independently of object categorization. For example, given that a novel artifact <has a blade>, one might infer that it is also <used for cutting>, or that it <has a handle>, without any other supporting evidence. To make this inference, the activation of knowledge of the visual form of a blade must be propagated along neural circuitry to cause the activation of one's knowledge of handles or functional knowledge of the action of cutting.

Figure 4 is a simplified shallow model with a single convergence zone and two representational modalities. Integrating two pieces of information from either a single modality (integrating two form features, such as *<has a blade>* and *<has a handle>*, Figure 4a) or cross-modally (integrating a form and a functional feature, such as *<has a blade>* and *<used for cutting>*, Figure 4b) requires two steps: one step between each representational unit and the convergence zone. Therefore, in timed tasks that are sensitive to the number of processing steps, this model predicts no difference between within-modal and cross-modal semantic integration. Naturally, the speed with which this inference is made could reflect differences in the degree to which features are related, perhaps owing to a more efficient or dedicated neural pathway between the structures that encode this information. However, it should take no longer to infer either *<has a handle>* or *<used for cutting>* from *<has a blade>*, all else being equal.

Figure 5 is a simplified hierarchically deep model incorporating both within-modal and cross-modal convergence zones. Within-modal integration (Figure 5a) requires two steps: one step between each representational unit and the unimodal convergence zone. This architecture predicts that cross-modal integration (Figure 5b) should take more time, as four processing steps are required: one from each representational unit to their corresponding unimodal convergence zone, and one from each unimodal convergence zone to the cross-modal convergence zone. Therefore, assuming that additional processing steps require additional time, it should take longer to infer the cross-modal functional feature *<used for cutting>* from the form feature *<has a blade>*, than it would to infer the within-modal form feature *<has a handle>*, all else being equal.

Figure 4. Within- and cross-modal feature-feature activation in a shallow hierarchy takes the same number of processing steps.

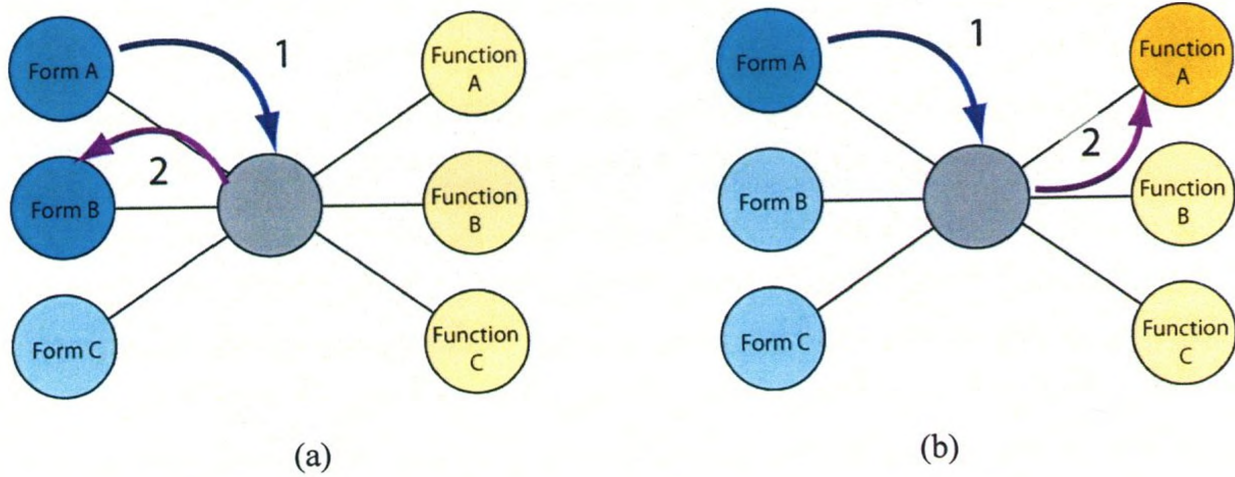
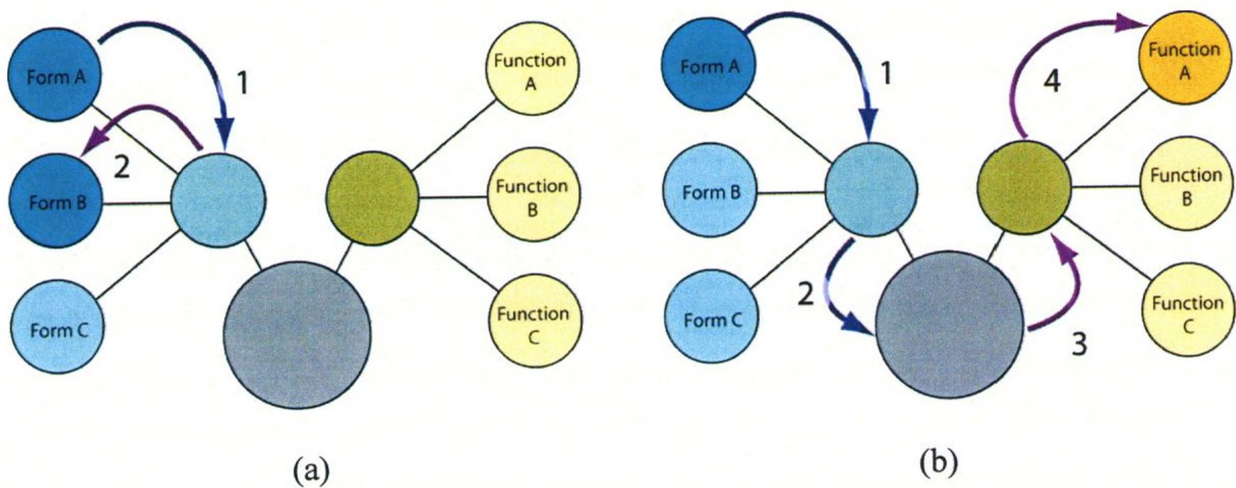


Figure 5. Feature-feature activation in a deep hierarchy takes fewer processing steps within-modally than cross-modally.



In amodal models, because features are not segregated according to sensorimotor knowledge type, there is no within- and cross-modal distinction, and therefore no predicted within- versus cross-modal advantage, all other factors being equal.

Activating Concepts from Features. For the purposes of this discussion, concept activation can be thought of in terms of pattern completion. In a system using distributed representations, an object's identity and relevant features can be deduced from a subset of features through pattern completion. For example, if one determines that an object *<has a blade>* and *<has a handle>*, a number of other features may be inferred; the blade of the object is probably *<made of metal>* and *<used for cutting>*. Over time, the pattern of activated features comes to resemble the representation for some class of objects, allowing the object to be categorized. In this example, one may coarsely identify the object as some sort of manipulable tool, and with additional processing time, identify it more precisely as some sort of knife.

Figure 6 illustrates a simplified hierarchically shallow model of a concept, consisting of just six features distributed across two modalities. In this example, I assume that concept identification requires that some proportion of the concept's features reach some activation threshold. Suppose in the first case that two features from the same modality are presented, as when a person is told that an object *<has a handle>* and *<has a blade>* (two form features, Form A and Form B, as in Figure 6a). Activation from the two feature units spreads in the first processing step to the common convergence zone, and to all associated features in the second step. All of the concept's features are soon active, *Figure 6.* Feature-concept activation in a shallow hierarchy takes the same number of processing steps from within-modal pairs as from cross-modal pairs.

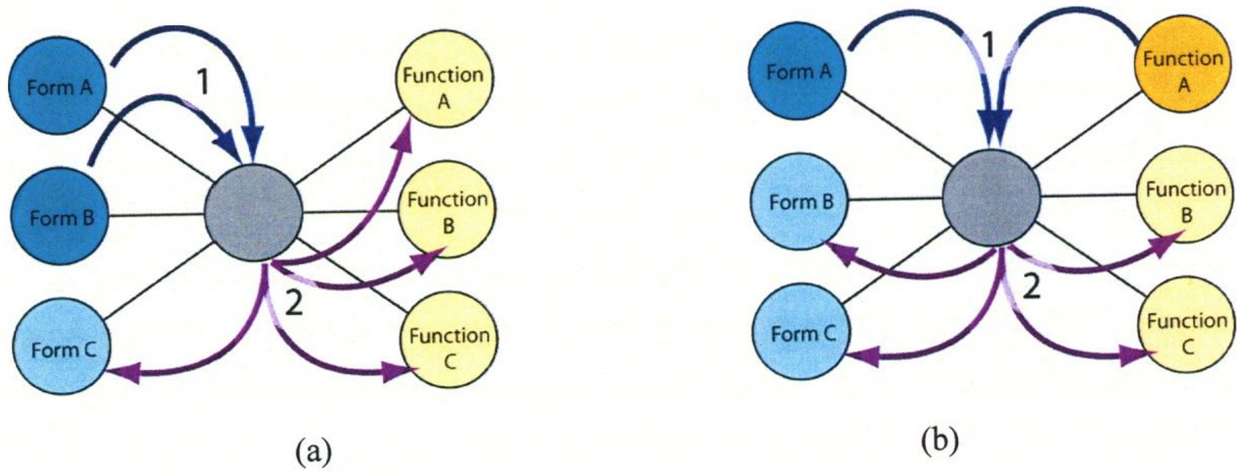
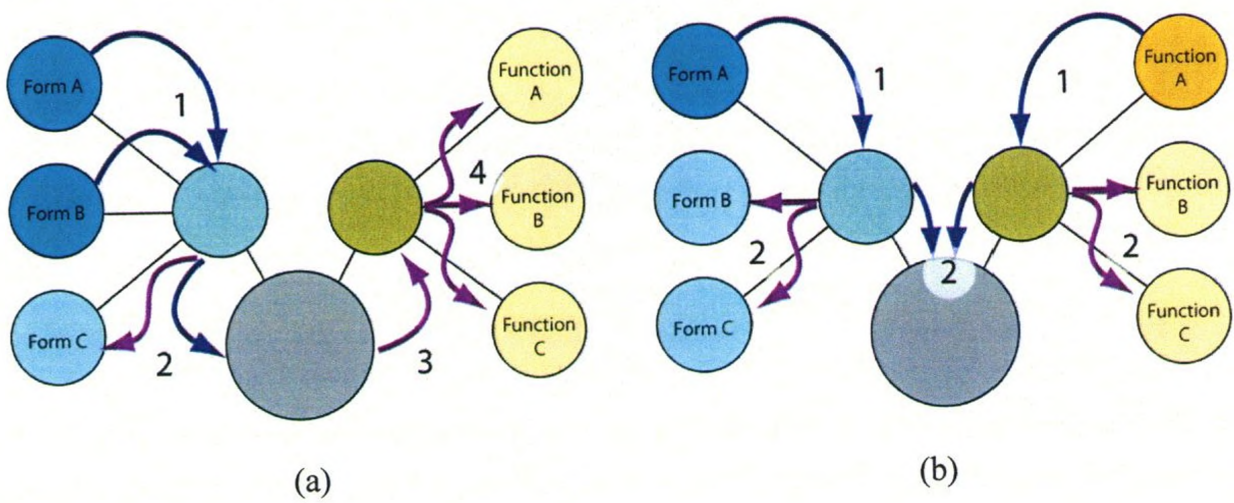


Figure 7. Feature-concept activation in a deep hierarchy takes more processing steps from within-modal pairs than from cross-modal pairs.



allowing the object to be identified. The cross-modal case, illustrated in Figure 6b, is similar. This might exemplify a situation in which a person is told that an object <has a blade> (form) and is <used for cutting> (function). Again, two processing steps are required to activate the remaining features in a shallow model. Activation of the two feature units spreads to the common convergence zone in the first step, and then to all associated features in the second. Thus, regardless of whether the presented information corresponds to a single or multiple modalities, shallow hierarchy models predict the same number of processing steps, and presumably the same amount of time, to activate a concept.

Figure 7 illustrates a simplified hierarchically deep model of a concept containing the same six representational units. Again, in the first scenario, two features from the same modality are presented (two form features, Form A and Form B, as in Figure 7a). After one processing step, activation has spread to the modality-specific convergence unit, and after two steps, it has begun to activate the remaining within-modal feature, Form C, and the top-level cross-modal convergence unit. It requires two additional processing steps, as activation passes first to the Function convergence unit, and then to the three function features, Function A, B and C, before all of the concept's features are activated.

Consider now the case illustrated in Figure 7b, in which the model is given a cross-modal pair of features (Form A and Function A). After one processing step, activation has spread from both features to their respective modality-specific convergence unit. After the second step, both within-modal convergence units pass activation forward to the cross-modal convergence unit, but more importantly, each within-modal convergence

unit has an opportunity to pass activation backward from the function convergence unit to other correlated functional units (i.e., Function B and C), and similarly from the form convergence unit to other correlated form units (i.e., Form B and C). Thus, in the hierarchically deep model, cross-modal feature pairs activate more of the network faster than do within-modal feature pairs because each within-modal convergence area allows the parallel activation of clusters of correlated features within its modality.

Finally, as above, lacking any modality information, amodal models predict no within- versus cross-modal speed advantage, all other factors being equal.

The preceding discussion illustrates how theories that exemplify the two classes can be differentiated by the predictions they make about performance on tasks requiring fast integration of semantic information. For time-sensitive feature inference tasks, amodal and hierarchically shallow models predict no difference between within-modal and cross-modal inferences. Hierarchically deep models, in contrast, predict a latency advantage for within-modal decisions because additional steps are required to pass information between modalities. For time-sensitive tasks requiring activation of a concept's meaning, amodal and hierarchically shallow models again predict no difference in the ability of within- or cross-modal feature pairs to facilitate pattern completion. In contrast hierarchically deep models predict a cross-modal advantage for pattern completion. One advantageous outcome of these tasks is that deep models predict opposite patterns of results. That is, within-modal feature inference is predicted to be faster than cross-modal feature inference, whereas cross-modal pattern completion is predicted to be faster than within-modal pattern completion.

It is important to reiterate that these predictions hold for tasks in which speed is an important factor (i.e., speeded tasks) because these effects are most likely to be observed during early semantic processing. Because the processing of even the slowest items is assumed to occur relatively quickly, differences related to the speed of processing should quickly drop off with additional processing time. Moreover, untimed tasks such as those that require participants to rate items on some dimension allow participants to engage in higher level reasoning and to use information from a number of sources that are assumed to be less sensitive to the temporal dynamics of semantic processing (McRae et al., 1997; Sloman, Love, & Ahn, 1998).

The present research uses an analysis of a large set of feature production norms (Study 1), along with predictions derived from a connectionist framework to guide the design of two sets of studies involving feature-to-feature (Studies 2 and 3) and feature-to-concept judgments (Study 4) to investigate and constrain assumptions made by neurally-based models of semantic representations. Results from these behavioural studies, in the context of what is known from the imaging and patient literatures, are used to make inferences about functional connectivity within the representational system in the brain. In Study 1, I set the stage by analyzing a large set of feature production norms to establish multiple statistics regarding the numbers of features of various modalities, and the statistical relationships among them. As outlined above, hierarchically shallow and deep models make different predictions for within- and cross-modal processing in the context of feature inference and conceptualization.

Chapter Two: Functional Connectivity and Feature Norms Analysis

Study 1: Feature Production Norms Analysis

The predictions for both neural architectures were derived assuming all other factors being equated. It is likely, however, that all pathways of communication between representational modalities are not equal. For example, the orbitofrontal cortex, the area of the brain most commonly associated in the literature with taste and smell processing, is located on the ventral anterior surface of the brain, whereas primary visual areas are located in the occipital-parietal cortex on the posterior surface (both dorsal and ventral). The brain region encoding taste is thus topographically distant from that which encodes visual features. From a neural proximity perspective, it is reasonable to imagine that cross-modal feature integration of gustatory and olfactory information occurs more rapidly and strongly than the integration of gustatory and visual information. On the other hand, there may be well-worn pathways that connect even distant processing areas that permit faster communication between these regions than between other, more proximal areas.

The purpose of Study 1 was to gain a better understanding of how various modality specific representational stores may be related. Our awareness that features, such as *<has wings>* and *<has feathers>* often co-occur in the world is captured by our knowledge of feature correlations. It is assumed that this knowledge is encoded through a process such as hebbian learning, by which the pathways between neural populations representing correlated features are strengthened over time. In the brain connectivity literature, the term *functional connectivity* is used to refer to a pattern of time-dependent correlated activity between distributed and potentially remote neural units (Sporns & Tononi, 2007).

Thus, the degree to which two features are statistically correlated should be positively correlated with the functional connectivity between the neural populations encoding them. Study 1 therefore used three aggregate measures of feature correlations occurring in a large set of feature norms under the assumption that these metrics provide an estimate of the functional connectivity between the sensorimotor modalities appearing in these norms. These measures informed the design of Studies 2, 3 and 4.

Materials and Procedure

Features were drawn from McRae et al.'s (2005) feature production norms (henceforth, "our norms"), which consist of over 2500 features produced for 541 concrete object concepts. In the norming task, participants listed features when given the names of approximately 20 dissimilar concepts (30 participants listed features for each concept). Aggregated across a number of respondents, the norms provide insight into what information is salient for most people for different concepts. Because the present investigation concerns communication between modality specific brain regions, these norms are useful because they have been categorized by Cree and McRae (2003) into ten knowledge types that are linked to modality-specific neural processing regions. Thus, for example, some authors have associated retrieval of object colour knowledge (as opposed to retrieval of other sorts of visual knowledge) with ventral regions of the posterior temporal lobes (Chao & Martin, 1999; Martin et al., 1995), and so features that describe an object's colour were assigned the visual-colour knowledge type. Three knowledge types corresponded to visual information (visual-color, visual-parts and surface features, and visual-motion), four corresponded to other primary sensory-processing channels (smell, sound, tactile, and taste), and one corresponded to functional/motor information

regarding the ways in which people interact with objects (function). In addition, one (taxonomic) corresponded to information regarding the categories to which concepts belong, and one (encyclopaedic) corresponded to other types of information that either could not be confidently mapped to particular neural processing regions given the state of the literature, or else did not form a coherent group of sufficient size to be of practical use in their analyses (see Cree & McRae, 2003, for a detailed discussion of the considerations underlying selecting and assigning these knowledge types). Because no particular brain regions are associated with taxonomic and encyclopaedic knowledge, these knowledge types were excluded from all analyses.

Feature correlations are a useful tool for investigating information integration because a correlation describes the result of feature integration, which involves co-occurrence detection. Learning associations between features is assumed to be a function of experience with objects containing both features. If people frequently encounter objects that both *<have wings>* and *<have a beak>*, then there will be many occasions on which these features are integrated, strengthening the learned relationship between them. Highly correlated features are those that tend to be listed together in the same concepts, and are assumed to be frequently experienced together and therefore be more strongly learned. Therefore, the shared variance (i.e., the squared correlation) between two features in the norms is taken as an index of the strength with which two features are associated in the mind. For example, within the norms, the pair *<has legs>* and *<eaten as meat>* has 5% shared variance. Neither feature intuitively implies (nor precludes) the other, and such a pair is likely to be weakly related in most people's minds. In contrast, because our sense of taste relies so heavily on the sense of smell, the pair *<smells*

strong> and <*tastes strong*> is almost tautological. The pair has 99.95% shared variance within the production norms, and would almost certainly be related in most people's minds.

Feature correlations were calculated between all features appearing in our norms. The entire set of norms included all features that were provided by at least 5 of 30 participants for at least one concept. A 541 concepts \times 2,526 features matrix was then constructed, where each matrix element corresponded to the number of participants listing a specific feature for a specific concept. Thus, each feature was represented by a 541-element vector so that a Pearson correlation could be computed between each feature pair. For the present study, spurious correlations were avoided by considering only pairs involving the 340 features that were listed for more than 3 concepts. The Pearson correlation between each feature pair was squared to obtain shared variance between features. Consistent with other research involving these production norms (McRae et al., 1997, 1999), the threshold for inclusion in the analyses presented in this section was arbitrarily set at a shared variance of 5%, which serves as an estimate of the minimum degree of statistical relatedness required to become a psychologically real association.

Just as the statistical correlations between features is assumed to approximate the strengths of the learned association between them, an analysis of the broader relationships among features from different modalities may provide insight into the overall functional connectivity between the brain regions that encode this information and the frequency with which information between any two modalities would be integrated. I classified the set of correlated feature pairs into 64 modality-pair types: 8 in which within-modal pairs were from the same knowledge type (e.g., pairs involving two

visual-colour features) and 56 types of cross-modal pairs in which features came from different knowledge types (e.g., pairs involving a visual-colour and a visual-form feature). Pair types with fewer than 6 members were omitted from the analyses that follow for reliability reasons. Numerosity, magnitude of shared variance, and a combination of these two measures were then calculated for each pair type.

Results and Discussion

Table 1 presents the number and proportion of the total number of each type of within-modal and cross-modal correlated feature pairs. Because people generally rely primarily on vision for information about the environment, a large proportion of the entries in the norms describe visual features. There was therefore a concomitant dominance, with respect to numerosity, of within- and cross-modal pairs involving visual features, especially those describing an object's form (*<has legs>*).

Table 2 presents the magnitude of shared variance for each type of within-modal and cross-modal correlated feature pair. Again, visual features tend to dominate, though for this measure, colour (*<is red>*) and motion (*<flies>*) features tend to have the strongest, if not the most numerous, statistical relationships among themselves and with other modalities. Colour, for example, is highly diagnostic of the taste and texture of many fruits and vegetables because it often indicates ripeness, which explains the strong correlations between colour and taste features. For similar reasons, the relatively strong correlations between motion and sound features may arise from features listed for animals, which are often saliently associated with a particular manner of motion (*<flies>*, *<slithers>*, *<runs>*) and characteristic sounds (*<chirps>*, *<hisses>*, *<roars>*).

Table 1

Numerosity and proportion of total correlations by correlated pair type

Pair Type	N	Proportion
Within-Modal		
Visual - Form & Surface / Visual - Form & Surface	314	.16
Function / Function	277	.14
Visual - Motion / Visual - Motion	44	.02
Touch / Touch	12	.01
Visual - Colour / Visual - Colour	9	.00
Sound / Sound	8	.00
Taste / Taste	6	.00
Cross-Modal		
Visual - Form & Surface / Function	611	.32
Visual - Form & Surface / Visual - Motion	132	.07
Visual - Form & Surface / Touch	77	.04
Visual - Colour / Visual - Form & Surface	67	.03
Touch / Function	66	.03
Visual - Colour / Function	57	.03
Visual - Form & Surface / Sound	35	.02
Sound / Function	32	.02
Visual - Motion / Function	30	.02
Visual - Form & Surface / Taste	26	.01
Touch / Function	21	.01
Visual - Motion / Sound	20	.01
Visual - Colour / Visual - Motion	19	.01
Visual - Colour / Touch	13	.01
Visual - Colour / Taste	13	.01
Visual - Colour / Sound	10	.01
Visual - Motion / Touch	9	.00

Table 2

Mean magnitude of shared variance by correlated pair type

Pair Type	Mean R^2	SE
Within-Modal		
Visual - Colour / Visual - Colour	39	34
Visual - Motion / Visual - Motion	34	19
Touch / Touch	32	25
Function / Function	20	17
Sound / Sound	17	8
Visual - Form & Surface / Visual - Form & Surface	15	13
Taste / Taste	12	5
Cross-Modal		
Visual - Colour / Taste	53	43
Visual - Motion / Sound	29	33
Visual - Colour / Visual - Form & Surface	26	22
Visual - Motion / Touch	23	27
Visual - Motion / Function	23	19
Visual - Form & Surface / Visual - Motion	22	20
Touch / Function	22	21
Visual - Form & Surface / Function	22	21
Visual - Colour / Visual - Motion	22	18
Touch / Function	20	14
Visual - Colour / Function	17	12
Sound / Function	17	19
Visual - Form & Surface / Touch	17	16
Visual - Colour / Sound	16	10
Visual - Form & Surface / Taste	16	15
Visual - Form & Surface / Sound	15	12
Visual - Colour / Touch	15	15

Finally, Table 3 presents the summed shared variance for each type of within- and cross-modal correlated feature pair, which summarizes the interaction between the previous two measures: pairs of modalities that have either many correlated pairs, or strong correlations, or both, have a higher summed shared variance. Thus, because neural functional connectivity between two modalities can be imagined as a function of both the number and strength of the pathways between them, it may best be estimated by this measure. Again, pairs involving visual form features dominate, but the largest proportion of estimated functional connectivity is between visual form and function feature pairs, which is consistent with the suggestion put forth by some researchers that form-function knowledge is particularly important in the representation of object concepts (Tyler & Moss, 2001).

These analyses suggest that functional connectivity may differ between modality-specific representational areas. According to these estimates, visual areas are highly interconnected with most other modalities, and among themselves. There would therefore be the greatest need to integrate visual information with other visual information, and with information from other modalities. Interestingly, the greatest estimated functional connectivity does not occur within a single visual modality or even between two visual modalities as might be predicted. Instead, the highest estimated functional connectivity occurs cross-modally between the visual form and function modalities, though the within-modal functional connectivity for the functional and the visual form modalities have the second and third highest estimated functional connectivity, respectively. Because of the apparent importance of, and strong relationship between, visual form and functional knowledge, and because a rich literature exists examining categories that differentially

depend on these knowledge types (Martin, 2007; Warrington & McCarthy, 1987), I began the investigation of the nature of cross-modal and within modal semantic integration with these modalities. Additional implications of these analyses are presented in the General Discussion.

Table 3

Summed shared variance by correlated pair type

Pair Type	Summed R^2
Within-Modal	
Function / Function	5533
Visual - Form & Surface / Visual - Form & Surface	4647
Visual - Motion / Visual - Motion	1499
Touch / Touch	384
Visual - Colour / Visual - Colour	355
Sound / Sound	132
Taste / Taste	69
Cross-Modal	
Visual - Form & Surface / Function	13640
Visual - Form & Surface / Visual - Motion	2951
Visual - Colour / Visual - Form & Surface	1719
Touch / Function	1313
Visual - Form & Surface / Touch	1281
Visual - Colour / Function	978
Visual - Colour / Taste	693
Visual - Motion / Function	682
Visual - Motion / Sound	571
Sound / Function	547
Visual - Form & Surface / Sound	533
Touch / Function	469
Visual - Colour / Visual - Motion	417
Visual - Form & Surface / Taste	411
Visual - Motion / Touch	207
Visual - Colour / Touch	192
Visual - Colour / Sound	159

Chapter Three: Feature Integration in Feature Inference

Study 2

Study 2 used a speeded relatedness decision task between within- and cross-modal feature pairs to test the role that modality plays in feature inference. The speeded nature of this study makes this task sensitive to factors that influence processing time.

Relatedness decisions were selected because, as has been argued elsewhere, they are a fairly transparent measure of people's knowledge of the relations between object features, and should therefore tap the sort of processes that are used during feature inference (McNorgan et al., 2007). Both amodal and hierarchically shallow models predict no difference in decision latencies between the two types of correlated pairs on the basis of the time taken to integrate the features. Hierarchically deep models, in contrast, predict a speed advantage for form-form pairs because modal representations for form and function features are processed by modality-specific convergence zones before that information is passed forward to cross-modal convergence zones.

Method

Participants. Twenty-one University of Western Ontario undergraduates received \$10 for participating in the speeded relatedness task. All participants in all of the studies reported herein were native English speakers and had either normal or corrected-to-normal visual acuity.

Materials. Twenty form-form and 20 function-form feature pairs were selected from our norms (see Appendix A). Because the task measures the time required to judge perceived feature relatedness, the groups were matched on a number of variables expected to influence both reading time and perceived relatedness (see Table 4).

Table 4

Equated variables in Study 2

Factor	<i>Function-form</i>	<i>Form-Form</i>	<i>t</i> (38)	<i>p</i>
	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)		
Mean % Shared Variance	18.5 (3.0)	18.7 (4.3)	-0.04	0.9
Number of Shared Concepts	5.5 (0.7)	4.9 (0.8)	0.63	0.5
Length in characters (1 st feature)	17.7 (0.4)	11.4 (0.7)	7.32	< .001
No. Concepts Unique to Feature (1 st feature)	10.3 (3.4)	11.5 (3.8)	-0.24	0.8
Concepts per feature (1 st feature)	15.8 (3.7)	16.3 (4.0)	-0.1	0.9
Length in characters (2 nd feature)	13.7 (.6)	13.0 (0.7)	0.57	0.5
Number of Unique (2 nd feature)	10.9 (3.0)	11.5 (3.2)	-0.14	0.8
No. Concepts Unique to Feature (2 nd feature)	16.4 (3.2)	16.3 (3.7)	0.01	0.9

The mean percentage of shared variance is the mean of the squared Pearson correlations between the feature production frequency vectors created from our norms. The number of shared concepts is the number of concepts in which both features appear together, and reflects the ease of generating a concept possessing both features. This variable was matched in addition to shared variance because shared variance was calculated on feature production values, thus a pair of features may appear together in many concepts but have a lower proportion of shared variance than another pair appearing in fewer concepts. The length in characters of the first and second feature indicates the number of characters, including spaces, of the feature name, and is assumed to influence reading time. The number of unique concepts listed for the first and second features is the number of concepts within our norms in which the one feature appears without the other. This indexes the likelihood that, if a feature was to prompt the retrieval of a concept, the participant would fail to note a co-occurrence between the feature pair. Finally, the number of concepts per feature is the number of concepts appearing in our norms for which the feature was listed by at least five participants.

There were no differences between form-form and function-form pairs on eight of the ten variables. In addition, each group had similar distributions of higher ($r^2 > .25$), medium ($.25 \leq r^2 \leq .15$) and lower ($r^2 < .15$) correlated pairs, such that the cross-modal group had 5, 3, and 12 and the within-modal group had 4, 2 and 14 high, medium and lower correlated pairs, respectively. Because the task was intended to measure judgments of relatedness between features without reference to any particular concepts, I avoided distinguishing features (i.e., features that are true of one or two concepts, such as *<moos>*). Rather, within- and cross-modal pairs included features that were true of an

average of approximately 16 concepts, and the number of concepts per feature did not differ between conditions. Note that the shorter length of the first feature of form-form pairs was not a critical issue for two reasons. First, decision latencies were measured from the onset of the presentation of the second feature. Second, the 500 ms SOA was expected provide participants ample time to read the first feature, regardless of length, so the longer feature names for function features were not expected to be a problem (see Cree et al., 2006). Note that equating for word length of the second feature ensured that there was no effect of reading time on decision latency.

Relatedness Ratings. Twenty-two participants not participating in the on-line task produced off-line relatedness ratings to ensure that differences in decision latencies were not attributable to differences in perceived relatedness when time pressure was not an issue. Participants rated the relatedness of each pair (i.e., judged how well each pair “goes together in common living and/or non-living things”) on a 7-point scale, ranging from 1 (“not at all related”) to 7 (“very highly related”). Note that hierarchical depth is predicted to influence the speed with which semantic information is retrieved, but with additional time, more slowly accessed information is expected to become available regardless of the underlying architecture. Therefore, a task such as feature relatedness rating in which participants take their time to decide on a particular rating value should not be sensitive to any influence of hierarchical depth. Because there was no time limit, this task was not expected to be sensitive to hierarchical depth, and accordingly, no predictions are made for shallow versus deep hierarchy models for the offline relatedness rating task.

Comparisons between relatedness ratings were conducted using modality (within vs. cross) as the independent variable. Modality was within participants (t_1) but between

items (t_2). Cross-modal pairs were judged to be more related ($M = 5.7, SE = 0.23$) than within-modal pairs ($M = 5.4, SE = 0.20$), though this difference was significant by participants, $t_1(21) = 3.62, p < .002$, but not by items, $t_2(38) = 0.71, p < .5$. This difference should lead, if anything, to the facilitation of decision latencies of cross-modal pairs.

Filler items. Because this was a binary decision task, an equal number of ‘yes’ and ‘no’ trials were used to avoid biasing the response. Thus, in addition to the 40 related pairs, I constructed 20 form-form and 20 function-form pairs that could not be construed as co-occurring in common objects. For example, *<has branches>* and *<has sharp fangs>* was used as an unrelated form-form pair because they do not occur together in common objects. I constructed an additional 15 related and 15 unrelated practice pairs using visual, functional, and other types of features. For example, *<lives in aquariums>* and *<swims>* conveyed information about location and motion. No feature used in the filler or practice pairs appeared in the experimental trials.

Procedure. Participants were tested using PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993) on a Macintosh PowerMac 8600 computer, equipped with a 17-inch colour monitor. Response latencies were recorded using a CMU button box that measured the time in milliseconds between the onset of the presentation of the second feature of each pair and the button press. Participants responded “yes” by pressing a button with the index finger of their dominant hand and “no” using the index finger of their non-dominant. Participants received written and verbal instructions concerning how relatedness decisions were to be made, as well as examples of related (*<is crunchy>* and *<grows in gardens>*) and unrelated pairs (*<covered in felt>* and *<used in salads>*). They were instructed to silently read each feature and respond as quickly and accurately as

possible as to whether the paired features were related – that is, whether they go together in living or non-living things. All participants received the same list of related and unrelated pairs in a random order.

Each trial proceeded as follows: First, a fixation point (+) appeared in the center of the screen for 500ms. The fixation point was then replaced by the first feature for 500ms, after which time the second feature appeared on the line below the first one so that both were present until the participant responded. Participants received the 30 practice trials followed by the 80 experimental trials. Each session took approximately 15 minutes.

Design

The dependent variables were decision latency and the square root of the number of errors (Myers, 1979). The independent variable was modality (within vs. cross). Modality was within participants (t_1) but between items (t_2).

Results and Discussion

Mean decision latencies were significantly faster for form-form ($M = 888$ ms, $SE = 51$ ms) than for function-form pairs ($M = 1032$ ms, $SE = 50$ ms), $t_1(20) = 5.94$, $p < .00001$, $t_2(38) = 1.93$, $p = .06$. Error rates were not expected to differ because both pair-types were judged to be at least moderately related in the off-line relatedness rating task, with mean relatedness ratings greater than 5 out of 7 for both conditions. The error rates for form-form pairs ($M = .05$, $SE = .01$) did not significantly differ from that for function-form pairs ($M = .07$, $SE = .01$), $t_1(20) = 2.04$, $p > .05$, $t_2(38) = 0.77$, $p > .4$.

Despite being judged off-line as significantly less strongly related, relatedness decision latencies were shorter for within-modal pairs. These results are consistent with the assumption of a deep integration hierarchy in which modally distributed information

is first integrated in single-modality convergence zones, which then feed into cross-modal convergence zones. The latency advantage for within-modal items is inconsistent with the assumption of a shallow integration hierarchy and with the predictions of amodal models, both of which predict no modality effect. Because statistical factors were equated and differences in perceived relatedness favoured cross-modal pairs, if anything, these models would predict a cross-modal advantage. Instead, the reverse pattern was found.

One potential concern regarding Study 2 is that, although the length of the second feature was controlled, multiple word stimuli may be problematic for two reasons. The form features used in the experiment frequently begin with the five-character (including the space between words) phrase “*has a*”, whereas the functional features frequently begin with the 8-character phrase “*used for*”. This introduces two problems, the first of which concerns timing. The tendency for functional features to begin with longer phrases potentially demands longer reading times on average for these features. Although this was not believed to be a problem for Study 2 because the SOA was chosen to exceed the expected reading time for the first feature, it is possible that the SOA did not provide enough time to read some of the initially-presented features in cross-modal pairs. A second potential issue is that initial phrases are repeated between within-modal features for 12 of 20 pairs (<*has fins*>, <*has gills*>), although the remaining 8 contained different initial phrases (<*has a lid*>, <*made of glass*>). Initial phrases were never repeated in cross-modal pairs. Study 3 addresses these potential issues in two ways. First, the same features were used as targets for both the within- and cross-modal pairs, thus automatically controlling for a number of variables. Second, multiword items such as <*has a handle*> were divided into stem (*has a*) and content (*handle*) components which

were presented sequentially. Importantly, decision latency was measured from the onset of the content component, nullifying any potential advantage of repeated stems within a pair.

Study 3

Study 3 replicates Study 2 using a more rigorously controlled set of items, and using a modified presentation paradigm. I expected to replicate the results, which supported hierarchically deep models.

Method

Participants. Thirty-eight University of Western Ontario undergraduates received either course credit or \$10 for their participation.

Materials. I selected 18 related form-form and 18 related function-form pairs from our norms (see Appendix B). The pairs were yoked to create function-form-form triples, such that one form feature was the second feature of a within- and cross-modal pair. For example, *<used for storing food>*, *<has windows>*, *<has doors>* is a triplet containing a form-form (*<has windows>*, *<has doors>*) and a function-form pair (*<used for storing food>*, *<has doors>*).

The pairs comprising each triplet were selected to be as similar as possible with respect to several variables expected to influence decision latencies. In addition to the variables identified in Study 2 as potential influences on decision latencies, the pairs were matched on distinctiveness, which is the inverse of the number of concepts in which the feature appears in our norms, and thus indexes the likelihood with which the feature could cue a particular basic level concept. Because stem and content words were presented separately, content word length should primarily influence reading times and

was therefore equated. In addition, each group had similar distributions of higher ($r^2 > .25$), medium ($.25 r^2 \leq .15$) and lower ($r^2 < .15$) correlated pairs, such that the cross-modal group had 10, 3, and 5 and the within-modal group had 9, 3 and 6 high, medium and lower correlated pairs, respectively. One of the 18 yoked triplets was dropped from the analyses that follow because, although judged to be moderately related in the offline relatedness rating task ($M = 3.1$ out of a possible score of 7), more than 40% of participants judged the within-modal pair (*<has a motor>*, *<has sails>*) to be unrelated in the speeded task. Item characteristics for the remaining 17 sets of yoked feature pairs are summarized in Table 5. Importantly, because the second features were identical for the within- and cross-modality groups, the groups are automatically matched on all variables concerning them. Within- and cross-modal pairs did not differ significantly on any variable, other than the number of concepts in which the features appeared together, and this difference favoured cross-modal pairs.

Relatedness Ratings. Forty University of Western Ontario students not participating in the main task provided off-line relatedness ratings. Each participant rated half of the within- and half of the cross-modal pairs, and saw only one pair from each triplet. The procedure and analyses were identical to the relatedness ratings in Study 2.

As in Study 2, the perceived relatedness of the cross-modal pairs ($M = 4.3$, $SE = 0.1$) was greater than that of the within-modal pairs ($M = 4.1$, $SE = 0.1$), though this difference was significant by participants $t_1(39) = 2.64$, $p < .02$, but not by items $t_2(16) = 0.39$, $p > .7$. Again, this difference was expected to facilitate decision latencies of cross-modal pairs, if anything.

Table 5

Equated variables in Study 3

Factor	<i>Function-Form</i>	<i>Form-Form</i>	<i>t</i> (16)	<i>p</i>
	<i>M</i> (<i>SE</i>)	<i>M</i> (<i>SE</i>)		
Mean % Shared Variance	29.4 (5.0)	27.2 (4.7)	0.73	>.4
Number of Shared Concepts	2.6 (0.2)	2.2 (0.1)	2.38	.03
Content Word Length (1 st feature)	6.6 (0.4)	6.2 (0.5)	0.78	>.4
Number of Unique (1 st feature)	2.5 (0.7)	2.5 (0.9)	0.00	>.9
Concepts per Feature (1 st feature)	5.2 (0.9)	4.8 (0.9)	0.57	>.5
Distinctiveness (1 st feature)	0.25 (0.03)	0.32 (0.04)	-1.77	>.1

The sets of yoked pairs were divided in half between two experimental lists, such that one list contained the within-modal pairs for half of the sets, and the cross-modal pairs for the other half. The remaining items were assigned to the second list. An equal number of unrelated filler pairs (nine form-form, and nine function-form) were constructed using features not appearing among the related pairs, and were used in both lists.

All features were divided into stem and content components. The stem comprised the initial segment of the feature name and included prepositions, conjunctions, and the verbs (*has a, used for, is*). The content component comprised the final one or two words that carried much of the feature's meaning. Because experimental and filler pairs were matched with respect to the number of times each stem appeared, $\chi^2(8) = 11.21, p > .15$, and an equal number of within- and cross-modal pairs appeared among the experimental and filler items, the stems did not cue the response. Participants therefore needed to wait for and process the content component to make the relatedness judgment, and response latencies were measured with respect to the onset of the content component of the second feature.

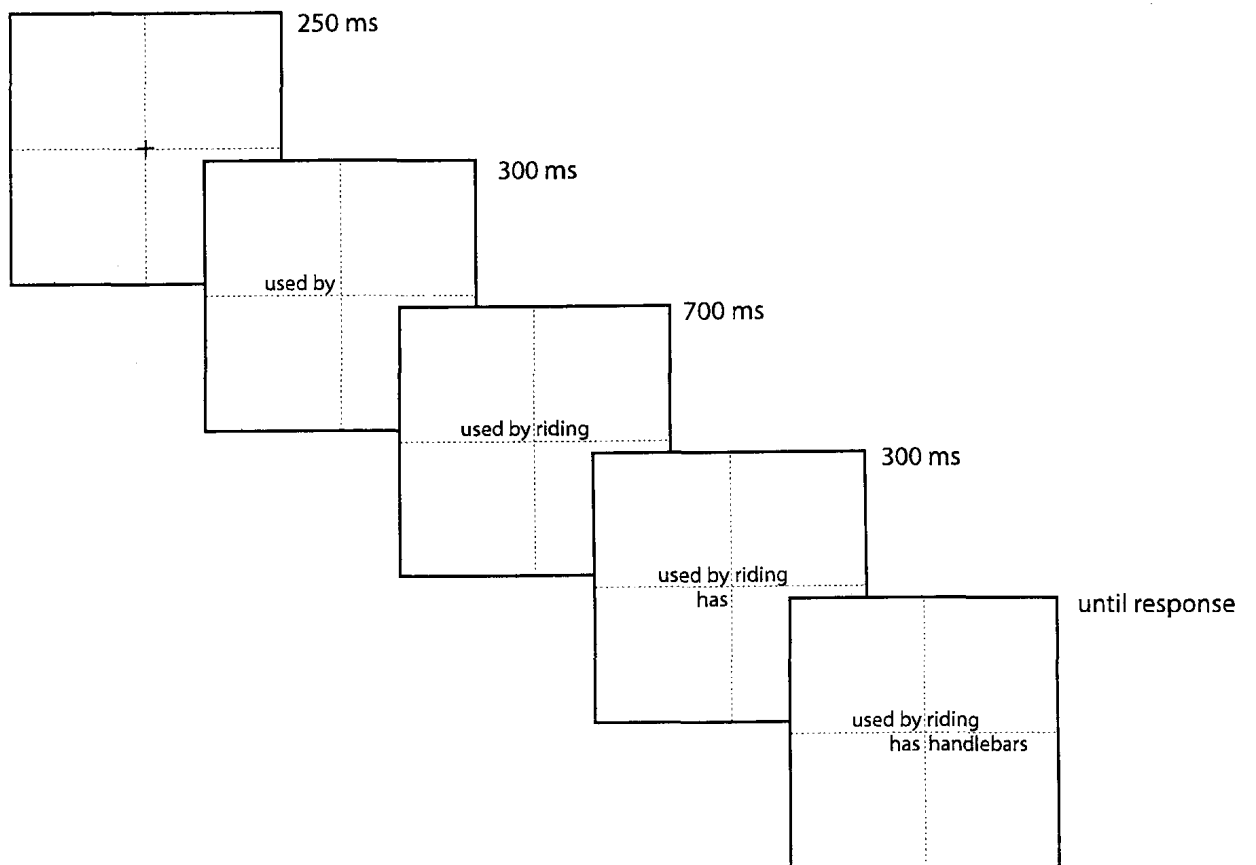
Procedure. Participants were tested using E-Prime (Psychology Software Tools Inc., 2002) on an AMD Athlon 64 3200+ personal computer, equipped with a 17-inch colour monitor. The instructions to participants were identical to Study 2.

Each trial proceeded as follows: First, a blank white screen was presented for 2000 ms, followed by a vertically and horizontally centered fixation cross ('+') for 250 ms, after which time it disappeared. For purposes of displaying the stem and content components of the first and second features, the screen was divided into vertically and

horizontally centered quadrants, though the quadrant boundaries were not visible (see Figure 8). Each quadrant was justified opposite to its position (i.e., the top-left quadrant was lower-right justified, etc.). The left quadrants were used for presenting stem component (*used by*) and the right quadrants were used for presenting the content components for each feature. The first feature's stem component was presented in the top left quadrant immediately following the removal of the fixation cross. After 300 ms, its content component (*riding*) was then presented in the top right quadrant so that the first feature (*used by riding*) was displayed in the top half of the screen, centered on the boundary between stem and content components. After 700 ms, the second feature's stem (*has*) appeared in the lower left quadrant. Finally, after 300 ms, its content (*handlebars*) appeared in the lower right quadrant, so that the second feature (*has handlebars*) was displayed in the lower half of the screen immediately below the first feature. Both features remained on the screen until the participant responded. The use of quadrants avoided cuing the length (and possibly the identity) of the content words, and allowed them to always appear in the same screen position. Decision latencies were recorded using a button box that measured with millisecond accuracy the time between the onset of the presentation of the second feature's content component and the button press.

Participants received five lead-in trials immediately followed by 36 experimental trials. The experiment took less than 15 minutes to complete.

Figure 8. Presentation sequence for stem and content components of stimuli in Study 3.



Design

Analyses of variance were conducted using participants (F_1) and items (F_2) as random variables. The dependent variables were decision latency and the square root of the number of errors (Myers, 1979). The independent variable was modality (within vs. cross), which was within both participants and items. List was included as a between-participants dummy variable and item rotation group as a between-items dummy variable to stabilize variance that may result from rotating participants and items over lists (Pollatsek & Well, 1995). Effects involving these dummy variables are not reported.

Results and Discussion

Incorrect trials and those with decision latencies greater than 3 standard deviations above the grand mean were removed from the analysis (3% of the trials). As indicated previously, one related pair (<has a motor>, <has sails>) had an error rate that exceeded 40%. Presumably this occurred because the motor on vehicles such as sailboats is not at all salient for most people without first-hand experience with these vehicles, and so participants saw these features as mutually exclusive. Thus, this pair and the corresponding yoked pair (<used for cruising>, <has sails>) were removed from the analyses.

For the remaining 17 pairs, relatedness decision latencies were greater for cross-modal ($M = 1087$ ms, $SE = 44$ ms) than for within-modal pairs ($M = 1013$ ms, $SE = 37$ ms), which was significant by participants, $F_1(1, 36) = 10.16, p < .004$, but not by items, $F_2(1, 15) = 2.36, p > .1$. Participants were quite accurate, and within-modal ($M = .07, SE = .01$) and cross-mode error rates ($M = .09, SE = .02$) did not differ, both F 's < 1 .

The pattern of results from Studies 2 and 3 suggests that multimodal feature representations activate one another relatively quickly in a hierarchically deep integration structure. Stimuli in both experiments were selected to be as closely matched as possible on several factors that are assumed to influence relatedness decision latencies, and where they differed, these variables tended to favour the cross-modal items. Nonetheless, there was an advantage for within-modal pairs in both studies.

One limitation of the design of Studies 2 and 3 is that feature inference predicts an advantage only for within-modal pairs. This introduces a potential issue in that cross-modal processing may involve some form of task-switching and thus incur a performance cost, reflected in longer decision latencies for cross-modal pairs. Modality-switch costs have been demonstrated in the perception literature (Spence, Nicholls, & Driver, 2000), and in the concepts literature using feature verification (Pecher, Zeelenberg, & Barsalou, 2003). Thus, it is unclear whether the within-modal advantage in Study 2 and 3 is the outcome of a deep integration hierarchy or modality switching costs. To alleviate this concern, Study 4 uses a task in which deep models predict a latency advantage for cross-modal trials.

A second potential issue is that functional features appeared only in cross-modal pairs, and it is possible that functional information takes longer to access. For example in Barsalou's (1999) perceptual symbol systems account, retrieval of an object's function (<*used for opening cans*>) would involve a *simulation*, or mental re-enactment of the sensorimotor elements of that function (visualizing the process of grabbing a can, applying a can-opener to the can, and then turning the key that causes the opener to follow and cut the rim of the can). Because these functions typically unfold over time, it

is reasonable to predict that retrieval of this information takes more time than retrieving relatively static form information, such as an object's shape (though retrieval of some form information in the perceptual symbol systems framework may involve a mental rotation from an object's canonical orientation, Edelman & Bühlhoff, 1992). Though function features were always presented first, it is possible that they imposed additional processing demands that carried over during the processing of the second feature. The within-modal speed advantage observed in Studies 2 and 3 could therefore be interpreted as an advantage for pairs involving only form features over pairs involving functional features.

A final issue concerns whether these results extend to other representational modalities. Because only visual form and functional features were used in Study 2 and 3, then it remains unclear whether evidence for deep hierarchy models would be found using other modalities. Study 4 was designed to deal with all of these concerns.

Chapter Four: Feature Integration in Pattern Completion

Study 4

People possess a rich knowledge of many objects that is not limited to just a single representational modality. Accordingly, when we identify an object, we generally do so using only a fraction of the information we possess about it, and yet are able to use this information to retrieve virtually any other knowledge that we have about the object. Thus, using knowledge of within- or cross-modal properties to retrieve concepts is a common event and is therefore an interesting test paradigm for a couple of reasons. First, like feature inference, it is a basic cognitive process in which we are constantly engaged, and therefore, elucidating the manner in which it occurs is of central importance to the study of cognitive function. Second, and most importantly in the present context, whereas deep hierarchy models predict a within-modal advantage for feature inference, they predict a cross-modal advantage for activating a concept from partial information. Finding a cross-modal advantage would therefore address the explanation that processing cross-modal information may be generally disadvantaged.

Study 4 tests the predictions for hierarchically deep and shallow models for a task involving pattern completion from incomplete information. Moreover, this study uses functional features in both within- and cross-modal conditions, and incorporates features from other knowledge types (smell, taste, sound, etc.). This design addresses the questions raised in the previous section: whether the results found in Studies 2 and 3 are attributable to a function versus form feature processing difference, and whether the results extend to other representational modalities. A novel dual feature verification task was used wherein two sequentially presented features were followed by a concept name.

Under the assumption that verbally presented features activate other features prior to the presentation of a concept name, and that this activation facilitates feature verification, the hierarchically deep and shallow models predict different patterns for concepts preceded by within- and cross-modal pairs. Hierarchically shallow models predict no difference between within- and cross-modal decision latencies, whereas hierarchically deep models predict a cross-modal advantage because two features should activate clusters of correlated features for two modalities in parallel. Amodal models again predict no effect of modality.

Because feature verification latencies have been shown to depend on the presence of feature clusters (McRae, de Sa, & Seidenberg, 1997), concept intercorrelational density was also manipulated. A feature cluster is a group of interrelated features. For example, birds such as *robin* and *seagull* are high-density concepts because there are many other objects in the world that possess similar sets of co-occurring features, such as *<has feathers>*, *<has wings>*, *<has a beak>*, *<flies>*. Because these features regularly occur together in many concepts, they are highly correlated with one another. Because they are mutually correlated, these features form a correlated feature cluster, one property of which being that activating any single feature in the cluster rapidly activates other features in the cluster. At the other end of the spectrum are concepts such as *ashtray*, which possesses features that may be idiosyncratic (*<used for cigarettes>*) or else occur with so many other features in other concepts as to be poor predictors for other features that the concept might possess (*<made of plastic>*). Such features are not part of correlated feature clusters, and therefore, activating such features would not be expected to lead to the rapid activation of other features belonging to the concept. Because the

advantage for cross-modal pairs predicted by hierarchically deep models relies on clusters of intercorrelated features in each modality in which activation may occur in parallel, I divided the target concepts into those with low and high intercorrelational density. Intercorrelational density is the sum of the percentage of shared variance across all of a concept's significantly correlated feature pairs (because it is a sum, it is no longer truly a percentage). Low and high density concepts were those with a standardized intercorrelational density less than -1 and greater than 1, respectively. Though any effect of target density does not distinguish among shallow, deep, or amodal theories, because high density concepts possess larger clusters of intercorrelated features on which pattern completion depends, this factor was included in the analyses. All feature-based models that allow for influences of correlated features predict a latency advantage for high-density items. However hierarchically deep models allow (but do not require) that target density may interact with pair type, such that greater facilitation may be found for high-density than for low-density concepts, whereas shallow and amodal models do not make this prediction.

Method

Participants. Twenty-six University of Western Ontario undergraduates received \$10 for their participation. Three participants were dropped because their mean response latency (2 participants) or error rate was greater than three standard deviations above the grand mean.

Materials. Standardized intercorrelational density was calculated for all concepts in our norms. Thirty-six concept-feature sets of the form {<A>, , <C>, TARGET} were selected such that {<A>, } were correlated within-modal features (<has a

sheath>, <*has a blade*>), {<*C*>, <*B*>} were correlated cross-modal features (<*used for cutting*>, <*has a blade*>), and {<*A*>, <*B*>, <*C*>} were listed for the TARGET concept (*sword*) in our norms (see Appendix C). Of these 36 item sets, half contained low-density and half contained high-density targets, which possessed standardized density scores of less than -1 and greater than 1, respectively.

To control for factors that might be expected to influence decision latencies, items were selected such that characteristics of the first feature, and the relationships between it and the second feature and target concept were equated (see Tables 6 and 7). As described below, two low density items were dropped because of error rates of approximately 50% on both within- and cross-modal trials, so the equated statistics in Tables 6 and 7 are for the remaining items only. Because the effect of density was of secondary interest, less emphasis was placed on matching low and high density items. Nonetheless, as Table 6 shows, the conditions were generally matched on most variables of interest. These variables included concept familiarity, which was the subjective rating, from 1 to 9 of the degree to which people feel familiar with the concept; mean shared variance, which was the average of the shared variances in the feature norms for yoked pairs; mean production frequency, which is the average of the production frequencies in the feature norms for the first features of the yoked pairs; mean CPF, which was the average of the number of concepts associated with the first features of the yoked pairs; and mean intercorrelational strength, which was the average of the sum of shared variances in the feature norms between the first features of the yoked pairs and other features of the target concept. The significantly higher target concept density scores for

Table 6

Equated variables in Study 4 by density condition

Factor	High Density		Low Density		t(32)	p
	M	SE	M	SE		
Target Concept Density	635	58	92	13	9.10	< .0001
Target Concept Familiarity	6.0	.5	6.9	.4	1.33	> .15
Mean % Shared Variance	17.7	2.3	21.4	5.9	0.61	> .5
Mean Production Frequency	10.9	0.8	11.2	1.1	0.21	> .8
Mean CPF	7.5	1.3	6.1	0.9	0.88	> .35
Mean Paired Concepts	4.6	0.7	2.9	0.2	2.30	< .05
Mean Intercorrelational Strength	264	33	220	35	0.94	> .35

Table 7

Equated variables in Study 4 by modality and density condition

Factor	<i>Within-Modal</i>		<i>Cross-Modal</i>		<i>t^a</i>	<i>p</i>
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>		
High Density						
% Shared Variance	18.3	2.8	17.1	2.8	0.40	> .7
Production Frequency	11.9	1.3	10.0	1.1	1.07	> .3
Paired Concepts	4.4	0.7	4.8	0.9	-0.59	> .5
Concepts per Feature	6.9	1.6	8.1	0.2	-0.55	> .5
Intercorrelational Strength	243.0	37.7	286.0	35.0	-1.32	> .2
Length	11.6	1.3	12.6	0.9	-0.59	> .5
Low Density						
Mean % Shared Variance	22.1	6.2	20.8	6.0	0.43	> .6
Production Frequency	10.6	1.2	11.8	1.3	-0.87	> .3
Paired Concepts	2.8	0.2	2.9	0.2	-0.37	> .7
Concepts per Feature	6.4	1.3	5.8	1.0	-0.42	> .6
Intercorrelational Strength	230	37	209	37	0.79	> .4
Length	11.3	1.1	14.1	1.3	-1.49	> .15

^a paired t-tests using 17 degrees of freedom for high density items, and 15 degrees of

freedom for low density items

the high density items confirms that the two sets of items indeed differed on this dimension. The mean number of concepts in which the features were paired together was significantly higher for high density versus low density items, however, because correlated features were used in filler items, this would tend to slow decision latencies for high- relative to low-density items because the high-density target features should cue more potentially distracting targets.

To ensure that within- and cross-modal conditions were matched in both density conditions, yoked items were pairwise-matched on length in characters (including spaces), percent of shared variance, the number of concepts in which the features were paired together, and the number of concepts per feature. Because the features were associated with particular concepts, items were additionally matched on the production frequency (the number of people who listed each feature for that concept) to control for the likelihood of the feature bringing to mind the target concept (Ashcraft, 1978). Additionally, because the initially presented features are assumed to activate the target concept by activating its features, and the shared variance among these features influence the speed with which this occurs (McRae et al., 1997), I matched items on intercorrelational strength, which is the sum of the shared variance between the presented features and other features of the concept. This ensured that differences in feature verification latencies were not attributable to differences between the groups with respect to the degree to which the initially presented features are correlated with other features of the concept. Because it was not relevant to the task, perceived relatedness of the feature pairs was not matched for Study 4. The speeded relatedness decision task used in Studies 2 and 3 could be seen as a detection task, in which participants were under time pressure

to detect the related correlated test items among the uncorrelated filler items. For these studies, relatedness ratings were a measure of extent, under the assumption that the greater the extent to which two features are perceived as related, the faster would be the detection of this relation under time pressure. In Study 4, correlated features were used in all trials – indeed, the correlated pairs used for some filler items was higher than that for some experimental items – and therefore, perceived relatedness would not cue the feature verification response. Instead, the extent to which the features belong to the concept was controlled by equating the conditions on the production frequency, paired concepts, and concepts per feature, all of which would be expected to influence the probability that each feature pair would cue the target concept. Because they are common to both members of a yoked pair, relations between the second feature and target concept were automatically equated. All knowledge types in our norms, with the exception of encyclopaedic and taxonomic features, were represented. Target concepts included both living and non-living concrete objects.

The 36 yoked within-modal and cross-modal items were assigned pseudo-randomly to two experimental lists, such that each list included 18 within- and 18 cross-modal trials, and no two yoked items appeared in the same list. Because each member of a yoked pair shared a target concept, half the items in each list had high-density targets, and half had low-density targets. Both lists contained an equal number of filler items in which it was not true that both features were true of the target concept. The filler items were divided into thirds (i.e., 12 of each kind) in which either the first feature, the second feature, or neither was true of the target concept. Filler items used features and medium-density target concepts not appearing among the experimental items. The use of

correlated features in the filler items ensured that there was at least one concept in which both features occurred for both experimental and filler trials, and therefore participants needed to wait for the presentation of the target concept to respond accurately. No feature or target concept appeared more than once in either list.

Procedure. Participants were tested using E-Prime (Psychology Software Tools Inc., 2002) on an AMD Athlon 64 3200+ personal computer, equipped with a 17-inch colour monitor. Each trial proceeded as follows. First, a blank white screen was presented for 1500 ms, followed by a vertically and horizontally centered fixation cross ('+') for 500 ms, after which time it disappeared. The first feature was displayed immediately above the position of the fixation cross, followed 1000 ms later by the second feature immediately below the first. Both features remained on the screen for an additional 1000 ms, after which time the target concept was presented in upper case letters on the line immediately below the second feature. The 1000 ms SOA for each feature ensured that participants had sufficient time to read the longest multiword features. The two features and the target concept remained on the screen until the participant responded. Responses were collected using a button-box that recorded the time between the onset of the target concept and the button press with millisecond accuracy. Participants responded "yes" by pressing a button with the index finger of their dominant hand and "no" by pressing another button using the index finger of their non-dominant hand.

There were 16 lead-in practice trials comprising a mixture of yes and no trials immediately followed by 72 experimental trials. The experiment took about 15 minutes to complete.

Design

Analyses of variance were conducted with decision latency and the square root of the number of errors as the dependent variables. The independent variables were modality (within vs. cross) and concept density (low vs. high). Modality was within participants (F_1) and items (F_2), whereas density was within participants but between items. List and item rotation group were again included as dummy variables.

Results and Discussion

Two low-density items with error rates approaching 50% were dropped from the analysis. Because the proportions of error responses were similar for within- and cross-modal conditions for both items, the errors were attributed to ambiguity of, or unfamiliarity with, the relationship of the second feature to the target concepts (*BEETS*, and *SCREWS*). Decision latencies greater than 3 standard deviations above the grand mean were replaced by the cutoff value (4% of the trials).

Mean verification latencies and error rates are presented in Table 7. Critically, feature verification latencies were shorter for cross-modal ($M = 773$ ms, $SE = 29$ ms) than for within-modal items ($M = 849$ ms, $SE = 34$ ms), $F_1(1,24) = 15.68, p < .0007, F_2(1,30) = 7.83, p < .009$. Verification latencies were shorter for high-density concepts ($M = 770$ ms, $SE = 29$ ms) than for low-density concepts ($M = 851$ ms, $SE = 33$ ms), which was significant by participants, $F_1(1,24) = 26.80, p < .0002$, but marginal by items, $F_1(1,30) = 3.46, p < .08$. Concept intercorrelational density did not interact with feature modality, $F_1(1,24) = 2.80, p > .1, F_2 < 1$.

Table 8

Feature verification latencies and error rates for Study 4

Factor	Error Rate		Latency (ms)	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
High Density				
Within-Modal	.05	.02	825	31
Cross-Modal	.05	.02	725	32
Low Density				
Within-Modal	.11	.02	883	39
Cross-Modal	.07	.02	826	30

Because it was hypothesized that a greater cross-modal advantage might be apparent for high-density items than for low-density items, planned comparisons were conducted between the modality conditions for both levels of density. For low density targets, the 58 ms advantage for cross-modal pairs was significant by participants, $F_1(1,48) = 4.62, p < .04$, but not by items, $F_2(1,30) = 2.11, p > .15$. For high density targets, the 99 ms advantage for cross-modal pairs was significant, $F_1(1,48) = 14.23, p < .0006, F_2(1,30) = 6.46, p < .02$. Thus, it appears that cross-modal pairs more quickly activated concepts than did within-modal pairs for concepts with relatively few and many clusters of intercorrelated features, though this difference was numerically but not statistically greater for higher density concepts.

Participants were generally quite accurate. There was no difference in error rates between within and cross modal pairs, $F_1 < 1, F_2(1,30) = 1.02, p > .3$. Error rates were marginally lower for high density than for low density concepts, $F_1(1,24) = 3.44, p < .08, F_2(1,30) = 3.57, p < .07$. The two factors did not interact, $F_1(1,24) = 1.48, p > .2, F_2 < 1$.

One potential concern was that the within-modal paired features could be construed as redundant for some items. For example, it is perhaps the case that little additional information is provided in pairing *<eaten in sandwiches>* with *<is edible>*, as the former implies the latter. One alternative explanation of these results, therefore, is that people are slower at responding to items for which less information is given. The above analyses were repeated with the removal of seven yoked pairs of items identified as potentially having features with overlapping meaning. A reanalysis of the item characteristics found the stimuli to be matched on all variables of interest, with the exception of the feature length (in characters) for low-density items, which favoured within-modal ($M = 10.2, SE$

= 0.8) over cross-modal ($M = 14.4$, $SE = 1.5$) items, $t(12) = 2.29$, $p = .04$. Despite this bias, feature verification latencies were again shorter for cross-modal ($M = 778$ ms, $SE = 31$ ms) than for within-modal items ($M = 875$ ms, $SE = 33$ ms), $F_1(1,24) = 24.29$, $p < .0001$, $F_2(1,23) = 10.13$, $p < .004$. Verification latencies were shorter for high-density concepts ($M = 794$ ms, $SE = 32$ ms) than for low-density concepts ($M = 860$ ms, $SE = 32$ ms), which was significant by participants, $F_1(1,24) = 12.43$, $p < .002$, but not by items, $F_1(1,23) = 1.45$, $p > .2$. Concept intercorrelational density interacted with feature modality by participants, $F_1(1,24) = 7.00$, $p < .015$, but not by items $F_2 < 1$. Error rates did not differ between conditions, nor was there an interaction in any analysis, all $F_s < 1$.

Study 4 advances our understanding of how pattern completion is influenced by the sensorimotor modality of available object features by suggesting a particular type of neural mechanisms that allow integration of distributed semantic representations into coherent concepts. Importantly, there was a clear cross-modal latency advantage that is predicted by deep hierarchy models. These results are inconsistent with the predictions of shallow hierarchy and amodal models. Although mode and density did not reliably interact, the cross-modal advantage was somewhat stronger for high density concepts.

One final thing to note is that there are other ways I could have measured the relationship between the feature pairs and the clusters of intercorrelated features within target concepts. The item groups were based on the intercorrelational density of the target concept, an index of the correlated feature clusters within the concept. However, McRae et al. (1997; 1999) showed that intercorrelational strength (i.e., the sum of the squared correlations between a feature and each other feature of a particular concept) influences feature verification latencies. For example, *<hunted by people>* is more strongly

intercorrelated with other features of *deer* than of *duck*, and verification latency was shorter for *deer*. Because the cross-modal advantage for pattern completion depends on these features activating feature clusters that overlap with the concept, the data were reanalyzed using two alternative methods. First, I classified items into high and low intercorrelational strength on the basis of the mean intercorrelational strength of the presented features, using the same logic underlying McRae et al. (1997; 1999). The second method incorporated both intercorrelational strength and intercorrelational density by multiplying these two values together to create a composite variable. For both measures, the placement into high and low groups remained the same for the vast majority of items, and consequently, the pattern of results after re-analyses of the verification latency data did not qualitatively differ from those presented above.

General Discussion

The present research used complementary behavioural tasks to identify the most likely neural architecture underlying integration of multimodal semantic representations. These tasks were of interest for numerous reasons. They were of practical interest because shallow and deep integration hierarchies make different predictions for them. From a broader perspective, however, they are of interest because of their ubiquity in cognitive processing. A sound or a smell can cause a concept to automatically come to mind, and allows access to a wealth of knowledge about that concept. Similarly, feature inference, both within and across sensorimotor modalities, is so routine that we commonly use texture words like “fuzzy” or “smooth” to describe objects we have seen but have never touched. Thus, understanding how the brain carries out these basic cognitive processes is central to our understanding of human behaviour.

An important implication of the ubiquity of feature inference is that the classification of features into modalities can be fuzzy. On one hand, features such as ‘red’ would seem to unambiguously describe an object’s colour. However, as mentioned the term ‘fuzzy’ itself can refer to a tactile experience, as when describing a tennis ball, or to a visual experience, as when describing the mould growing on something one would rather not touch. In McRae et al. (2005), single knowledge type labels were assigned to features by consensus. Thus, because some verbal descriptors, such as ‘fuzzy’ can refer to multiple modalities, this equivocation leads to uncertainty regarding the modality of the information retrieved when reading feature descriptors, such as *<is fuzzy>*. Given that support was found for deep hierarchy models in each of the studies, one concern might be whether these results might be attributable to an ambiguity regarding the knowledge

types for some stimuli. I do not believe this to be the case for a few reasons. First, these knowledge-type classifications have been used to explain a number of behavioural phenomena, including susceptibility to various neurological impairments (Cree & McRae, 2003), and speeded semantic decisions for concrete objects (Grondin, Lupker & McRae, in press), which therefore suggests that the knowledge type classifications are generally coherent. More importantly, however is that because multiple meanings are activated when polysemous words are read (Pexman, Hargreaves, Bodner & Pope, 2008), it is likely that all knowledge type senses are available when an ambiguous feature descriptor is read. However, the predicted effects depend not on the particular modalities, but instead on whether the relationship between the features is within- or cross-modal. Thus, the biggest concern is that there might be some sense in which a cross-modal feature is also within-modal, and vice versa. However, even if this were indeed the case for every item used in Studies 2, 3 and 4, then the contrasts in each of these experiments would be between items from the same condition, which should *reduce* between condition variability. In other words, knowledge type fuzziness should make it *less* likely to find an effect. Thus, while it is desirable that the knowledge type labels used in this research is an accurate description of the features to which they are applied, and I believe that, for the most part they are, a degree of ambiguity does not take away from the major findings of this work.

Studies 2 and 3 used a feature relatedness task. The results were consistent with the within-modal advantage predicted by hierarchically deep models. However, necessary aspects of these experiments' design yielded open questions. The first concern was that functional features, which may take longer to retrieve, only appeared in the slower cross-

modal condition, and so it was unclear whether the results reflected a general disadvantage for processing functional features. The second was that cross-modal items may have incurred a task-switching cost and correspondingly longer decision latencies. Finally, I wanted to extend the results to knowledge types other than visual form and function information. Study 4 addressed all three of these concerns using a double feature verification task, and again the results supported a deep integration hierarchy.

These results constrain semantic memory models in a few ways. First, multimodal models assume that people do not literally store lists of features, such as *<has wings>* or *<buzzes>* to represent concepts. Rather, verbal labels such as these are used as a shorthand to refer to the underlying visual or auditory information. Thus, these results confirm that the orthographic presentation of these features can activate associated sensorimotor feature knowledge implicitly learned through real world experience. Second, these results indicate that, as feature knowledge becomes available, it activates clusters of associated semantic information distributed across a network of brain areas that gradually come to approximate a concept. This is not a trivial point because one may imagine that feature-to-feature and feature-to-concept activation is carried out using two different systems. In dual-coding theory (Paivio, 2007), for example, concrete concepts have both lexical and perceptual representations, and a concept's lexical representation can act as a pointer to all of its features. Thus, brain areas specialized for language processing could act as a single convergence area during concept activation when a concept name is read or heard. The pattern of results across four studies was consistent with the assumption that multimodal semantic integration occurs in a deep integration hierarchy, wherein information encoded by modality-specific representational units is integrated in proximal

modally-tuned convergence zones which in turn pass activation forward to higher order convergence zones that integrate multiple modalities.

Relation to Previous Research

The results are clearly compatible with models such as those put forward by Damasio (1989) and Simmons and Barsalou (2003) that explicitly specify a deep organization. In these models, higher level convergence zones integrate more information sources and can therefore encode more abstract relationships. Because both Damasio's and Simmons and Barsalou's models predict that higher level convergence zones appear in more anterior brain regions, these results are also compatible with research showing that these regions support the learning of abstract concepts such as reward learning (Fuster et al., 2000), and feature conjunctions (Eacott & Gaffin, 2005). Furthermore, the anterior shift described by Thompson-Schill (2003), which is the association between conceptual processing and activity in brain regions just anterior to perceptual areas, may reflect activity among focused integration units as they provide re-entrant activation to perceptual areas in what Barsalou (1999) would call a simulation of the concept.

The effects of modality found in the present experiments are inconsistent with amodal representational systems (Caramazza, 1991; Moss & Tyler, 2001). Lacking any modality distinction, semantic processing largely driven by statistical properties such as feature correlations. Because these factors are assumed also to play a role in multimodal models, stimuli used in Studies 2 to 4 were matched on these factors, making it difficult for these results to be explained by amodal models. This research therefore adds to the growing body of literature that challenges the argument that our representational system stores only abstract amodal representations.

The support for a deep integration hierarchy is also inconsistent with the ways in which our own models have been implemented. For example, Cree et al. (2006) simulated influences of distinctive versus shared semantic features in a connectionist model that used direct connections among its feature units. That is, convergence zones were not incorporated in any way (which could be implemented as sets of hidden units in connectionist models). In fact, our implemented models have not even instantiated modality-specificity, but we have argued elsewhere that sensorimotor modality is an important organizing force in how the brain represents information (Cree & McRae, 2003; McRae, 2004). Thus, the present results suggest that our future models should include intramodal hidden unit clusters that feed into a higher level convergence zone.

Humphreys and Forde's (2001) HIT model was designed to capture some of the properties of Farah and McClelland's (1991) fully interconnected network in a hierarchical model in which processing from a top-level layer cascades back to earlier functional and sensory representational units. Both are shallow models as presented, and would therefore need to be extended somewhat to be consistent with our results. Farah and McClelland's model was an existence proof that category specific deficits can arise out of a segregation of sensory and functional knowledge and was not intended to be an argument about other aspects of the underlying neural architecture. The HIT model was designed to confirm that the same patterns hold in a system that allows reciprocal activation (or "cascading") from a higher-level structure that acts as a convergence zone. Their model, however, includes only three types of knowledge: structural descriptions of the general visual form of objects, functional and inter-object associative information that they called "semantic" knowledge, and name representations. However, it has been

shown here and elsewhere that knowledge of other sensory modalities such as taste are important aspects of concepts for some categories, such as food (Simmons, Martin, & Barsalou, 2005). The main focus of Humphreys and Forde's argument was the importance of the role of reentrant activation from higher-level integration sites, and the role of the various sensorimotor modalities in the model was secondary. Thus, they sketched alternative versions illustrating the role of cascading activation in different tasks relying primarily on different modalities, for example, the role of auditory knowledge when identifying a guitar by its characteristic sound. The present results show how the mechanism of cascading reentrant activation functions in a single deep hierarchy model capable of accounting for those phenomena that the HIT model was designed to address, and further refines the model by demonstrating that this cascade may function differently within and across modalities. These results are also consistent with a model of form-colour synaesthetic activation proposed by Smilek, Dixon, Cudahy and Merikle (2001), in which reentrant activation of form information leads to activation of colour information before the visual form of a visually-presented character has been completely resolved. Because these results suggest that a similar process occurs during normal concept processing for multiple sensorimotor modalities, the model proposed by Smilek et al. may apply to other types of synaesthesias. Moreover, it suggests that the critical difference between synaesthetic and typical cross-modal processing may be with respect to speed: if cross-modal activation occurs rapidly enough, a simultaneous experience of two modalities may result, whereas slower cross-modal activation would allow within-modal inhibition of cross-modal experiences, and allow a sufficiently large temporal

delay as to promote the subjective feeling that any cross-modal information that reaches awareness was the result of an inference.

Plaut (2002) describes a hybrid semantic model in which functional specialization of semantic representations are graded, rather than all-or-none. Although this model may initially appear to make the assumption of a shallow integration hierarchy, a functionally quasi-deep integration hierarchy emerges through training. Although Plaut's model uses a single convergence area, it is biased to form short connections, which creates pools of connections that conjoin features from the same input modality adjacent to each sensory input layer. Longer connections tend to form mid-way, in multidimensional space, between input layers for multiple modalities. Because a connection in the model is represented by a single numerical value representing its weight, physical distance is not directly encoded in this network, but is instead simulated by weakening connections that link units that are more distant in multidimensional space. Because the ability for one unit to activate another depends on the strength of the connection between them, weakening this connection will reduce the efficiency with which the units may activate one another, thereby increasing the time required for activation to spread from one to the other.

It is not a requirement that convergence zones be clearly defined in the brain, and the graded organization in Plaut's model may be an important property of convergence zones. Thus, though it lacks explicit boundaries between within- and cross-modal convergence areas (i.e., both types of units appear in the same layer), functionally equivalent pools of connections emerge as a result of this bias: integration units closer in multidimensional space to each input modality and approximate a within-modal convergence zone, and units that are more distant will form weaker cross-modal

connections. Because transmission speed is directly related to connectivity strength, this model predicts faster within-modal than cross-modal communication between two features, which is predicted by the deep, but not the shallow hierarchy assumption. It is therefore probable that Studies 2 and 3 could be simulated in a connectionist network without explicitly defined within- and cross-modal integration areas, similar to that described by Plaut. The model also predicts faster pattern completion from cross-modal than within-modal pairs for the same reason, because cross-modal input can activate other within-modal features for both modalities in parallel via the relatively stronger within-modal connections, whereas within-modal would propagate along relatively weaker cross-modal connections in order to activate other modalities. A model similar to Plaut's would therefore also be likely capable of simulating Study 4. Note, however, that simulating Studies 2 to 4 in a quasi-hierarchical model would not imply that these studies fail to distinguish between shallow and deep hierarchical models. Rather, it would further support the central prediction of hierarchical models: that connective distance varies by modality and determines how information from each modality influences conceptual processing.

Finally, the present results are inconsistent with Patterson et al.'s (2007) claim that anterior temporal cortex is the sole hub through which semantic memory is routed because, as presented, theirs is a shallow hierarchy model. Although the evidence for deep hierarchy models does not refute their claim that this area plays an important role in integrating semantic information, it does suggest that this area is perhaps a top-level multimodal convergence zone in a deep integration hierarchy, and that their theory would

need to be modified to include lower-level unimodal convergence zones to account for these results.

Insights into Functional Connectivity

Mapping functional connectivity among various functional areas in the brain is a rather difficult undertaking, owing in part to the absence of natural anatomic boundaries in white matter, aside from notable exceptions, such as the corpus callosum. In recent years, techniques have been developed to estimate functional connectivity between brain areas by examining inter-regional correlations in fMRI time series (fcMRI), or measuring the shared variance between EEG signals measured simultaneously at different scalp locations. In addition to identifying those inter- and intra-modal relationships that would be most sensitive to the manipulations used in Studies 2 to 4, the functional connectivity estimates calculated in Study 1 provide testable hypotheses for those studying functional connectivity at the macroscopic level using anatomical and neuroimaging techniques. For example, Griffiths, Green, Rees, and Rees (2000) imaged participants while they listened to sounds with sources that appeared to move to one side of the head, or else remained apparently stationary. Activity within a network of areas involving right parietal cortex was involved in processing perceived sound movement, the implication being that information travels from auditory cortex to parietal cortex (an area implicated in motion perception) and presumably back to auditory cortex.

The role of the dorsal stream – a pathway of functional connectivity beginning in visual cortex and extending forward to parietal cortex – in the selection and programming of appropriate motor acts on objects has been studied extensively over the past two decades (Goodale & Milner, 1992). One relatively new method of assessing functional

connectivity involves combining transcranial magnetic stimulation (TMS) with measurements of event-related potentials (ERP) to investigate how activation of one brain region is causally responsible for activation in another region (Taylor, Walsh, & Eimer, 2008). Using the motion-sound and motion-vision functional connectivity estimates from Table 3, the functional connectivity between motion and vision processing areas may be as much as five times greater than between motion and sound areas, if one makes the simplifying assumption that these estimates apply bidirectionally. Because the imaging research described above provide potential loci for TMS and ERP measurements, one could assess whether the extent to which current measured at electrodes placed over visual cortex differs from that measured at electrodes placed over auditory cortex following the administration of an electromagnetic pulse to posterior parietal cortex, and if so, if it differs to the extent predicted by the functional connectivity estimates from Study 1.

Another application of the functional connectivity estimates from Study 1 is to the synaesthesia literature. In synaesthesia, sensations from one sensory modality are reliably experienced either as a direct superimposition (as in colours superimposed over numbers or letters in colour-grapheme synaesthesia) or as salient and reliable visualizations (as in perceiving time as physical locations in space in temporal-spatial synaesthesia). One proposed explanation is that synaesthesia results from a breakdown of the neural pruning that typically occurs in the developing brain (Baron-Cohen, 1996). According to this hypothesis, whereas many communication pathways between functional areas die off over the course of normal development, this does not happen normally in synaesthetes, allowing atypical cross-talk between perceptual processing systems. Thus, processing of

information in one modality causes atypical perceptual experiences in another modality. If synaesthesia is related to the degree of functional connectivity between brain regions, then perhaps the functional connectivity estimates from Study 1 could be used to predict the prevalence of various types of synaesthesias. Interestingly, using the average magnitude of shared variance measures in Table 2, colour has the highest estimated functional connectivity with other modalities, the noteworthiness of which being that forms of synaesthesia involving the perception of colour imposed on other senses are the most prevalent. Related statistically-based investigations of cross-modal correlation learning in synaesthetes may provide insight into the development of neural pathways in this population, and in general.

Perceptual and Conceptual Integration

I suggested in the Introduction that the question of how multimodal conceptual knowledge is integrated into unified concepts is the conceptual analog of the perceptual binding problem (Triesman, 1996), which argues that object perception requires a mechanism through which multiple input streams across multiple modalities are incorporated into single perceptual objects. A critical assumption underlying the present research and the body of research supporting distributed multimodal semantic representations is that the brain regions specialized for perceptual processing in each modality are also used in processing semantic information from the corresponding modality. From this perspective, the primary difference between perception and semantic processing is in the source of the inputs driving processing. In the case of perception, processing is initially driven by environmental inputs (though top-down processing introduces learned information not present in the environment into the processing

stream), whereas conceptual processing may be primarily driven by internal representations which act as inputs (though environmental cues provide context that may guide the retrieval of information and provide additional information). If concepts and percepts are indeed two sides of the same coin, then the present research and the investigation of multimodal perception mutually inform one another.

This issue is perhaps most clearly apparent in the synaesthesia literature, where there is some question of whether synaesthetic experiences reflect perceptual or semantic processes: In particular, Smilek and Dixon (2002) suggest that, because non-synaesthetes can be made to demonstrate patterns of interference in a Stroop task similar to those found in synaesthetes through extensive overtraining (e.g., naming a colour in response to a visually presented digit over thousands of trials), cross-modal synaesthetic activations may reflect semantic associations. However, as the authors point out, if one considers the subjective reports of such experiences—namely, that digit-colour synaesthetes such as C see coloured digits overlaid atop digits visually presented in black, and therefore sees two different colours. Because overtraining non-synaesthetes does not lead to such experiences, Stroop effects arising from automatic cross-modal activation induced in non-synaesthetes likely arises from different processes than those induced by synaesthetic experiences (Smilek & Dixon, 2002).

Because I used words as stimuli, rather than pictures or sounds, our studies clearly involve multimodal integration in the conceptual system. Nonetheless, in a perceptual symbols system framework (Barsalou, 1999), retrieval of the underlying meaning of these stimuli is assumed to induce a pattern of activation in the primary sensory areas similar to that experienced during perception, albeit presumably less vividly. If this is so,

then the same neural circuits could be used in both conception and perception, and therefore be subject to the same processes. The support for hierarchically deep representational models suggests a similar organization for the routing of multimodal information in perceptual processing.

Similarly, the idea of inhibitory processes is not new to either the perception or concept literature. Nonetheless, research into cross-modal integration circuits in perception have identified two ways in which cross-modal inhibition is accomplished during perception, and may provide insight into how multimodal inhibitory processes may influence concept processing. For example, the present research used only positively correlated features from the McRae et al. (2005) norms. In part, this is because these feature norms contain very few negatively correlated features, partly as a result of how people provide features for concepts – people overwhelmingly list properties that are true of concepts, rather than those that are not. One would expect negatively correlated features to be reciprocally inhibitory: if one feature occurs, the other does not. Thus, assuming one could overcome the challenge of generating an appropriate set of positively and negatively correlated features from different modalities, one might adapt the methodologies used here to investigate the role of within- and cross-modal inhibition in multimodal integration.

Alternative Experimental Paradigms and Future Directions

Given that this research fundamentally concerns the neural architecture underlying semantic processing, one might question the suitability of behavioural methods for such an investigation. For example, neuroimaging data appears to allow more transparent inferences of brain structure than do decision latencies. Aside from the relatively low

expense in terms of time and money, there were a number of advantages, however, to the methods employed in the present research. First, the consistency among the studies suggests that the experimental tasks reliably measure the mental processes that they are supposed to measure, and by extension, give insight into the underlying brain function. A second related advantage was that the assumed hierarchies provide clear behavioural predictions that were borne out in the data. As discussed in the Introduction, there are a number of candidate brain regions that may act as convergence zones, and thus any investigation that explicitly examined particular brain structures would require a number of additional assumptions regarding the particular functions of various brain regions. Additionally, each methodology carries with it its own set of considerations.

Neuroimaging. The ERP literature is included under this heading along with fMRI investigations of brain function, all of which have become mainstream techniques in recent years. The strength of ERP research rests in its excellent temporal resolution, with the ability to measure changes in brain electrical activity measured with millisecond resolution. Spatial resolution for ERP, however, is rather poor, limiting localization of brain activity to rather coarse discriminations of anterior versus posterior, or left versus right. Thus, although the temporal resolution of ERP would be suitable for studying the time-course of multimodal integration, its poor spatial resolution makes it inadequate for identifying the particular neural structures involved.

fMRI, in contrast, has the opposite profile, with excellent spatial but rather poor temporal resolution, measured in seconds. The behavioural data presented here found that decisions involving the integration of multiple semantic features typically occurred under one second, and therefore poses a problem for this experimental paradigm. Moreover,

even if one were to use a design able to overcome this limitation, another problem remains in that until the technology is able to identify the activity of individual neurons, one would have to make the unrealistic assumption that the clusters of neurons comprising individual voxels are homogenous. In the discussion of candidate convergence zone loci, it was suggested that the anterior shift described by Thompson-Schill (2003) may reflect activity in modality-specific convergence zones.

In Figure 3, the modality-specific convergence zones are represented as distinct areas placed ahead of the representational areas. It is just as likely, however, that the neurons that encode information and the neurons that integrate input from multiple neurons are not cleanly partitioned. One possibility is that the neurons are arranged in a gradient with distal areas containing predominantly either encoding or integration populations, but with the majority of the volume containing a relatively heterogeneous population. Such an arrangement would make distinguishing the different neural populations quite difficult.

TMS. Transcranial magnetic stimulation, or TMS, involves the application of an electromagnetic pulse (EMP) to the scalp, causing a temporary disruption of brain activity for a relatively small area, and has been used to create “virtual brain lesions”, which can be used to investigate the causal relationship between activity in a particular brain area and different types of behaviour. This technique has straightforward applications to the investigation of multimodal integration; by comparing performance on tasks, such as those used in the present experiment, that involve the integration of information from different modalities with and without virtual lesions, one can gain insight into what brain areas might be responsible for integrating information from different sensorimotor modalities. There are a few limitations of this approach however.

The first is that, as discussed above, the population of within-modal integration neurons may be interspersed among the representation neurons for that modality, making it impossible to selectively lesion integrative neurons. Thus, this technique may be more appropriate for identifying loci for cross-modal integration, which may be sufficiently distant from the areas that represent either modality. On the other hand, potential multimodal integration sites found in ventral and medial areas such as the perirhinal cortex cannot be investigated using this technique because the EMP delivered by TMS is limited to brain areas underlying the scalp.

Behavioural Paradigms. The null hypothesis was the assumption of a shallow integration hierarchy for all experiments described in this investigation. Thus, the present set of studies was not ideally suited to testing the predictions of a shallow integration hierarchy. A more suitable test of the shallow hierarchy assumption might involve some type of disruptive attentional task. For example, if there were but a single integration site, then the processing load imposed by cross-modal feature integration would be expected to also influence within-modal integration. Under the assumption of a deep hierarchy, however, one might expect that even if a top-level integration site were under load, within-modal processing might be relatively unaffected because it uses different structures. In such an experiment, an effect is predicted for shallow models in the disruption versus no disruption for both within- and cross-modal integration, whereas the effect of disruption under the deep hierarchy assumption would depend on the nature of the disruption and the integration modality.

Conclusion

In the present research, I combined insights regarding the influence of distributional statistics of concepts motivated by connectionist modeling with insights from neurally-based theories of conceptual organization to test ideas regarding neural functional connectivity using complementary behavioural tasks. My studies provide clear evidence for the existence of a deep hierarchy in a multimodal distributed semantic memory system.

References

- Allport, D. A. (1985). Distributed memory, modular systems and dysphasia. In S. K. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 32 – 60). Edinburgh, Scotland: Churchill Livingstone.
- Baier, B., Kleinschmidt, A., & Müller, N. G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *The Journal of Neuroscience*, *26*, 12260 – 12265.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences*, *22*, 577 – 660.
- Chao, L. L., & Martin, A. (1999). Cortical representation of perception, naming, and knowing about colour. *Journal of Cognitive Neuroscience*, *11*, 25–35.
- Cree, G.S., McNorgan, C., McRae, K. (2006). Distinctive features hold a privileged status in the computation of word meaning: Implications for theories of semantic memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *32*, 643-658.
- Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163 – 201.
- Cree, G. S., McRae, K., & McNorgan, C. (1999). An attractor model of lexical conceptual processing: Simulating semantic priming. *Cognitive Science*, *23*, 371-414.
- Damasio, A. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25 – 62.

- Damasio, H., Tranel, D., Grabowska, T., Adolphs, R., & Damasio, A. (2004). Neural systems behind word and concept retrieval. *Cognition*, *92*, 179 – 229.
- Edelman, S. & Bühlhoff, H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision Research*, *32*, 2385 – 2400.
- Eckert, M. A., Kamdar, N. V., Chang, C. E., Beckmann, C. F., Greicius, M. D., & Menon, V. (2008). A cross-modal system linking primary auditory and visual cortices: Evidence from intrinsic fMRI. *Human Brain Mapping*, *29*, 848 – 857.
- Farah, M. J., & McClelland, J. L. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. *Journal of Experimental Psychology: General*, *120*, 339 – 357.
- Fuster, J. M., Bodner, M., & Kroger, J. K. (2000). Cross-modal and cross-temporal association in neurons of frontal cortex. *Nature*, *405*, 347 – 351.
- Goldberg, R.F., Perfetti, C.A., & Schneider, W. (2006a). Perceptual knowledge retrieval activates sensory brain regions. *The Journal of Neuroscience*, *26*, 4917 – 4921.
- Goldberg, R.F., Perfetti, C.A., & Schneider, W. (2006b). Distinct and common cortical activations for multimodal semantic categories. *Cognitive, Affective and Behavioural Neuroscience*, *6*, 214 – 222.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*, 20 - 25.
- Green, A. E., Fugelsang, J. A., Kraemer, D. J. M., Shamos, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, *1096*, 125 – 137.

- Grondin, R., Lupker, S. J., & McRae, K. (in press). Shared features dominate semantic richness effects for concrete concepts. *Journal of Memory & Language*.
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity and interactivity in object recognition: "Category specific" neuropsychological deficits. *Behavioural and Brain Sciences*, 24, 453 – 509.
- Kaas, J. H., & Collins, C. E. (2004). The resurrection of multisensory cortex in primates: Connection patterns that integrate modalities. In Calvert, G. A., Spence, C., & Stein, B. E. (Eds.) *The Handbook of Multisensory Processes*. (pp. 285 – 294). Cambridge, MA, US: MIT Press.
- Kellenbach, M. L., Brett, M., & Patterson, K. (2001). Large, colorful, or noisy? Attribute- and modality-specific activations during retrieval of perceptual attribute knowledge. *Cognitive, Affective & Behavioural Neuroscience*, 1, 207 – 221.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211 – 240.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25-45.
- Martin, A., Haxby, J. V., Lalonde, F. M., Wiggs, C. L., et al. (1995). *Science*, 270, 102 – 105.
- McNorgan, C., Kotack, R. A., Meehan, D. C., & McRae, K. (2007). Feature-feature causal relations and statistical co-occurrences in object concepts. *Memory & Cognition*, 33, 418 – 431.

- McRae, K., de Sa, V., & Seidenberg, M. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, *126*, 99-130.
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods*, *37*, 547-559.
- Muggleton, N., Tsakanikos, E., Walsh, V., & Ward, J. (2007). Disruption of synaesthesia following TMS of the right posterior parietal cortex. *Neuropsychologia*, *45*, 1582 – 1585.
- Mummery, C. J., Patterson, K., Hodges, J. R., & Price, C. J. (1998). Functional neuroanatomy of the semantic system: Divisible by what? *Journal of Cognitive Neuroscience*, *10*, 766 – 777.
- Parker, A., & Easton, A. (2004). Cross-modal memory in primates: The neural basis of learning about the multisensory properties of objects and events. In Calvert, G. A., Spence, C., & Stein, B. E. (Eds.) *The Handbook of Multisensory Processes*. (pp. 333 – 342). Cambridge, MA, US: MIT Press.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*, 976 – 987.
- Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science*, *14*, 119 – 124.

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008).

There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*, 161 – 167

Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A

computational account of optic aphasia. *Cognitive Neuropsychology*, *19*, 603 – 639.

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive

research: A suggestion for a better and more powerful analysis. *Journal of*

Experimental Psychology: Learning, Memory, & Cognition, *21*, 785-794.

Riddoch, M.J., Humphreys, G.W., Coltheart, M., & Funnell, E. (1988). Semantic systems

or system? Neuropsychological evidence re-examined. *Cognitive Neuropsychology*, *5*,

3 – 25.

Simmons, K. W., & Barsalou, L. (2003). The similarity-in-topography principle:

Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, *20*, 451 –

486.

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., Scott,

K., & Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal

experiences. *Perception*, *35*, 1024 – 1033.

Simmons, K. W., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods

activate gustatory cortices for taste and reward. *Cerebral Cortex*, *15*, 1602 – 1608.

Simmons, W. K., Ramjee, V., Beauchamp, M., McRae, K., Martin, A., & Barsalou, L.

W. (2007). A common neural substrate for perceiving and knowing about color.

Neuropsychologia, *45*, 2802 – 2810.

- Sitnikova T., West, W. C., Kuperberg, G. R., & Holcomb, P. J. (2006). The neural organization of semantic memory: Electrophysiological activity suggests feature-based segregation. *Biological Psychology*, *71*, 326 – 340.
- Sloman, S. A., Love, B. C., & Ahn, W. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, *22*, 189 - 228.
- Smilek, D., & Dixon, M. J. (2002). Towards a synergistic understanding of synaesthesia: Combining current experimental findings with synaesthetes' subjective descriptions. *Psyche*, *8(01)*. Retrieved October 12, 2008, from <http://psyche.cs.monash.edu.au/v8/psyche-8-01-smilek.html>.
- Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *Journal of Cognitive Neuroscience*, *13*, 930-936.
- Spence, C., Nicholls, M.E.R., & Driver, J. (2000). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*, 330–336.
- Sporns, O., & Tononi, G. (2007): Structural determinants of brain dynamics. In Kirska, V. K., & McIntosh, A. R. (Eds). *Handbook of brain connectivity*. (pp. 117 – 148). New York, NY, US: Springer.
- Taylor, P. C. J., Walsh, V., & Eimer, M. (2008). Combining TMS and EEG to study cognitive function and cortico–cortico interactions. *Behavioural Brain Research*, *191*, 141 – 147.
- Thompson-Schill, S. L. (2003). Neuroimaging studies of semantic memory: Inferring “how” from “where”. *Neuropsychologia*, *41*, 280 – 292.

- Tranel, D., Grabowski, T. J., Lyon, J., & Damasio, H. (2005). Naming the same entities from visual or from auditory stimulation engages similar regions of left inferotemporal cortices. *Journal of Cognitive Neuroscience*, *17*, 1293 – 1305.
- Triesman, A. (1996) The binding problem. *Current Opinion in Neurobiology*, *6*, 171 – 178.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327 – 352.
- Tyler, L. & Moss, H. E. (2001). Towards a distributed account of conceptual knowledge. *Trends in Cognitive Sciences*, *5*, 244 – 252.
- Warrington. E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, *110*, 1273 – 1296

Appendix A

Within- and cross-modal pairs for Study 2

	Feature A	Feature B
Function-Form	worn for warmth	made of wool
	used in bands	made of brass
	used for storage	has doors
	used for killing	has a trigger
	worn by women	made of material
	used for passengers	has an engine
	worn on feet	made of leather
	used for transportation	has 4 wheels
	eaten in pies	has seeds
	worn by men	has pockets
	worn around neck	made of gold
	used for cutting	has a blade
	used by riding	has a seat
	hunted by people	has 4 legs
	eaten in salads	has leaves
	used for holding things	made of plastic
	eaten by people	has gills
	used for cooking	is electrical
	used for cleaning	has a handle
	used for eating	made of ceramic
Form-Form	has fins	has gills
	has a lid	made of glass
	has sleeves	made of cotton
	made of brick	has a roof

has a blade	has a wooden handle
has pockets	has buttons
has blades	is electrical
has windows	has a roof
has 4 wheels	has an engine
has a tail	has hooves
has buttons	made of material
made of wood	has drawers
has scales	has teeth
has a seat	has wheels
has a tail	has legs
is flat	is rectangular
has seeds	is round
has doors	has windows
has fur	has whiskers
has horns	has hooves

Appendix B

Yoked within- and cross-modal pairs used in Study 3

Within-Modal Feature A	Cross-Modal Feature A	Target Feature B
has two wheels	used by riding	has handle bars
has a stinger	used for travel	has wings
has a clasp	used for gifts	made of gold
has buckles	worn for walking	made of leather
has hinges	used by opening	has a lock
has sections	used for juice	has pulp
made of diamonds	worn around neck	made of silver
has buttons	worn by men	has cuffs
has pockets	worn for covering	has zippers
made of fur	worn for rain	has a hood
has a metal blade	used for gardening	has a handle
has heels	worn on feet	has soles
made of rubber	used for games	is round
has a roof	used for living in	made of brick
has cushions	used for relaxing	has armrests
has pedals	used for cargo	has wheels
has doors	used for storage	has shelves

Appendix C

Yoked within- and cross-modal pairs and target concepts used in Study 4.

Within-Modal Feature A	Cross-Modal Feature A	Feature B	Density	Target
has seeds	tastes sweet	is round	High	ORANGE
fires	used by the police	is loud	High	PISTOL
has a thumb	worn on hands	made of wool	High	GLOVES
has zippers	worn for covering legs	has buttons	High	TROUSERS
used for war	has a trigger	used for killing	High	GUN
talks	eats seeds	sings	High	PARAKEET
has feet	waddles	has a beak	High	DUCK
has laces	worn on feet	made of leather	High	BOOTS
has pockets	worn for warmth	has sleeves	High	COAT
eaten in sandwiches	has fins	is edible	High	TUNA
tastes tart	has peel	tastes sour	High	LEMON
eats rodents	has talons	flies	High	FALCON
has a blade	used for digging	has a wooden handle	High	HOE
used for watching television	has armrests	used for sleeping	High	COUCH
is soft	used for relaxing	is comfortable	High	SOFA
has a pit	is juicy	has skin	High	PLUM
used for passengers	has an engine	used for transportation	High	JET
has a collar	worn by women	made of cotton	High	BLOUSE
has a cushion	used by sitting on	has legs	Low	STOOL
put on ceilings	has lightbulbs	used by hanging	Low	LAMP
is smooth	used by throwing	is hard	Low	STONE

worn on heads	has a strap	worn for protection	Low	HELMET
is flat	used for carrying things	made of plastic	Low	TRAY
used for construction	has a pointed end	used for holding	Low	SCREWS
is damp	is dark	is cold	Low	BASEMENT
tastes different flavours	eaten by baking	tastes good	Low	PIE
has fur	used in experiments	has whiskers	Low	MOUSE
has a plug	is hot	is electrical	Low	TOASTER
has taps	used for washing	has a drain	Low	BATHTUB
has fangs	slithers	has scales	Low	PYTHON
used for music	is gold	used in bands	Low	SAXOPHONE
has shelves	used for storing food	has doors	Low	CUPBOARD
is pointed	used for writing	is thin	Low	PEN
produces sound	has strings	produces music	Low	GUITAR
is square	used for storage	is rectangular	Low	BOX
eaten as pickles	has leaves	eaten in salads	Low	BEETS

Appendix D

Ethics Approval Notice



Department of Psychology The University of Western Ontario
 Room 7418 Social Sciences Centre,
 London, ON, Canada N6A 5C1
 Telephone: (519) 661-2067/Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	08 03 01	Approval Date	08 03 04
Principal Investigator	Ken McRae/Chris McNorgan	End Date	08 08 31
Protocol Title	Property relatedness decisions		
Sponsor	n/a		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: <http://www.uwo.ca/research/ethics/>)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:

- a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
- b) all adverse and unexpected experiences or events that are both serious and unexpected;
- c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2007-2008 PREB are: Mike Atkinson, David Dozois, Bill Fisher and Matthew Maxwell-Smith

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files