

Chapter 11. Perceptual and conceptual interactions in object recognition and expertise

Thomas W. James, Indiana University

George S. Cree, University of Toronto, Scarborough

Corresponding author

Thomas W. James

1011 E Tenth St.

Bloomington, IN 47401

812-856-0841

thwjames@indiana.edu

1 Introduction

Visual object recognition is a fundamental cognitive operation performed countless times each day. Yet despite the ease with which we are able to recognize familiar objects in our environment, object recognition is extremely complicated, and the computations necessary to recognize objects have eluded cognitive modelers for decades. This complexity can be observed in the fact that engaging in even the simplest of object recognition tasks recruits widespread regions of the cortex (for review, see Grill-Spector, 2003).

Curiously, even though many of the regions active during object recognition lie outside of traditional visual cortex, object recognition is, to a large extent, thought of as a mainly visual process. Most (if not all) current theories of general object recognition are based almost solely on data from visual experiments (Biederman, 2000; Deco & Rolls, 2004; Grossberg, 1999; Humphreys & Forde, 2001; Tarr & Bulthoff, 1998). Furthermore, since the early 1980s, object recognition has often been considered a bottom-up process that occurs separately from other cognitive operations (Marr, 1982; Pylyshyn, 1999). But object recognition involves much more than dissecting an object into its visual features; we rapidly, and often involuntarily, access non-visual information about the object as well.

What role does this non-visual information play in object recognition? One could argue that access to this “semantic” information comes only after the visual processes have finished, and this would be consistent with the theory that object recognition is visually driven and that semantic memory is a cognitive capacity separate from vision. But, we will argue that conceiving of object recognition as solely visual misrepresents the flexible and adaptive nature of the mechanisms that support it.

We believe that perceptual and conceptual processes, and representations are integrated into a distributed sensorimotor system that underlies both object recognition *and* semantic memory. In the rest of the chapter we will discuss evidence that object recognition is best conceived of within a framework wherein both conceptual influences and multisensory interactions play a central role in object recognition. A majority of the data that we discuss will focus on what has been learned from the study of people who have become experts through training in the lab, or are already recognized experts in the domain of interest. We will first consider the interface of perceptual and conceptual processing, and how the two interact. We will then turn to structure within the visual/semantic pathways, and review evidence for the existence of highly-interactive, modality-specific processing streams.

2 Expertise, and Perceptual and Conceptual Processes for Object Recognition

In the past 15 years we have witnessed a shift away from the traditional view of conceptual representation as amodal, verbal, and proposition-like, to the view that it is firmly grounded in modality-specific processing channels, and represented in memory areas that are contiguous with, and perhaps even overlapping, the sensory pathways through which information is acquired (for example, see Barsalou, 1999). These pathways have been described as a continuum, varying in how much, and what sort of, top-down processing has been done to the bottom-up input, and the degree of productivity afforded by the resultant representations (Goldstone & Barsalou, 1998). This change in viewpoint has brought about a search for new behavioral paradigms that allow us to explore both the representations and computations that occur throughout these processing channels.

One promising approach is to explore how processing and representation change as one gains perceptual and/or conceptual experience with a set of objects. For us, this idea was inspired by the perceptual expertise approach to studying object recognition in which one studies the processing and representational changes that occur as participants gain experience in dealing with closed-class sets of complex visual objects, such as Greebles (Gauthier & Tarr, 1997, 2002). Examples of this kind of approach also exist in the concepts and categorization literature and the expertise literature.

In the concepts and categorization literature there have been numerous studies designed to look at how the goals and demands of categorization influence perceptual processing. Researchers have examined, for example, how knowledge influences the interpretation of visual features (Wisniewski & Medin, 1994), and how information about the concepts to be acquired influences the creation of new perceptual features (Shyns, Goldstone, & Thibaut, 1998). Within categorization learning paradigms, using small sets of stimuli, Goldstone and colleagues have been able to address several interesting phenomena, including sensitization of existing perceptual dimensions, sensitization of novel perceptual dimensions, perceptual reorganization, and unitization (for review, see Goldstone, Steyvers, Spencer-Smith, & Kersten, 2000).

In the expertise literature there have been many studies looking at how perceptual abilities change as people gain experience within their domain of expertise. There is considerable evidence, for example, that experts parse the perceptual features of the world differently than do novices, this having been demonstrated in the domains of chess (De Groot,

1966), radiology (Myles-Worsley, Johnston, & Simons, 1988), sexing day old chicks (Biederman & Shiffrar, 1987), and beer tasting (Peron & Allen, 1988), to name a few prominent examples. Although insightful, the approaches taken in this literature are limited for our purposes because they rarely involve the use of object classes that allow for systematic manipulation of the perceptual components of the objects of expertise, and the knowledge that is associated with those objects.

In the next section we will discuss evidence from two different paradigms used in our labs that show ways in which conceptual knowledge can influence perceptual processing, even in tasks that have traditionally been thought to be primarily perceptual in nature.

2.1 Behavioral Studies with Experts

Why is the processing of faces disrupted in experts when the ability to process configural/holistic information is impaired? The standard answer is that faces are visually complex objects that share common part configurations, for which to discriminate among them one must learn about not only specific parts and when they occur together, noting very subtle differences, but also the arrangements of those parts in relation to one another. Learning the specific positions of parts in relation to one another is thought to be crucial for discrimination because the problem of learning the parts themselves, and the variation in parts across individuals, is simply too difficult for the common observer. The difficulty is easy to see when you consider that each part can vary on multiple dimensions (e.g., height, width, color, etc.), and perception of very subtle differences in parts, and combinations of parts, is needed for identification. In such cases, implicit learning mechanisms that pick up on statistical variations in all variables having to do with the parts, including their global configurations, may be the most efficient learning mechanism. In short, the learner is awash in a sea of parts, and may automatically grab hold of any information they can that will help keep them afloat.

One source of information that is typically not available in perceptual expertise learning tasks in the laboratory, but that could act to anchor perceptual part-based information, making learning of the parts easier, is conceptual knowledge. If the subject understood why the parts vary as they do, and why specific combinations of parts are likely to occur together, then this could make the learning problem substantially easier. Knowledge of genetics, and why specific combinations of morphological traits are likely to occur together, are examples of such knowledge.

Furthermore, during the task of identification, conceptual knowledge could help the learner focus on specific visual parts that serve as cues to appropriate semantic knowledge about the object. Consider, for example, the rock hyrax, a small animal that to the untrained eye appears visually similar to a groundhog. An appropriately trained biologist would not be fooled by the surface visual similarity, and would point out that the clubbed foot of the rock hyrax serves as an obvious visual cue to the fact that the hyrax is more closely genetically related to the elephant than to the groundhog. These two kinds of conceptual knowledge are most likely not available for ordinary undergraduates who are tested on visual memory for faces, and may or may not be available for other perceptual experts such as birders or automobile experts. If true, then one could predict that people with appropriate conceptual knowledge would not be left afloat in the sea of parts, but rather, might prefer a parts-based strategy to recognizing objects from their domain of expertise. In other words, the visual recognition mechanisms in play could vary for different kinds of visual experts. One domain in which it is possible to test this prediction is that of birds, in which there are both enthusiasts, known as birders, and scholars, known as ornithologists.

Although both birders and ornithologists share a passion for birds, the learning goals of each are clearly different. Birders must often identify birds in the field, in sub-optimal conditions, and under time pressure. This may lead them to rely heavily on information about shape and distinctive visual features. Additionally, although birders are known to possess conceptual knowledge about birds, such as knowledge about habitats, flying patterns, and feeding habits, it is reasonable to suppose that the knowledge would be of a different sort than that possessed by ornithologists. Specifically, birders may lack knowledge about how morphological traits reveal clues about other conceptual knowledge. Ornithologists, on the other hand, are more likely to have detailed semantic knowledge about birds due to extensive book-based research and experience. Most importantly, they have been specially trained to understand the rich variation in morphological traits across species, and how and why those traits are related to other properties, many of which may not be visually present on the exterior of the bird. Thus morphological traits may serve as visual anchors around which ornithologists organize their knowledge about birds. If true, then ornithologists may be capable of recognizing birds in an analytic, parts-based fashion, rather than, or perhaps in addition to, holistically.

Ozubko, Cree, and Bub (2005) reported two experiments designed to test these ideas (Ozubko, Cree, & Bub, 2005). Experiment 1 tested depth of conceptual knowledge about birds by probing knowledge of bird taxonomies. It was reasoned that ornithologists should perform the

best, given their extensive book-based training with birds, followed by birders, and finally novices. Experiment 2 tested memory for pictures of inverted vs. upright birds. It is well documented that recognition of many stimulus classes thought to be processed holistically, such as faces, is impaired if the images are inverted (Yin, 1969). Therefore, individuals recognizing birds in a holistic manner should perform worse with inverted birds than upright birds. It was predicted that birders would be most likely to process birds holistically, and so should have the worst memory for inverted birds. Ornithologists, in comparison, were predicted to perform equally well for inverted and upright birds, due to their reliance on parts-based processing.

Twenty-three novices, 16 birders, and 5 ornithologists participated in the study. The novices self-reported having no specialized knowledge or interest in birds, the birders all had at least 10 years of birding experience (average 26 years of experience with birds), and the ornithologists were all practicing professionals from either the Royal Ontario Museum or the University of Guelph (average 31 years of experience with birds).

Eighty images of birds local to the Toronto area were selected from the Sibley Guide to Birds (Sibley, 2000). For Experiment 1, the Patuxent Bird Identification Infocenter (Gough, Sauer, & Iliff, 1998) was used to classify the birds into biologically related families. Twenty-four triads were constructed such that two of the birds looked similar to one another whereas the third looked different. For consistent triads, the two visually similar birds were also biologically related, whereas the dissimilar bird was unrelated. For inconsistent triads, the two birds that looked similar were not biologically related, and the visually dissimilar bird was related to one of the other two. For Experiment 2, 60 of the bird images were used, 48 of which were of unique bird species and 12 of which were repetitions of species already selected. For the 12 repeated species, juvenile, breeding plumage, and/or female pictures of the species were selected, whereas the other birds were all adult, non-breeding plumage, male birds.

In Experiment 1 participants were told that three birds would be presented, one near the top of the screen, one near the left, and one near the right. The bird on the top was called the host bird. Participants were told to press the “Z” key if they believed that the bird on the left was more similar to the host than the other bird, and to press the “/” key if they believed that the bird on the right was most similar. The few participants that asked what criteria to use when deciding on similarity (a subset of the ornithologists) were told by the experimenter to “use whatever criterion you think is best.”

In Experiment 2 participants were told that images of birds would be presented on the screen, one at a time. They were informed that there would be a memory test, and were

instructed to remember the individual bird images, and not just species names, as they were warned that different images of the same bird species would appear at test. This technique was adapted from Diamond and Carey (Diamond & Carey, 1986), who used it to prevent experts from using mnemonic techniques. Thirty-six images of different birds were shown to participants on the computer during the study phase. No bird species was repeated, so each image was of a unique species. Images were presented one at a time, and presented for 3 s with a 0.5 s inter-stimulus interval where the screen was blank. The first six and last six images were used as buffers to reduce primacy and recency effects, and were not used at test. The images were presented in a randomized order, half inverted, the other half upright. After the study phase, bird images were shown on the screen, two at a time, and participants were instructed to press the “Z” key if they believed they had seen the bird on the left at study, and the “/” key if they believed they had seen the bird on the right. Birds were presented in the same orientation at study and test, and were paired with one other bird in the same orientation. In every pair, one of the birds had been presented at study and one had not. Of the birds not present at study, 12 were the same species as birds seen at study, and were highly visually similar, but were not the same image.

The performance of novices, birders, and ornithologists in Experiment 1 can be seen in Figure 1. The degree to which participants correctly classified birds in inconsistent triads can be viewed as a measure of the degree to which participants were relying on conceptual knowledge of birds, and not on visual similarities. Conservative, nonparametric tests (i.e. the Mann-Whitney and Wilcoxon Signed Ranks tests) were used to analyze the data because inspection revealed that both the assumptions of homogeneity of variance and normality were violated. Mann-Whitney tests were used to look for differences between the groups. Birders performed significantly better than novices on inconsistent triads, $U = 11.0$, $p < .001$, but no difference was found for consistent triads, $U = 176.0$, $p = .832$. Ornithologists performed significantly better than novices on inconsistent triads, $U = 0$, $p < .001$, but no difference was found for consistent triads, $U = 55.0$, $p = .908$. Finally, ornithologists and birders were found to behave similarly on both inconsistent triads, $U = 19.5$, $p = .091$, and consistent triads, $U = 40.0$, $p = 1.00$. These results suggest that, in general, both birders and ornithologists possess more knowledge of bird taxonomy than novices, and, contrary to our predictions, that birders and ornithologists possess the same level of conceptual knowledge about birds as each other. This suggests that both birders and ornithologists are quite knowledgeable when it comes to taxonomic conceptual knowledge regarding birds.

The second experiment examined the memory of participants for inverted versus upright birds. The recognition accuracy of the three groups can be seen in Figure 2. A Mann-Whitney test was conducted to examine possible differences between novices and birders. A significant difference was found between the two groups for both inverted birds, $U = 80.50$, $p = .002$, and for upright birds, $U = 9.0$, $p < .001$. Therefore, as expected, birders were found to have better recognition accuracy for birds than novices, regardless of whether the birds were inverted or upright. A second Mann-Whitney test examined the differences between novices and ornithologists for both inverted birds, $U = 5.50$, $p < .001$, and for upright birds, $U = 1.5$, $p < .001$. Ornithologists, therefore, also demonstrated better recognition accuracy for inverted and upright birds than novices. In general, these two results confirm that birders and ornithologists were “expert” groups inasmuch as both birders and ornithologists had an easier time remembering birds than novices.

A more theoretically interesting difference was also found between birders and ornithologists in the second experiment. A significant difference was found for inverted birds, $U = 15.5$, $p = .04$, but not for upright birds, $U = 39.0$, $p = .968$. This suggests that although birders and ornithologists had similar recognition accuracy for upright birds, birders had poorer recognition accuracy for inverted birds than did ornithologists. Even more interestingly, whereas birders may have performed significantly worse when birds were inverted, ornithologists may have been unaffected. To test this hypothesis, a Wilcoxon Signed Ranks test was used to examine the difference between ornithologists’ recognition accuracy for inverted versus upright birds. No significant difference was found in recognition accuracy for ornithologists for these two stimulus types, $Z = 0$, $p = 1.00$. However, a second test was used to examine the difference between birders’ recognition accuracy for inverted versus upright birds, and a significant difference was found, $Z = -2.338$, $p = .019$. These results suggest that ornithologists were unaffected by the orientation of the birds, performing similarly regardless of whether the bird was inverted or upright, but birders were affected by the orientation of the birds, performing worse when birds were inverted. Thus, birders showed a typical “inversion” effect, but ornithologists did not. The results held even when one considered only the trials on which a very highly visually similar distractor of the same species was present, suggesting that a labeling strategy was not in play.

Configural processing is considered a hallmark of visual expertise. However, only birders in this study demonstrated a reliable inversion effect, which is a marker of configural processing. Ornithologists, who are also experts with birds, did not demonstrate evidence of configural processing, by way of an inversion effect. These data speak against the idea that visual

perceptual expertise necessarily leads to a reliance on just configural processing. But, if configural processing does not necessarily arise from visual expertise, in what circumstances does it arise, and in what cases do experts engage other visual object recognition mechanisms?

One possibility is that configural processing is engaged for objects for which people have very little conceptual knowledge. Greeble training is one example. However, the results of Experiment 1 suggest that things may not be this simple, as birders and ornithologists may have similar levels of conceptual knowledge for local birds. It would therefore be difficult to argue that configural processing arises due to a lack of semantic knowledge, because, to the extent that it was tested, both birders and ornithologists appear to have similar background knowledge about birds.

An alternate idea, and one that has been developed elsewhere by other authors (Bukach, Gauthier, & Tarr, 2006; James, James, Jobard, Wong, & Gauthier, 2005), is that different types of experts have different goals that recruit different perceptual processes. For example, although birders and ornithologists may have similar semantic knowledge about birds, they may have different perceptual goals. Birders may be primarily interested in being able to quickly identify birds in the field. Once a birder spots a bird in the field, they have limited time to identify it before it leaves. Thus, birders may have more utilitarian, quick identification goals when it comes to birds. They may therefore adopt a more holistic method of processing birds out of necessity. However, ornithologists, who may spend more time examining birds in textbooks or in laboratories, may not be interested in being able to quickly identify birds, but instead be interested in being able to draw fine-grained distinctions between similar looking birds who differ on only a few attributes. Goals such as these could bias ornithologists to process birds in a more part-based, analytic manner than birders. Overall, configural processing may arise in situations where individuals seek to simply identify members of a stimulus class, whereas analytic processing may be more likely to arise in situations where experts seek to make classifications of members of a class based on access to conceptual knowledge.

There are a number of important caveats that must accompany these data. First, the sample of ornithologists is very small, and it is important that these results be confirmed in a larger sample before firm conclusions can be drawn. Second, although Experiment 1 was designed to be a difficult task, it is possible that it was actually too easy for these participants. Further analyses of the data revealed that some birders demonstrated taxonomic knowledge as good as, if not better than, some ornithologists, whereas others demonstrated near-novice levels. A ceiling effect may also have been responsible for the fact that ornithologists did not significantly

outperform birders. That is, ornithologists may possess more taxonomical knowledge than birders, but because the task was easy enough for some birders to achieve perfect scores, ornithologists were unable to perform better than birders. It will be important to probe for other kinds of conceptual knowledge about birds in the three groups before conclusions can be drawn about whether birders and ornithologists have comparable conceptual knowledge about birds (which seems unlikely). Third, the Sibley guide is a very popular birding book, and so it is possible that the participants were familiar with these pictures. Finally, this was a correlational study, in which participants self-selected into groups, and so it is impossible to draw firm conclusions. With all that being said, it is still interesting that the inversion effect was not found in the ornithologists, and it will be interesting to see if this finding holds in future work.

2.2 fMRI Studies of Conceptual Learning

The above experiments showed that both birders and ornithologists use non-visual information learned about birds to make classification decisions. Although there has been considerable work documenting the brain networks involved in the perceptual aspects of expert categorization, little is known about the brain networks that support the use of conceptual knowledge for decision-making. To investigate this question, some researchers have compared the brain activation produced by common as opposed to novel objects, with the assumption that brain regions activated more by common objects are being recruited to process the semantic information associated with those objects. Two studies that used this method (Martin, 1999; Vuilleumier, Henson, Driver, & Dolan, 2002) revealed a common focus of activation in the inferior frontal cortex, suggesting that this region is involved in processing semantic information as opposed to visual information. One concern with this method, however, is that novel objects may produce different patterns of activation simply because they are less familiar visually. This problem was addressed in one study (Leveroni et al., 2000) in which participants were familiarized with novel faces before testing. Activation with these newly-learned faces was compared to activation with famous faces. Similar to the other two studies, Leveroni, et al. (2000) found greater activation in the inferior frontal cortex. These findings suggest that the inferior frontal cortex provides a significant contribution to the processing of non-visual information associated with objects.

Another method for investigating the neural substrates of visual and non-visual contributions to object recognition is to compare the effects of perceptual and conceptual priming. In neuroimaging, priming effects usually manifest as decreased activation with a repeated stimulus

(repetition priming), or to a similar stimulus (for review, see Schacter & Buckner, 1998; Wiggs & Martin, 1998). Perceptual priming effects arise when the manipulation of the stimulus between study and test is restricted to the perceptual attributes of the stimulus, for instance, a chair seen from two different viewpoints. Conceptual priming takes place between stimuli that are related conceptually. For instance, chairs, couches, and stools are perceptually distinct, but are conceptually related, because they are all used for sitting. Two studies that compared conceptual and perceptual priming (Koutstaal et al., 2001; Vuilleumier et al., 2002) found a common site of activation in the inferior frontal cortex that was affected by conceptual priming, but not perceptual priming. One concern with this method, however, is that conceptually similar items are often perceptually similar (e.g., a stool is more similar to a chair than to a lamp) and consequently, a conceptual priming effect may actually reflect a combination of conceptual and perceptual factors, or perhaps perceptual factors alone.

Despite the methodological concerns with the two sets of studies reviewed above, the findings converge to suggest that an important neural processing site for non-visual information associated with objects exists in the inferior frontal cortex. Using a third methodology for dissociating visual and non-visual contributions to object recognition, we recently provided further evidence that the inferior frontal cortex is important for the processing of non-visual information (James & Gauthier, 2004). In this experiment, subjects associated verbal information with sets of objects (Figure 3). In the semantic condition (SEM), the information was a nickname, such as Carl, and three non-visual semantic features, such as strong, soft and friendly. In the name condition (NAM), the information was a first, middle and last name, such as Michael Francis Sutherland. Thus, each object in these two conditions could be individuated by the verbal information with which it was associated. In the SEM condition, however, this information not only individuated the objects, it also carried strong semantic associations. Two significant benefits of our conceptual learning paradigm over existing methods of studying semantic associations are: 1) the conditions are equated on familiarity, and 2) the semantic features and the visual features of the objects are related arbitrarily, and therefore are not confounded.

During neuroimaging, subjects performed a same/different perceptual match task on pairs of objects from within the same condition shown at a learned orientation. Because of the ease of performing the perceptual match task, it was highly improbable that subjects explicitly attempted to recall the verbal information that was associated with the objects to help individuate them. It is more likely that they relied only on the visual information. Despite not explicitly using the verbal

information, a comparison of the SEM and NAM conditions showed significantly greater activation of the inferior frontal cortex for the SEM condition (Figure 3).

This finding converges with the studies described above, and suggests that the inferior frontal cortex plays a role in processing semantic associations with object stimuli. Semantic associations are recruited during object recognition, even when the information is not necessary to perform the task. It is worthwhile mentioning that the novel semantic associations in this study were developed over a period of a few days, yet our findings converge with evidence from objects for which semantic associations were developed over a lifetime of experience. Therefore, the role that the inferior frontal cortex plays in processing semantic associations seems to be flexible with respect to the learning of new information.

2.3 Conclusions

The evidence described here suggests that experts can make use of both perceptual and conceptual information to classify objects. Conceptual information appears to be recruited even for tasks that are traditionally thought to be primarily perceptual. The findings also strengthens the claim that the learning goal at least partially determines the type of perceptual processes that are recruited during learning. That is, experts will use the perceptual features of objects in ways that maximize their chances of success, given a particular goal. The combination of different kinds of knowledge, and goals, may lead to the engagement of different perceptual processes in experts. For instance, given a wealth of conceptual information with which to aid identification, preliminary evidence suggests that birders and ornithologists still appeared to recruit different perceptual processes in pursuit of their task goals. Finally, our findings converge with other methodologies, showing that an area of the inferior frontal cortex is automatically recruited during object recognition. This finding suggests that a common executive process is may be recruited when non-visual information is available to aid recognition.

3 Knowledge Types

A growing body of evidence supports the claim that some important components of conceptual knowledge, specifically the knowledge associated with the perceptual and functional attributes of object concepts, is represented as distributed patterns of activity across multiple modality-specific processing pathways in the brain. This idea is not new (Lissauer, 1890), but the data required to support it have become much more readily available with the advent of modern neuroimaging techniques. The basic claim is that cognition is embodied, grounded in

both perception and action, and that conceptual knowledge is stored in brain regions that overlap with those involved in acquiring and using that knowledge (Barsalou, 1999; Damasio, 1989; Warrington & Shallice, 1984). These brain regions are specialized for processing specific types of information (e.g., form, colour, taste, function, etc.), and thus conceptual knowledge for a concept is distributed across these regions to the extent that each knowledge type is relevant to the representation of that concept (e.g., taste is salient for foods, but not vehicles). Current research is focused on determining which modality-specific pathways are important in conceptual representation, how they differ from perceptual representations, where they are in the brain, and how they interact with the known functional organization of the brain.

Neuroimaging studies have provided a wealth of evidence supporting the idea that perceptual and conceptual knowledge are stored in distinct, overlapping brain regions, and that different types of knowledge are stored in different modality-specific processing pathways (for review, see Martin, 2007). In one of the first such studies (Martin, Haxby, Lalonde, Wiggs, & Ungerleider, 1995), participants were presented with either black and white pictures of common objects, or the written names of those objects, and asked them to generate words for action or colour associates. Generating actions, relative to colors, led to heightened activity in several regions that were common for both pictures and words, including left middle temporal gyrus (pMTG) just anterior to the primary visual motion processing area. Color word generation, relative to action words, activated the fusiform gyrus anterior to regions associated with color and object perception. More recent work has extended these findings showing that as the task becomes less perceptual and more conceptual (e.g., verifying properties of concepts presented in written form), activation can be observed extending from early occipital areas into temporal areas thought to be involved in processing conceptual knowledge (Beauchamp, Haxby, Jennings, & DeYoe, 1999; Simmons, Ramjee, McRae, Martin, & Barsalou, 2006). This overlap has been observed for several types of knowledge, including visual, auditory, tactile, and gustatory information (Goldberg, Perfetti, & Schneider, 2006).

3.1 Analysis of Feature Norms

A useful source of information about the modalities important in object concept representations, and how salient each knowledge type is for different categories of concepts, is analysis of the verbal features people use to describe concrete nouns. For example, Cree and McRae (2003) classified a large set of verbal feature production norms for 541 concepts into two different knowledge type taxonomies. The first taxonomy came from an independent source (Wu

& Barsalou, personal communication) and was inspired by a cognitive analysis of how objects are described and used in language. The second was inspired by neuropsychological and neuroimaging evidence regarding possible brain-based modality-specific sensorimotor processing pathways, and included 9 knowledge types: visual form and surface properties, visual color, visual motion, sound, smell, taste, touch, function, and encyclopaedic knowledge. Cree and McRae used hierarchical cluster analyses to examine the relative salience of the different knowledge types in concepts from each of 34 different categories. Remarkably similar results were found using the two knowledge type taxonomies. Of most interest, they found that the groupings of categories observed in the cluster analyses reflected the patterns of impairment observed across patients with category-specific semantic deficits. The most striking result was that the categories clustered roughly into three major domains: animals, fruits and vegetables, and nonliving things. These are also the three major clusters of categories likely to be impaired/spared in patients (Capitani, Laiacona, Mahon, & Caramazza, 2003). Furthermore, in the cluster analysis, the musical instruments category clustered with living things, and foods clustered with nonliving things. These two patterns were again similar to what is often observed in the patient literature (for summary, see Cree & McRae, 2003). These findings validated the brain-based knowledge type taxonomy that had been employed, and provided support for the idea that conceptual knowledge may be distributed across these knowledge types in the brain such that damage to specific clusters of knowledge type pathways could lead to the patterns of impairment observed in patients. Thus, feature norms should provide a valuable tool for designing experiments that can probe for the existence of modality-specific sensorimotor knowledge-type pathways in the brain using verbal features.

3.2 fMRI studies of Feature Learning

In an earlier section, we described a conceptual learning paradigm in which words describing semantic features were associated with novel objects during learning. This conceptual learning paradigm provides a unique methodology for exploring the brain regions involved in representing modality-specific sensorimotor conceptual knowledge. In addition to the benefits of the conceptual learning paradigm described above – 1) the conditions are equated on familiarity, and 2) the semantic features and the visual features of the objects are related arbitrarily and therefore not confounded – there is another benefit of using conceptual learning to study knowledge types. As described in the preceding section, most categories of objects are represented across several knowledge types. This makes studying knowledge types with

common objects difficult, because all categories of objects are described by several types of knowledge. The conceptual learning paradigm removes this constraint, because the types of knowledge that are associated with a set of objects are arbitrary. Thus, a set of objects can be trained with information from one, and only one, knowledge type, or can purposefully be trained with combinations of knowledge types. In the following paragraphs, we describe two experiments in which we attempted to isolate the neural substrates involved in modality-specific semantic processing using the conceptual learning paradigm.

In our first experiment (James & Gauthier, 2003), subjects associated verbal information from two separate knowledge types with two sets of objects. One set of objects was trained with auditory or “sound” features (SND) and the other set of objects was trained with visual motion or “action” features (ACT; Figure 4). According to the neural systems knowledge-type taxonomy described above, cortical regions involved in the perceptual processing of sounds should be involved in processing semantic sound features. Likewise, cortical regions involved in the perceptual processing of biological motion should be involved in processing semantic action features.

During neuroimaging, subjects performed “localizer” tasks aimed at locating brain regions involved in processing environmental sounds and brain regions involved in processing biological motion. For the sound localizer, environmental sounds were compared with the same sounds phase-scrambled in Fourier space. An area of the superior temporal gyrus (STG) showed greater activation for intact than scrambled sounds (Figure 4). For the biological motion localizer, point-light displays of people performing various movements were compared with scrambled point-light displays. An area of the posterior superior temporal sulcus (STSp) showed greater activation for intact than scrambled biological motion displays.

The subjects also performed a same/different perceptual matching task on pairs of objects from the same learning set. Again, even though the subjects did not explicitly use the verbal information to individuate the objects and aid their performance on the perceptual matching task, a comparison of the different learning conditions produced significant differences in brain activation. In the STG, the SND condition produced significantly greater activation than the ACT condition; in the STSp, the ACT condition produced significantly greater activation than the SND condition (Figure 4).

In our second experiment, we attempted to broaden the scope of the neural systems knowledge-type taxonomy to categories other than sensorimotor. Much of our conceptual knowledge of the world, and especially of other human beings, is affective or social in nature.

We are able to readily perceive affective nuances in facial and vocal expression and in body posture. Our social interactions with individuals, or groups of other humans are influenced by our previous experiences with those individuals or groups. To investigate the possible neural substrates underlying these social semantic feature types, we had subjects associate social personality traits with one set of objects, and inanimate properties with another set of objects (Figure 5). Our intention was to make the first set of objects seem like living creatures to the subjects, complete with unique personalities, and social predispositions. The second set of objects would seem like non-living statues (or perhaps bookends or doorstops), but again, with characteristics that made each one unique.

Our hypothesis was based on the work of Shultz and colleagues (Schultz et al., 2003). By comparing a social attribution task with a control task, they were able to show the brain networks involved in social information processing. Although many of these brain regions overlap with other work on social cognition (Adolphs, 2001), the key contribution of their study was the finding that the fusiform gyrus is involved in social attribution. The fusiform gyrus is best known for its specialized role in the recognition of faces (Kanwisher, McDermott, & Chun, 1997), and other objects of expertise (Gauthier, Skudlarski, Gore, & Anderson, 2000), but it is worthwhile considering its role in social cognition, and how this overlaps with perceptual and conceptual expertise. If, as we argue, objects are not just processed using visual information, but also conceptual knowledge associated with the object, then perhaps the fusiform gyrus does not represent a purely perceptual stage in visual processing, but instead represents a conceptual stage of object processing.

Based on this premise, we were not surprised when we compared objects trained with social traits with objects trained with inanimate characteristics, and found that the fusiform gyrus evoked stronger activation with social objects (Figure 5). Although there is considerable evidence that parts of the fusiform gyrus are heavily recruited for perceptual processing, our findings converge with others to suggest that the fusiform gyrus may be involved in more than just perceptual processing. Specifically, areas of the fusiform gyrus may be specialized for processing the social and affective importance of objects.

Our results also suggest that the view of cognition embodied only in sensorimotor processes is likely too limited. This bias probably comes from the heavy focus on object concepts in most behavioral, neuroimaging, and feature norming studies of semantic memory. By including the neural systems involved in affective and social cognitive processes, perhaps we edge closer to explaining conceptual processing in full.

3.3 Conclusions

The evidence reported here converges on the idea that knowledge can be categorized into different types and that these types are represented and/or processed in different neural structures. Thus, knowledge types are a useful construct for understanding the organization of semantic memory. Furthermore, at this stage of our understanding, using knowledge types based on neural processing systems may be more predictive, in terms of elucidating mind/brain functioning, than knowledge types based on more abstract cognitive categories.

4 Summary

In this chapter we have presented one view, supported by convergent empirical findings of behavioral and neuroimaging experiments, of the relationships between perceptual and conceptual cognitive and neural processes. The results of these experiments have prompted us to make two controversial claims. First, early visual perceptual processes are coupled with other higher-level processes involved in the representation of object concepts. That is, vision is not encapsulated from other cognitive processes. Second, the neural processes involved in conceptual processing are organized according to a sensorimotor framework. That is, the sensorimotor neural machinery involved in perceptual processing is also involved in conceptual processing. Of course, the number of experiments on which these arguments is based is small. Thus, a complete understanding of the complex interplay of conceptual and perceptual processing, both on the cognitive and the neural level, will require much more investigation.

5 Acknowledgements

This research was supported by grants from the NSF, NIH, JSMF, NSERC, and CIHR. Cindy Bukach, Lu Bukach, Danielle Brown, and Danielle Nikolaiczky helped with fMRI data collection, and Jason Ozubko with the birder/ornithologist behavioral data. Daniel Bub, Isabel Gauthier, Bob Shultz, Karin Harman James, Cindy Bukach, and other members of the Perceptual Expertise Network gave helpful advice and comments.

6 References

- Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinion in Neurobiology*, 11, 231-239.
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-660.

- Beauchamp, M. S., Haxby, J. V., Jennings, J. E., & DeYoe, E. A. (1999). An fMRI version of the Farnsworth-Munsell 100-Hue test reveals multiple color-sensitive areas in human ventral occipitotemporal cortex. *Cerebral Cortex*, 9, 257-263.
- Biederman, I. (2000). Recognizing depth-rotated objects: a review of recent research and theory. *Spatial Vision*, 13(2-3), 241-253.
- Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 13, 640-645.
- Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: the power of an expertise framework. *Trends in Cognitive Sciences*, 10(4), 159-166.
- Capitani, E., Laiacona, M., Mahon, B., & Caramazza, A. (2003). What are the effects of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology*, 20, 213-261.
- 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(2), 163-201.
- Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, 33(1-2), 25-62.
- De Groot, A. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem Solving: Research, method, and theory*. New York: Wiley.
- Deco, G., & Rolls, E. T. (2004). A neurodynamical cortical model of visual attention and invariant object recognition. *Vision Research*, 44(6), 621-642.
- Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107-117.
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197.
- Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: exploring mechanisms for face recognition. *Vision Research*, 37(12), 1673-1682.
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, 28(2), 431-446.
- Goldberg, R. F., Perfetti, C. A., & Schneider, W. (2006). Perceptual knowledge retrieval activates sensory brain regions. *Journal of Neuroscience*, 26, 4917-4921.
- Goldstone, R. L., & Barsalou, L. W. (1998). Reuniting perception and conception. *Cognition*, 65, 231-262.
- Goldstone, R. L., Steyvers, M., Spencer-Smith, J., & Kersten, A. (2000). Interactions between perceptual and conceptual learning. In E. Diettrich & A. B. Markman (Eds.), *Cognitive Dynamics: Conceptual Change in Humans and Machines*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Gough, G. A., Sauer, J. R., & Iliff, M. (1998). Patuxent Bird Identification Infocenter (Version 97.1). Laurel, MD: Patuxent Wildlife Research Center.
- Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology*, 13, 1-8.
- Grossberg, S. (1999). The link between brain learning, attention, and consciousness. *Conscious Cognition*, 8(1), 1-44.
- Humphreys, G. W., & Forde, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences*, 24(3), 453-509.

- James, K. H., James, T. W., Jobard, G., Wong, C.-N., & Gauthier, I. (2005). Letter processing in the visual system: Different activation patterns for single letters and strings. *Cognitive, Affective and Behavioral Neuroscience*, 5(4), 452-466.
- James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology*, 13, 1792-1796.
- James, T. W., & Gauthier, I. (2004). Brain regions engaged during visual judgments by involuntary access to novel semantic information. *Vision Research*, 44, 429-439.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17, 4302-4311.
- Koutstaal, W., Wagner, A. D., Rotte, M., Maril, A., Buckner, R. L., & Schacter, D. L. (2001). Perceptual specificity in visual object priming: functional magnetic resonance imaging evidence for a laterality difference in fusiform cortex. *Neuropsychologia*, 39, 184-199.
- Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *The Journal of Neuroscience*, 20(2), 878-886.
- Lissauer, H. (1890). Ein Fall von Seelenblindheit nebst einem Beitrag zur Theorie derselben. *Archiv fur Psychiatrie*, 21, 222-270.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman.
- Martin, A. (1999). Automatic activation of the medial temporal lobe during encoding: lateralized influences of meaning and novelty. *Hippocampus*, 9, 62-70.
- 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., & Ungerleider, L. G. (1995). Discrete cortical regions associated with knowledge of color and knowledge of action. *Science*, 270, 102-105.
- Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-Ray image processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(3), 553-557.
- Ozubko, J., Cree, G. S., & Bub, D. (2005). *Not all Visual Experts are Created Equal*. Paper presented at the Fifteenth Annual Meeting of the Canadian Society for Brain, Behavior and Cognitive Science.
- Peron, R. M., & Allen, G. L. (1988). Attempts to train novices for beer flavor discrimination: A matter of taste. *The Journal of General Psychology*, 115, 403-418.
- Pylyshyn, Z. (1999). Is vision continuous with cognition? The case of impenetrability of visual perception. *Behavioral and Brain Sciences*, 22, 341-423.
- Schacter, D. L., & Buckner, R. L. (1998). Priming and the brain. *Neuron*, 20(2), 185-195.
- Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., et al. (2003). The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society of London: Series B*, 358, 415-427.
- Shyns, P. G., Goldstone, R. L., & Thibaut, J.-P. (1998). Development of features in object concepts. *Behavioral and Brain Sciences*, 21, 1-54.
- Sibley, D. A. (2000). *The Sibley Guide to Birds*. New York: Algreed A. Knopf, Inc.
- Simmons, W. K., Ramjee, V., McRae, K., Martin, A., & Barsalou, L. W. (2006). fMRI evidence for an overlap in the neural bases of color perception and color knowledge. *Neuroimage*, 31, S182.

- Tarr, M. J., & Bulthoff, H. H. (1998). Image-based object recognition in man, monkey and machine. *Cognition*, 67(1-2), 1-20.
- Vuilleumier, P., Henson, R. N., Driver, J., & Dolan, R. J. (2002). Multiple levels of visual object constancy revealed by event-related fMRI of repetition priming. *Nature Neuroscience*, 5(5), 491-499.
- Warrington, E. K., & Shallice, T. (1984). Category-specific semantic impairments. *Brain*, 107, 829-854.
- Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology*, 8, 227-233.
- Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, 18, 221-281.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, 81, 141-145.

7 Figure captions

Figure 1. Taxonomical Accuracy for Consistent and Inconsistent Triads of Birds from Experiment 1 (behavioral).

Figure 2. Recognition Accuracy for Inverted and Upright Birds from Experiment 2 (behavioral).

Figure 3. Stimuli and results for the Semantic (SEM) and Name (NAM) conditions from Experiment 1 (fMRI).

Figure 4. Stimuli and results for the Auditory (AUD) and Action (ACT) conditions from Experiment 2 (fMRI). Height of bars indicate percent signal change in Superior Temporal Gyrus (STG; primary & secondary auditory cortex) and Superior Temporal Sulcus (STS).

Figure 5. Stimuli and results for the Social (SOC) and Inanimate (INA) conditions from Experiment 3 (fMRI). Clockwise from the top-left are sagittal, coronal, and axial slices of an average structural MRI.

Figure 1

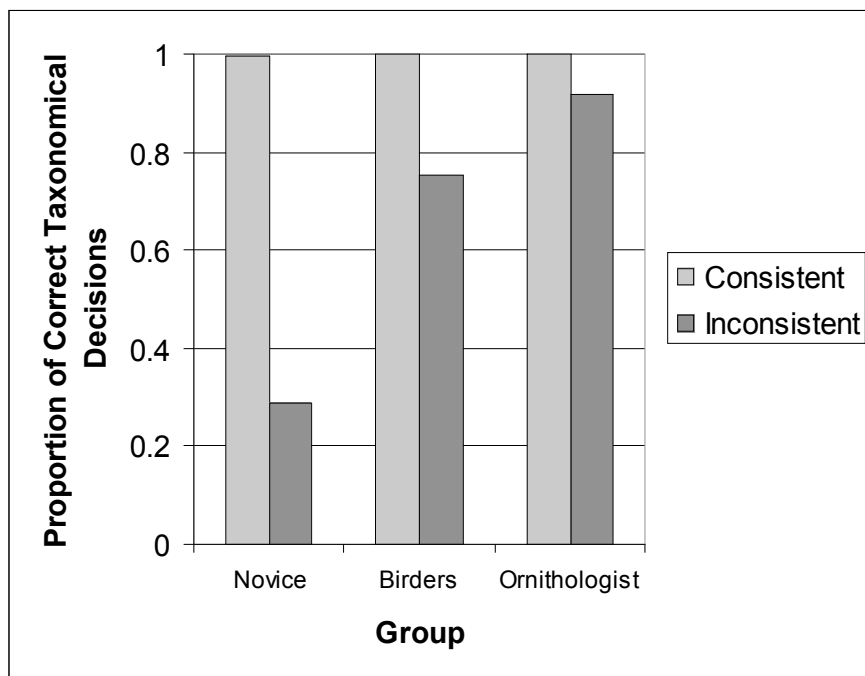


Figure 2

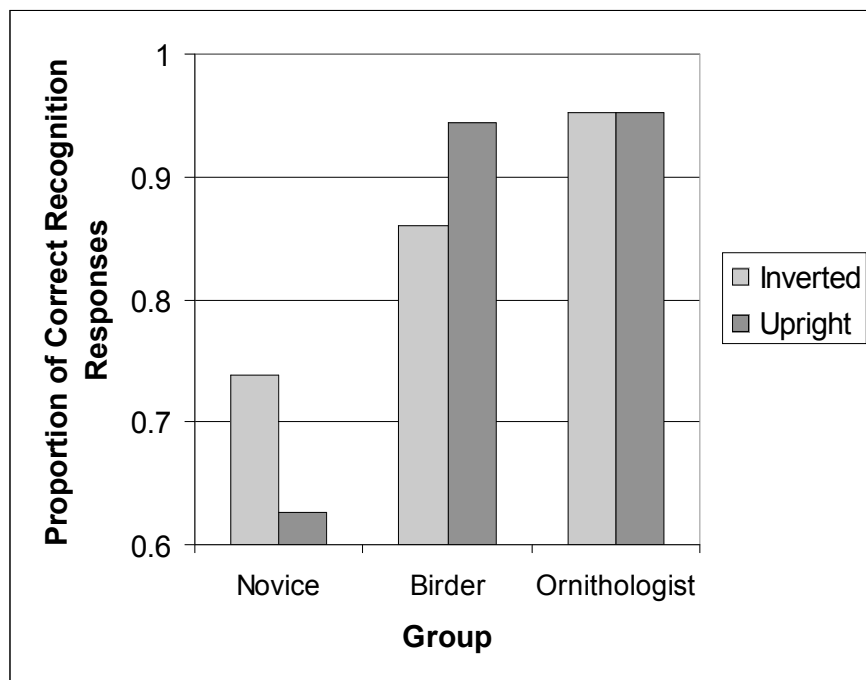


Figure 3

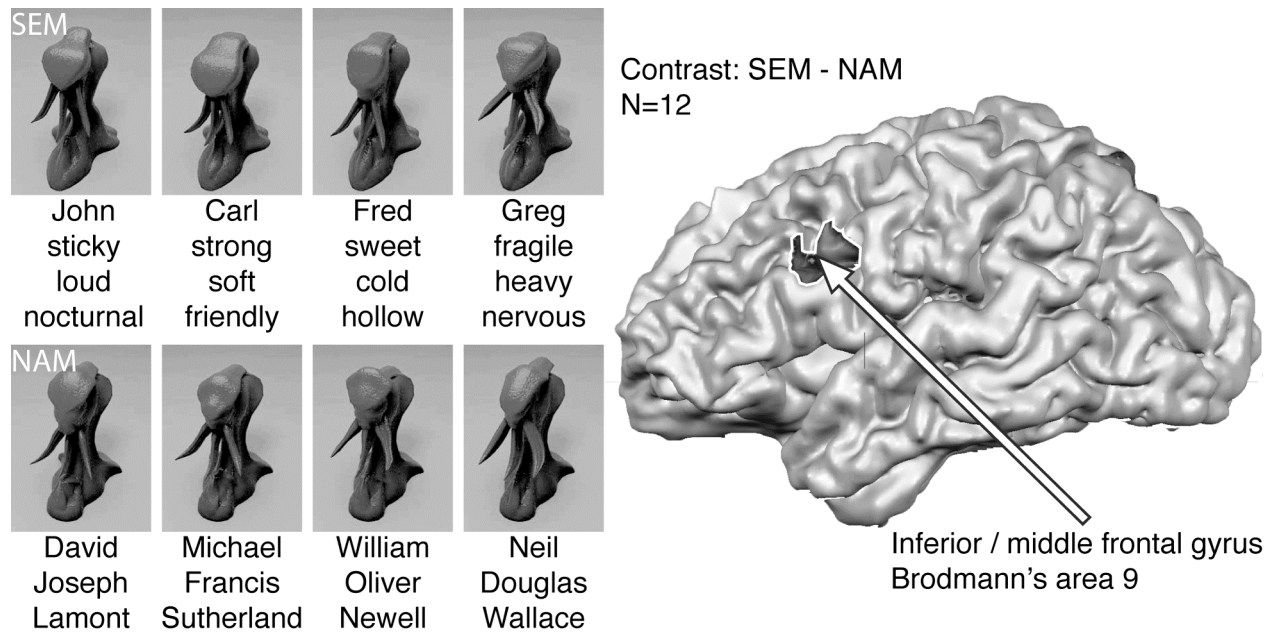


Figure 4

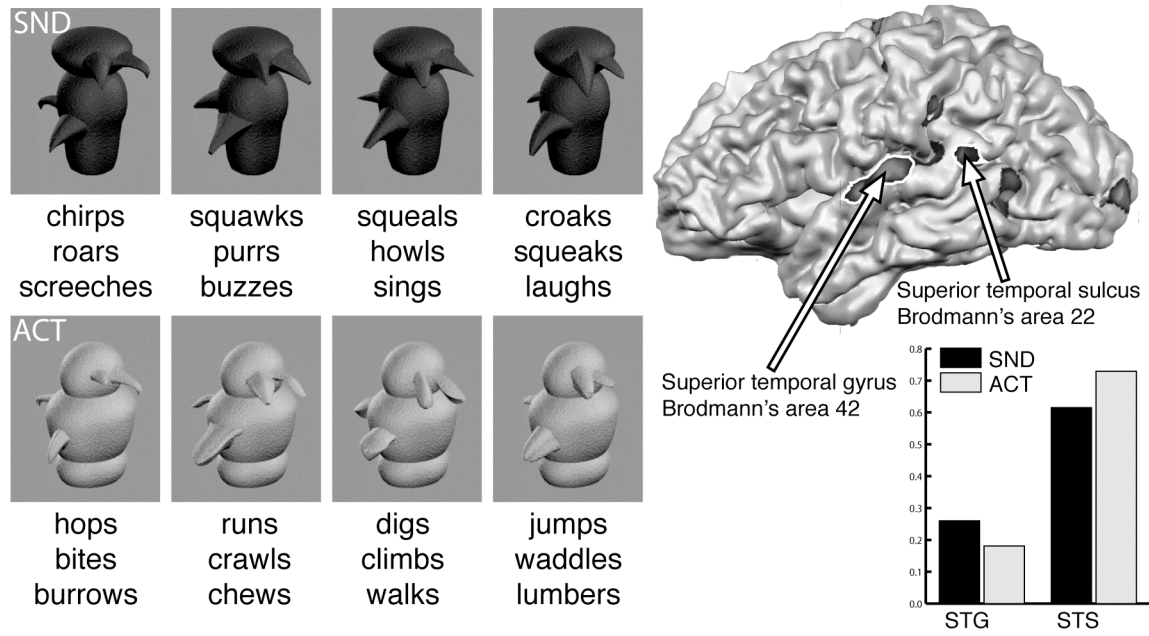


Figure 5

