THE INTERACTIONS BETWEEN VISUAL APPEARANCE AND CONCEPTUAL KNOWLEDGE IN THE ACQUISITION OF PERCEPTUAL EXPERTISE

By

Olivia Sin-Chi Cheung

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirement

for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

August, 2010

Nashville, Tennessee

Approved:
Professor Isabel Gauthier
Professor Timothy P. McNamara
Professor James W. Tanaka
Professor Frank Tong

ACKNOWLEDGEMENTS

I am extremely grateful to my advisor, Isabel Gauthier, for her complete trust and support during my graduate career and especially during the preparation for this dissertation. I am also indebted to my committee, Tim McNamara, Jim Tanaka and Frank Tong, for their invaluable comments on and their interest in the work reported in this dissertation. Thanks also to Jennifer Richler, Yetta Wong, Rankin Williams McGugin, Jonathan Folstein, Magen Speegle, Sarah Muller, Daryl Fougnie, Eunice Yang and Min-Suk Kang for help in various aspects of this project.

I would also like to thank past and present members in the Object Perception Laboratory, the Category Laboratory, the Perceptual Expertise Network and the Human Information Processing Laboratory for many simulating discussions on research. I feel very fortunate to be able to work with such inspiring and friendly researchers.

On a personal level, I thank Daryl Fougnie for things too numerous to enumerate here. Finally, I give many thanks to my parents and my sister for their support and encouragement.

TABLE OF CONTENTS

Page

Chapter

Appendix

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Our ability to recognize objects in the world is remarkable: we can recognize an object under various viewing conditions, regardless of variations in viewpoint, size, lighting, or location. Also remarkable is the amount of information we can accumulate about an object. Apart from its physical appearance, we acquire other semantic, often abstract, knowledge regarding separate modalities, functions, contextual setting, and valence. In this dissertation, I address this specific question: to what extent is object recognition constrained by the *interactions* between visual and conceptual information? There are many aspects of shape that may influence how we learn conceptual information about objects. Objects can possess affordances that may facilitate learning information that is consistent with these affordances (Gibson & Walker, 1984). For instance, it makes sense to be given information about how to grasp a hand-sized tool-like object, whereas the same information may also be associated with a large boulder, but arguably in a very different manner. In addition, conceptual information can build on experience. For example, some of the properties we are willing to attribute to computers are not appreciable as direct affordances but rather depend on our knowledge of electronic equipment. In this project, the large and important difference between animate and inanimate objects is considered. Note that our expectations concerning the semantic attributes that fit these two broad categories and how they correlate with shape can be both innate and acquired. Regardless of the sources of our expectations, the experiment outlined here will consider whether and how such expectations can constrain the way people learn to individuate objects from novel categories.

In this chapter, I will first highlight relevant literature related to the roles of shape and conceptual knowledge in object recognition. I will then provide an introduction to the design and goals of the current study.

The roles of visual and conceptual properties in object recognition

Much research on object recognition has focused on the processing of visual features (Biederman, 1987; Humphreys & Forde, 2001; Tarr & Bülthoff, 1998; Tarr, Kersten, & Bülthoff, 1998). Since the 1980's, it has been suggested that object recognition is a bottom-up visual process that occurs independently from other cognitive operations (Marr, 1982; Pylyshyn, 1999). Essentially, several influential object recognition models emphasize exclusively the role of shape or visual features (e.g., recognition-by-component theory, Biederman, 1987; Biederman & Gerhardstein, 1993; view-based theories, Tarr, 1995; Bulthoff & Edelman, 1992; hierarchical models, Perrett & Oram, 1993; Riesenhuber & Poggio, 1999). However, there is also evidence that (1) object recognition involves much more than dissecting an object into its visual features and that (2) conceptual knowledge about object categories is also represented in the visual system (e.g., living vs. non-living things, Chao et al., 1999; Mahon & Caramazza, 2009; Martin, 2007). For this reason, a great deal of work on the nature of the representations and mechanisms involved in object recognition has been done with novel objects. For instance, Tarr & Pinker

(1989) used novel objects to contrast predictions based on viewpoint-dependent vs. structural description theories in order to eliminate the confound of experience with multiple viewpoints, while assuming that the lack of conceptual associations was irrelevant towards testing these theories. However, this assumption may be invalid, as I will demonstrate in this study that object recognition is influenced by how an object looks and the inferences afforded by visual properties.

Since objects in various categories often differ in how they look, it is obvious that visual appearance is part of the knowledge we have about objects. However, conceptual knowledge is traditionally thought to be amodal, propositional and verbal in nature (see Glaser, 1982 for a review). Some suggested that perceptual and conceptual representations of objects should be distinguished (Mandler, 2000; Carey, 2000): perceptual processing computes perceptual similarity between objects, whereas conceptual processing consists of the mental transformation of perceptual information into concepts (Mandler, 2000). Nonetheless, while conceptual knowledge about objects can be acquired through verbal associations (Nelson, 1974; Nelson et al., 2008), it can also be grounded on perception of physical properties and affordances (Gibson, 1988; Goldstone & Barsalou, 1998; Graham & Poulin-Dubois, 1999; Jones & Smith, 1993). In particular, perceptual cues are critical to differentiate animate vs. inanimate categories even for infants and young children (e.g., Keil, 1991; Massey & Gelman, 1988). While object motion is a salient cue for animacy (Gao et al., 2009; Gelman & Spelke, 1981; Scholl & Tremoulet, 2000; Tremoulet & Feldman, 2000; Wheatley et al., 2007), object shape is also particularly useful in detecting animacy (Booth & Waxman, 2002; Jones et al., 1991; Jones &

Smith, 1998). For instance, bilateral face or body symmetry appears to be a natural and crucial indicator of animacy. Facial and body symmetry signals good health and developmental stability in humans (Livshits & Kobyliansky, 1991; Rhodes et al., 2001) and nonhuman animals (Parsons, 1990; Polak, 2003). Therefore, symmetry advertises mate quality as symmetrical faces or bodies are more attractive than asymmetrical faces (Concar, 1995; Gangestad & Simpson, 2000; Thornhill & Gangestad, 1994; 1999). Interestingly, symmetrical faces are also more likely to be perceived as having better skin color and textures than asymmetrical faces (Jones et al., 2004). In fact, healthy facial color and textural cues alone are sufficient to increase facial attractiveness (Jones et al., 2004; Perrett et al., 1999; Rhodes et al., 1998; see also Rhodes et al., 2005). On the other hand, perception and manipulation of tools or man-made objects also emerge at an early age (e.g., McCarty et al., 2001). It requires extraction of sensory information about object properties and what they afford (e.g., a hammer affords pounding; Gibson, 1978), as well as reasoning for how the objects should be used (Lockman, 2000; see also Osiurak et al., 2010). Depending on their functions, the shape of man-made objects varies greatly and bilateral symmetry is a less relevant visual cue for inanimate than animate categories.

Although there is no doubt that there exists a close relationship between perceptual and conceptual properties of objects, little is known about how perceptual and conceptual representations interact during visual processing of objects. Specifically, how does the visual appearance of objects affect the learning of new concepts about the objects, and how do learned conceptual associations influence perception of objects that differ in shape?

Possible influences between visual and conceptual properties

Rapid interactions likely occur between visual and conceptual representations of objects during visual processing. Even when we encounter an object for the first time, we likely make inferences depending on visual properties such as shape and material. Just at a glance, smooth-shaped objects are preferred over sharp objects (Bar & Neta, 2006), as sharp objects indicate potential threat (Bar & Neta, 2007). Other visual information such as symmetry, contrast, and complexity also influence our preferences (Reber et al., 2004). Moreover, given a minimal amount of experience, semantic associations with objects can influence the way objects are perceived. For example, there is considerable evidence that just the act of placing objects into categories can change perception by enhancing perceptual discrimination of dimensions that are diagnostic for categorization (Goldstone, 1994: Goldstone & Steyvers, 2001; Op de Beeck et al., 2003). Critically, however, the meaning of the categories we create may also matter. Assuming we divide a group of novel objects into two categories, does it matter whether we call objects on one side "animals" and objects on the other side "tools"? Doing so would not change what dimensions are diagnostic for the categorization of the objects, but by anchoring categories into our existing conceptual networks, different features may become relevant within each category. Indeed, specific semantic associations can bias interpretation of neutral stimuli (Bentin & Golland, 2002; Wheatley et al., 2007;

Hilliar & Kemp, 2008), facilitate perceptual categorization (Lin & Murphy, 1997; Palmeri & Flanery, 1999; Wisniewski & Medin, 1994), and improve discrimination in RSVP and visual search tasks (Dux & Coltheart, 2005; Lupyan & Spivey, 2008; Telling et al., 2010).

The facilitation of visual processing by semantic information may reflect a “perceptual-semantic continuum” in object representations (Sergent & Poncet, 1990; Sergent & Signoret, 1992; Young et al., 1989; but see Newcombe et al., 1989). According to this idea, psychological similarity between objects would depend on both visual and conceptual properties. While objects that share visual features are more difficult to discriminate than objects that have distinct visual features, objects that are closely related semantically are also more difficult to discriminate compared to objects that are loosely related semantically (Dixon, Bub & Arguin, 1997; 1998; Schweizer & Dixon, 2006). For instance, shapes or faces are harder to discriminate if they are associated with semantically similar sounds (e.g., owl's hoot, crow's caw, robin's song) compared to semantically distinct sounds (e.g., banjo, robin etc, Dixon et al., 1997). Similar effects arise when objects are paired with words from the same vs. different categories (Dixon et al., 1997), when faces are paired with names of people in the same vs. different professions (Dixon et al., 1998), or when novel objects are associated with overlapping vs. distinct concepts (Gauthier et al., 2003a). Given this sort of evidence, to try to understand perception outside of the context of conceptual processing would appear to be a fool’s errand.

Brain imaging provides us with a window into the creation of associations between shapes and concepts. Implicit activation of conceptual knowledge in various brain areas during visual object processing has been demonstrated in a series of studies where different types of semantics were associated with novel objects (Gauthier et al., 2003a; James & Gauthier, 2003; 2004; Bukach et al., unpublished). In a short training paradigm, participants learned to associate different non-visual attributes with a few novel 3D objects (i.e., Greebles or Yufos; Figure 1). In James and Gauthier (2003), participants learned associations for three sets of four novel objects (Greebles), attributing a different type of semantic features for each category. That is, each object was associated with three words that described either auditory (e.g., squawks, purrs, buzzes), action (e.g., runs, crawls, chews) or encyclopedic (e.g., long-lived, sly, gentle) attributes. At test, participants were required to visually discriminate pairs of Greebles in a simultaneous matching task. Although semantic information was task-irrelevant, it nonetheless influenced the neural responses to the objects.


Figure 1. Sample novel 3D objects used in Gauthier et al. (2003) and James & Gauthier (2003; 2004). Left: a Greeble. Right: a Yufo.

During this purely visual task, auditory and action semantics were engaged, with greater activity in the sound-responsive superior temporal sulcus (STS) areas when Greebles associated with auditory attributes were shown, and greater activity in motion-responsive STS areas when Greebles associated with action semantics were presented. Using similar training procedures, brain areas implicated in semantic processing (e.g., left inferior frontal lobe) were also more active during visual judgments for objects associated with concepts from various domains (e.g., Fred, sweet, cold, hollow) compared to objects that were only associated with names (e.g., Jonathan Wesley Abraham; James & Gauthier, 2004). These results suggest that semantic representations can be automatically engaged during a visual task. However, it is unknown to what extent semantics can be activated as a result of the visual properties of the objects implying certain relevant semantic associations (e.g., auditory vs. action descriptions applied to creature-like objects).

Even more relevant to the study presented in this dissertation was Bukach et al. (unpublished). In this study, participants learned to associate social (e.g., friendly, shy, smart) or inanimate features (e.g., matte, sharp, made of clay) to 3D novel objects (Greebles). During a visual task in the scanner, the objects associated with social semantics showed higher activity in the "fusiform face area" (FFA) than the objects associated with inanimate semantics. The FFA responds more to faces compared to other common objects (Kanwisher et al., 1997; Puce et al., 1996; but see Grill-Spector et al., 2006). But critically, novel objects not shown during the training phase showed comparable FFA activity with the objects associated with social semantics. It was speculated that the response in FFA for the untrained objects was to some extent influenced by their symmetrical shape, leading these objects to be more easily interpreted as social entities. This study does not tell us whether this assumption would have been made prior to any experience with similar objects. An alternative is that, because all the objects in this study were animate-like novel objects with a common configuration of parts, the experience of associating social semantics with some of the objects was in some way more powerful than the experience of associating inanimate semantics with the other subset of objects, and dominated the inferences made about novel exemplars. These two competing interpretations of the results from Bukach et al. (unpublished) highlights how little we know about how semantic information acquired for single objects influences perceptual and conceptual processing at the category level.

Effects of perceptual experience

Apart from visual and conceptual properties, another critical factor that has been proposed to account for activity in the ventral visual system is experience with different perceptual strategies. According to the process-map hypothesis (Bukach et al., 2006; Gauthier et al., 2000; Tarr & Gauthier, 2000), the visual system learns to recruit the best neural substrates to fulfill the specific perceptual demands associated with different object categories through experience. For instance, experience individuating objects in a category leads to greater activity in regions of the lateral fusiform gyrus (Behrmann et al., 2005; Gauthier et al., 1999; 2000; Xu, 2005; Wong et al., 2009b); extensive reading experience results in specialization in areas on the left occipital cortex for visually presented words and letters in proficient readers of different languages (Baker et al., 2007; Wong et al., 2009c) in addition to activity in motor areas engaged by writing (James & Gauthier, 2006); skilled readers of musical notation recruit a multimodal brain network when viewing musical notations relative to viewing letters or symbols, presumably due to the specifics of musical training (Wong & Gauthier, 2010).

How does visual appearance and conceptual knowledge facilitate and constrain object representations in the brain during the acquisition of expertise for a category of objects? One interesting locus where the interactions may occur is in the fusiform face area (FFA). It is unclear what factors cause faces to be "special" in recruiting this area. Some have suggested that the evolutionary importance of faces explains the existence of a face-specific cortical module, and that face-like symmetrical appearance is critical to engage the "face processer" (Kanwisher, 2000; Yovel & Kanwisher, 2004).

At least two separate sets of findings converge to implicate abstract social semantics in the FFA. First, individuals with Autism, whose primary deficit is one of social functioning, show hypoactivation in the FFA during face perception (Grelotti et al., 2002; 2005; Schultz et al., 2005). Second, the perception of human-like interactions among geometric shapes engages this area, as part of an entire social semantic processing network that also includes the amygdala, STS, temporal pole, medial prefrontal and inferolateral cortices (Schultz et al., 2003; Wheatley et al., 2007). It is difficult to explain both of these results solely in terms of the perceptual processes that are critical for face individuation, which are the focus of much of the discussion concerning the function of the FFA (Gauthier, 2000; Grill-Spector, Knouf & Kanwhisher, 2004).

Extensive experience individuating objects has been argued to be sufficient to account for selectivity in the FFA. Such expertise with non-face objects can recruit this area (Gauthier et al., 1999; 2000; Xu, 2005; see also Harley et al., 2009; van der Linden et al., 2009), and moderate amounts of individuation experience recruits the immediately surrounding cortex in the fusiform gyrus (Wong et al., 2009b). But is activity in this area also modulated by visual and conceptual properties of object categories? Strikingly, the expertise effect observed in the FFA has also been found with inanimate objects (e.g., cars, Gauthier et al., 2000; Xu, 2005) and novel objects that are asymmetric (Kung et al., 2007). However, no study has directly compared performance or neural activity between different categories of objects of expertise (e.g., faces vs. cars in car experts; cf. Behrmann et al., 2005): therefore, it is still unclear how visual appearance and conceptual knowledge may interact with perceptual expertise to engage the FFA.

An interesting but untested possibility is that the recruitment of the FFA for an object category (and by extension, the recruitment of any visual area) may be due to an interaction between visual appearance, conceptual knowledge and training experience. This idea has been the topic of speculations that try to account for anecdotal findings. In particular, in a study where participants were trained extensively to discriminate novel objects, greater FFA activity was observed only when an observer thought that one of the classes of novel objects used, which happened to be smooth, resembled "women wearing hats" (Op de Beeck et al., 2006). Similarly, the FFA was recruited for images with blurred blobs atop human bodies, but not for the same blobs presented in a context that did not encourage participants to think of them as faces (Cox et al., 2004). These results suggest that the processing of both visual and conceptual information of stimuli based on prior experience may contribute to an object's representation in the visual system, including the recruitment of an area such as the FFA.

The current study

The current study is unique in that it is designed to systematically investigate the interactions between visual appearance and semantic knowledge in object recognition, and to study how the interactions change or develop throughout the course of extensive training in terms of behavioral performance and neural activity in the ventral object recognition system. Because almost no work has been done that speaks directly to this question, to a large extent the hope is that this study will provide a proof of concept that such interactions exist and have implications both for behavior and for the representations of objects in the visual system. Just as the demonstration that participants could acquire expertise in the laboratory (Gauthier & Tarr, 1997) opened the way to the experimental study of perceptual expertise, I hope the methods developed here will give our field traction to address the complex nature of interactions between shape and semantics. This question is critical for understanding visual object recognition in general, for understanding the functional organization of the ventral visual system, and to test practical assumptions about our use of novel objects in experimental psychology.

Here, I use the terms “conceptual knowledge” or “semantics” to refer to relatively abstract meaning about the character of a person or the quality of a man-made object. In this study, the concepts are learned through a list of words without any interactive contact with the objects (see Gauthier et al., 2003; James & Gauthier, 2003; 2004). Some words may indicate surface properties of objects (e.g., “shiny”, “portable”), but most words may not be immediately realized from the shape of an object (e.g., “shy”, “thoughtful”, or “multipurpose”, “well-made”). These descriptions are then arbitrarily associated with a set of objects that resemble either living or non-living things. Note also that even if “cute” or “shiny” may be appropriate to describe all objects in a set, each word or concept was only explicitly associated with one of the items. Of interest is whether such conceptual information can influence perception of different categories of objects.

This study first examines whether an interaction between visual appearance and conceptual knowledge for novel objects exists prior to training, and then examines whether the interaction between these factors emerges or changes after semantic training. These effects are again examined after more extensive visual individuation training where conceptual knowledge is no longer relevant.

To examine the interactions, visual appearance and conceptual knowledge will be manipulated in a two-stage training study. Four categories of novel objects will be used, with objects in two of the categories sharing a symmetrical configuration of parts to suggest an animate category, and objects in the other two categories sharing a asymmetrical configuration of parts to suggest man-made objects or tools. In this study, I chose to use symmetry to define animate objects since faces are presumably the most critical visual stimuli involved in social interactions and symmetry is an important and desired feature for faces (Perrett et al., 1999; Rhodes et al., 1998). In contrast, man-made objects or tools are more likely to have a canonical orientation in which the objects appear asymmetrical (e.g., a hammer or a pen). Additionally, textural and color information are also critical for determining animacy (even for young children of 2-3 years of age; Booth & Waxman, 2002; Jones et al., 1991; Jones & Smith, 1998). Thus, the current animate vs.

inanimate object sets are also rendered in textures and colors appropriate for the categories. In other words, the manipulation of the visual properties of different object categories very likely implies specific conceptual categories before any semantic association is suggested by the experimenter.

It is often assumed - but not tested - that recognition of novel objects does not involve semantic processing (e.g., Bar & Neta, 2006; Op de Beeck et al., 2008a; Op de Beeck et al., 2008b). The existence of implicit semantic biases based on visual object properties will be examined here in a pre-training behavioral test.

In the first training phase of the study (semantic training), two types of verbal semantic associations will be used: social and inanimate. Two groups of participants will learn the same objects and the same semantic associations. The critical manipulation is the explicit *pairing* of the types of objects and semantics. Specifically, the Typical Pairing group learned to associate social semantics with animal-like (symmetric) objects and inanimate semantics with tool-like (asymmetric) objects, whereas the Reversed Pairing group learned to associate inanimate semantics with animal-like objects and social semantics with tool-like objects. The semantic training is based on procedures used in prior work (Gauthier et al., 2003a; James & Gauthier, 2003; 2004; Bukach et al., unpublished), in which each trained object was associated with three verbal descriptions. In other words, the experimental conditions in these studies demonstrated that associations could be created rapidly between words and objects, and that these associations carried behavioral and neural consequences, even when these associations were entirely arbitrary. Here however, I am specifically interested in the nature of the associations that are created between shapes and words when these associations are consistent vs. inconsistent with implicit biases based on object shape.

While arbitrary semantic associations can be learned for objects (Gauthier et al., 2003a; James & Gauthier, 2004; Dixon et al., 1997; 1998; Scheweizer & Dixon, 2006), it is unknown whether learning efficiency can be enhanced if the associations are expected based on the visual properties of the objects. Alternatively, it is possible that acquiring atypical semantic associations, because it is more difficult, facilitates perceptual processing through increased attention. Although some have speculated that the mechanisms through which we conceive of shape can influence visual representations (Cox et al., 2004; Op de Beek et al., 2006), the most explicit neurally plausible models of object recognition assume that semantic associations bear no influence on the nature of visual representations in the ventral visual pathway (Riesenhuber & Poggio et al., 1999; 2000). Perhaps most extreme is the claim that "…IT [inferotemporal] cortex contains a large-scale map of shape that is largely independent of meaning, familiarity, and behavioral task." (Op de Beeck et al., 2008, p.1676). In the context of such claims, it is clearly important to document interactions between shape and semantics, if found to exist.

After the semantic training, visual individuation training will be introduced. The training procedure is modeled after studies that emphasize rapid subordinate-level categorization of individual exemplars in a category (Gauthier et al., 1999; Gauthier & Tarr, 2002; Scott et al., 2006; 2008; Tanaka et al., 2005; Wong et al., 2009a). Critically, in this second training phase, participants in both Typical and Reversed Pairing groups will be trained in exactly the same way, with both animal-like and tool-like objects, and in a procedure that does not include any semantic training. The only difference between the two groups at this stage will be their distinct experience in the earlier semantic training, which may be carried into the visual training by the inclusion of the objects that had been explicitly associated with either social or inanimate semantics. One question of interest at this stage will be whether the nature of the semantic associations generalizes to some extent to new exemplars introduced during visual learning. More generally, it is interesting to ask whether the influence of the visual-semantic associations will disappear because they are no longer used, or whether they will create a conceptual bias that has lasting – or even growing – consequences, as participants acquire further expertise with object categories. Behaviorally, while recognition performance is expected to improve during the visual learning, it will be interesting to test whether the improvement will be greater for participants who received typical associations earlier on. In terms of neural activity, we can specifically address the speculation according to which the FFA, a locus of perceptual expertise, is most sensitive to objects resembling living creatures associated with social semantics (Op de Beeck et al., 2006).

A note should be made about generalization of training effects. An important aspect of perceptual expertise is the generalization of expertise effects to new exemplars in the trained categories (Gauthier & Tarr, 1997; 2002; McGugin et al., submitted; Scott et al., 2006; 2008; Tanaka et al., 2005; Wong et al., 2009a; Wong et al., 2009b) but not to exemplars in other categories (e.g., modern vs. antique cars, Bukach et al., in press; see also Grill-Spector et al., 2004; Yue et al., 2006). On the other hand, only a few studies have included conditions to examine generalization of semantic training effects in visual processing, and the results are inconclusive (Bukach et al., unpublished; James & Gauthier, 2003; 2004). In this study, generalization of the interaction effects between visual appearance and conceptual knowledge from trained objects to new objects will also be examined in two types of transfer objects. Specifically, the first type of transfer object will consist of unlearned exemplars in the trained categories, whereas the other type will be novel exemplars of what should be considered different categories based on salient differences in shape and color. To preview one of the unexpected findings of this dissertation, the results will include some evidence that these large differences in shape and color will be ignored and that similarity of individual features and their configuration will cause a non-trivial degree of generalization to these objects. These considerations will be discussed more fully in Chapter 3: they are important to the extent that they limit the degree to which I can localize category-specific expertise effects in the fMRI study.

CHAPTER II

INTERACTIONS BETWEEN VISUAL APPEARANCE AND CONCEPTUAL KNOWLEDGE IN OBJECT RECOGNITION: BEHAVIORAL MEASURES

Introduction

As discussed in Chapter 1, one of the main questions addressed in this dissertation is how visual and conceptual properties of objects interact to influence object recognition. To directly measure the interactions during object learning, I combined two training paradigms used in prior work: a semantic training procedure (Gauthier et al., 2003a; James & Gauthier, 2003; 2004) and a visual expertise training procedure (Gauthier & Tarr, 1997; Gauthier et al., 1998; Gauthier et al., 1999; Wong et al., 2009a; Wong et al., 2009b). Each participant learned semantic associations for objects and practiced individuating objects from two different categories. In the first part of the study, participants were randomly assigned to one of two groups for semantic training, in which they learned to associate social and inanimate semantic features with animal-like and tool-like novel objects, in a manner that was either consistent or inconsistent with intuition. Specifically, I assumed that social semantics were more easily expected to apply to animal-like objects than to asymmetric objects. In the second part of the study, all participants experienced the same visual individuation training in which they learned to individuate several exemplars of the two types of objects, without any mention of semantic associations.

In this chapter, I investigate the possible interactions between visual appearance and conceptual knowledge using behavioral measures. The first question I address is whether implicit associations between visual and semantic features can be revealed prior to any exposure or training with a set of novel objects. If the visual appearance of an object predicts the semantic category to which the object belongs, performance in a lexical judgment task may be affected by a simultaneously presented object, even though the object is task-irrelevant. The next question I address is how new pairings of conceptual knowledge with the objects, introduced during the semantic training, will affect any pre-existing visual-semantic biases or will produce new biases if they are not found before training. Finally, I will ask whether any behavioral changes resulting from the semantic training can be modulated by the following visual individuation training.

In addition to considering how objects become associated with social vs. inanimate categories, I am also interested in how this conceptual information may affect perceptual measures. In particular, how will visual and conceptual features contribute and interact to affect two hallmarks of perceptual expertise: the reduction of basic-level advantage and changes in holistic processing? These two effects have been demonstrated in experts of various animate and inanimate categories (e.g., dogs: Diamond & Carey, 1986; birds: Tanaka et al., 2005; Scott et al., 2006; 2008; cars: Gauthier et al., 2003b; Xu, 2005; novel 3D objects: Gauthier et al., 1998; Gauthier & Tarr, 2002; Wong et al., 2009a). However, these effects have not been directly compared across two or more different object categories (e.g., faces vs.

cars in car experts), and thus it remains unclear whether visual and/or semantic features can modulate the effects.

Both semantic and visual training phases require participants to learn to individuate objects within a visually homogenous category (although speeded individuation is only required for the visual training, and individual names are only provided for objects then). Because individuation training leads to the development of perceptual expertise, recognition performance at the subordinate-level should improve and holistic processing may emerge for the trained categories. A main goal here is to test whether certain types of visual features can facilitate subordinate-level recognition and holistic processing, and whether explicit semantic associations can also affect performance. These effects are examined in a visual matching task at the basic- or subordinate-level, and in a part judgment task that measures holistic processing.

Overview of the Study

To examine the interactions between visual appearance and conceptual knowledge, participants were randomly assigned to one of the two training groups. All stimuli and trial sequences during training and testing were matched between the two groups. The only difference between the two groups was the nature of the object-word pairings during the semantic training sessions: the object-word pairing was either in a typical manner (e.g., social semantics with symmetric, animal-like objects) or in a reversed manner (e.g., social semantics with asymmetric, tool-like objects).

The object set

Figure 2. Left panel: Example Greebles with a symmetric configuration and a smooth texture (S-Greebles). Right panel: Example Greebles with an asymmetric configuration and an artificial texture (A-Greebles).

One key manipulation in the current study was the visual appearance of the objects. The stimuli used in this study were modified "Greebles" (e.g., Gauthier & Tarr, 1997; Gauthier et al., 1998; Kung et al., 2007), novel objects each consisting of a large central part and four small peripheral parts (Figure 2, Appendix A). Some researchers have suggested that Greebles with peripheral parts organized symmetrically appear face-like (Kanwisher, 2000). Greebles with an asymmetric configuration avoid this confound while preserving the visual homogeneity of the set. Indeed, even after a long visual training, participants did not perceive asymmetric Greebles to be face-like (Kung et al., 2007). To investigate the influence of such differences in visual appearance, this study used both symmetric and asymmetric Greebles (S-Greebles, A-Greebles). The animate vs. inanimate character of the objects was emphasized further by the use of colors and textures appropriate for living creatures or man-made objects.

The word set

Explicit manipulation of conceptual associations was the second key manipulation in the study. A list of social and inanimate semantic features was generated in a pilot study, where participants were asked to generate words that describe non-visual attributes of people, animals, tools and man-made objects. Social and inanimate features were then selected and ambiguous words that could be used to describe either social or inanimate categories were avoided (e.g., “sharp”, “fast”). Appendix B shows the complete list of words used in the study.

Training overview

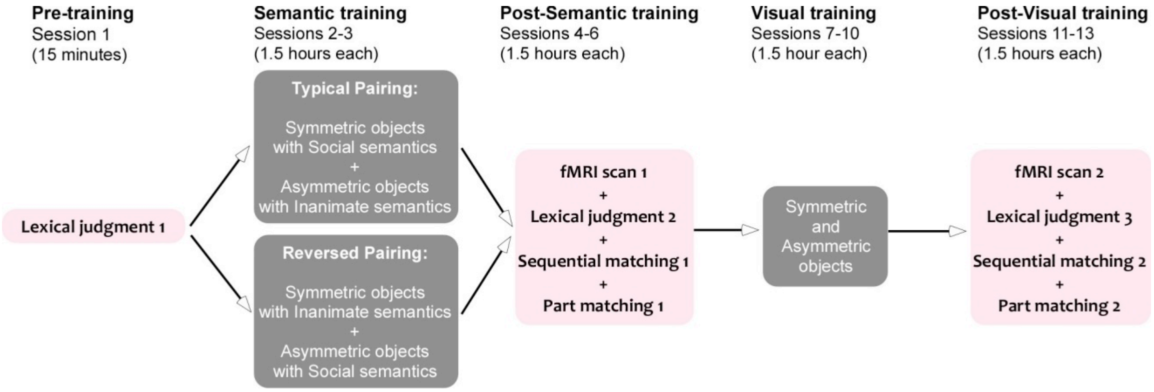


Figure 3. Sequence of training and testing sessions. Note that the only difference in procedure between the two groups occurred during the semantic training.

The current study included two stages of training. First came semantic training, conducted in two 1.5-hour sessions. The semantic training followed published procedures in which participants learned to associate different types of semantic features with objects (Gauthier et al., 2003a; James & Gauthier, 2003;

2004). Here, participants learned to associate two types of semantic features with individual objects from two categories. While the objects and semantic features used were identical across the two training groups, the object-word *pairing* differed between them. The Typical Pairing group learned to associate inanimate semantics with A-Greebles and social semantics with S-Greebles. The Reversed Pairing group learned to associate social semantics with A-Greebles and inanimate semantics with S-Greebles.

The second phase of the study consisted of four 1.5-hour sessions of visual training at the subordinate-level, with half of the training objects already seen and familiarized during the semantic training. The training procedure was based on previous perceptual expertise training studies involving speeded subordinate-level categorization (Gauthier & Tarr, 1997; 2002; Kung et al., 2007; Rossion et al., 2004; Scott et al., 2006; 2008; Tanaka et al., 2005; Wong et al., 2009a; Wong et al., 2009b). In most of these studies, participants were trained with only one homogenous set of novel or real-world objects (e.g., symmetric Greebles, Gauthier & Tarr, 1997; owls, Tanaka et al., 2005). The design of the current study differed in two important ways. First, two categories of training objects (S-Greebles and A-Greebles) were used for each participant. Second, the visual training occurred after semantic associations had been explicitly learned for half of the objects used in the visual training. This allowed for the investigation of interactions between visual object properties (and the inferences they lead to) and explicit conceptual associations, both immediately after the semantic training and following the visual individuation training.

Pre- and post-training tests

To examine the effects of training, one behavioral task was administered prior to the training and three behavioral tasks were conducted after semantic training and after visual expertise training. Specifically, a lexical judgment task (15-minutes) tested whether a task-irrelevant object image presented with a word would affect the speed in judging the nature of a word, depending on pairing of the object and word. Second, a sequential matching task (45-minutes) evaluated matching performance at the basic-level vs. subordinate-level. Third, a part-matching task (60 minutes) commonly used to measure holistic processing related to processing of faces and objects of expertise was included (e.g., Cheung et al., 2008; Richler et al., 2008a; Gauthier et al., 2003b; Young et al., 1987). To limit exposure to the objects prior to training, only the 15-minute lexical judgment task but not the other two tasks was conducted prior to training. An additional control group with no training also participated in the sequential matching task to provide a no training baseline.

To explore the neural changes accompanying the behavioral training effects, a functional magnetic resonance imaging (fMRI) scan was also conducted after the semantic training, and another fMRI scan was conducted after the visual training. Further details are included in Chapter 3.

Methods

Participants

Participants were 36 young adults from the Vanderbilt University community. Twelve participants (6 females, age M=22.58, SD=4.32) were randomly assigned to the Typical Pairing group, 12 participants (4 females, age M=23.67, SD=4.29) to the Reversed Pairing group, and the last 12 participants (5 females, age M=22.67, SD=3.08) to a control group for the sequential matching task. All participants reported normal or corrected-to-normal vision and none had been exposed to Greebles before. The participants in the two training groups were each compensated with $250 for their participation in the 16-hour study. Those in the control group were each compensated with $12/hour for the behavioral test.

Apparatus and Stimuli

All behavioral training and testing were conducted on Mac mini computers using Matlab (MathWorks, Natick MA) with the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). Stimuli were presented on 19″ CRT monitors with a 1280 × 960 resolution.

Objects

Two versions of 48 novel objects were created using 3D Studio Max 9 (Autodesk, Inc., http://usa.autodesk.com) with the “Greeble Generator” script (Tarr lab, 2002). Each participant was only presented with one version of the objects: half of the participants were shown the first version, and the rest were shown the second version (see Appendix A). For each version, all S-Greebles had an identical *symmetric* configuration of top peripheral parts and were rendered with the same organic textures. All A-Greebles shared the same *asymmetric* configuration of peripheral parts and were rendered in identical metallic textures. Note that configuration symmetry refers to object, and not image, symmetry. There were 24 S-Greebles (18 had the same body shape and color and 6 had a different body shape and color) and 24 A-Greebles (18 had the same body shape and color and 6 had a different body shape and color). The two versions had the same central and peripheral parts but differed in the part assignment to the trained and transfer objects.

The Greebles within each set differed in the shape of the peripheral parts: S-Greebles generally had smooth-edged peripheral parts pointing downward and A-Greebles generally had sharp-edged peripheral parts pointing upward. The four sub-categories of Greebles in each version were relatively easy to discriminate at the category level, as they had different central parts and were rendered in different colors. All Greebles were rendered with realistic lighting and shading on a white background at four viewpoints (0˚, 6˚, 12˚, and 18˚. The 0˚ viewpoint was an arbitrarily defined canonical orientation). Each object image was about 9.5 cm × 5.6 cm in size (a visual angle of about 6˚ × 3.6˚ from a viewing distance of 90 cm). To avoid image-based effects, objects used during training were shown at either 0˚ or 18˚. During the behavioral and fMRI tests, the objects were presented at either 6˚ or 12˚.

The selection of objects for use in the training and testing was counterbalanced across participants. For each participant, 6 S-Greebles and 6 A-Greebles from the sets of 18 (the trained sets) were first used during semantic training. Because visual individuation training typically involves a large number of individual exemplars (e.g., Gauthier et al., 1998; Wong et al., 2009a), 6 additional S-Greebles and 6 additional A-Greebles were also included as part of the trained sets during the visual training. Implicit effects of the kind of semantic training implemented here have so far only been tested with the trained exemplars (James & Gauthier, 2003; 2004) whereas visual expertise training effects have been observed for unseen exemplars within the expert category (e.g., new car models for a car expert; Gauthier et al., 2003b; Curby & Gauthier, 2009; Gauthier & Tarr, 1997; 2002; Wong et al., 2009a; Wong et al., 2009b). Thus, the untrained objects in the trained sets (transfer-1 objects) were included to investigate the generalization of training effects to objects that shared all properties of the trained categories expect for the shape of their individual parts. The transfer-1 objects were expected to produce, as in prior work, considerable perceptual expertise effects. In fact, they are often the only kind of objects used to test expertise effects (e.g., Gauthier & Tarr, 1997; 1999; Wong et al., 2009a) so it will be interesting to compare expertise effects on the transfer-1 objects to those for trained exemplars. Since all objects were used during testing, there were 12 transfer-1 S-Greebles and 12 transfer-1 A-Greebles after the semantic training, while the number of transfer-1 objects for each Greeble type was reduced to 6 after the visual training once some of these objects were experienced in the visual training.

The rest of the 6 S-Greebles and 6 A-Greebles, which were different in color and central shape but shared a configuration of these parts with the trained objects, were also used for testing generalization to more visually distinct objects (transfer-2 objects). Although the differences between the transfer-2 and trained objects were salient in terms of color and body shape, these characteristics were not diagnostic during either the semantic or visual training tasks. Thus, it will be interesting to test whether differences in color and body shape limit generalization. In other words, this is one way to ask the question of what defines the boundaries of the trained categories. Visual expertise training effects do not usually transfer to objects that are outside the shape space of the expert category (Bukach et al., in press; cf. Grill-Spector et al., 2004; Yue et al., 2006), but control objects tend to differ both on superficial characteristics and in dimensions relevant during training.

Semantic features

A hundred and twenty words were used, including 60 social features and 60 inanimate features (see Appendix B). These words describe *non-visual* attributes that could be used to describe people or man-made objects, generated by the experimenter and 20 participants in a pilot study. Word length was controlled across the social (M=7.35 letters, SD=2.15) vs. inanimate (M=7.67 letters, SD=1.92) sets. I was not able to match word frequency across the two sets because of the large number of words used in the study and the selection constraint regarding category ambiguity (e.g., “sharp” may be used to describe both people and objects). According to the $\text{SUBTLEX}_{\text{us}}$ word frequency database (Brysbaert & New, 2009), the mean frequency was higher for the social (M=33.5, SD=58.36) vs. inanimate (M=3.33, SD=5.56) sets. Nonetheless, since the critical manipulation in this study was the object-word *pairing* and identical words were used for both training groups, word frequency alone would not be able to account for differences between the training groups.

Non-overlapping subsets of these words were used in the lexical judgment task (24 of each kind), during semantic training (18 of each kind) and the rest of the 18 words for each set were used during the fMRI semantic localizer.

Semantic training procedures

The semantic training (Figure 4) was modeled after prior work (Gauthier et al., 2003a; James & Gauthier, 2003; 2004). The current study involved learning 18 social or 18 inanimate verbal associations for 6 A-Greebles and 6 S-Greebles (3 words for each object) in two 1.5-hour sessions. All participants in the two training groups were trained with the same words. The selection of trained objects were counter-balanced across participants within each group and matched between groups. Thus, the only difference between the two groups was the *pairing* between words and objects. The Typical Pairing group learned 3 social features for each of the 6 S-Greebles and 3 inanimate features for each of the 6 A-Greebles. The Reversed Pairing group instead learned 3 inanimate features for each of the 6 S-Greebles and 3 social features for each of the 6 A-Greebles.

In the first session, 4 S-Greebles and 4 A-Greebles were learned. In the second session, 2 Greebles from each category were added. The S-Greebles and A-

Greebles were learned in separate blocks. As in several previous studies (James & Gauthier, 2003; 2004; Bukach et al., unpublished), the current training included 4 tasks (Figure 4): 1) passive viewing, 2) three-feature matching, 3) single-feature verification, and 4) fill-in-the-blanks in order to prevent task-specific learning effects and promote robust associations. Details of the four tasks are discussed below.

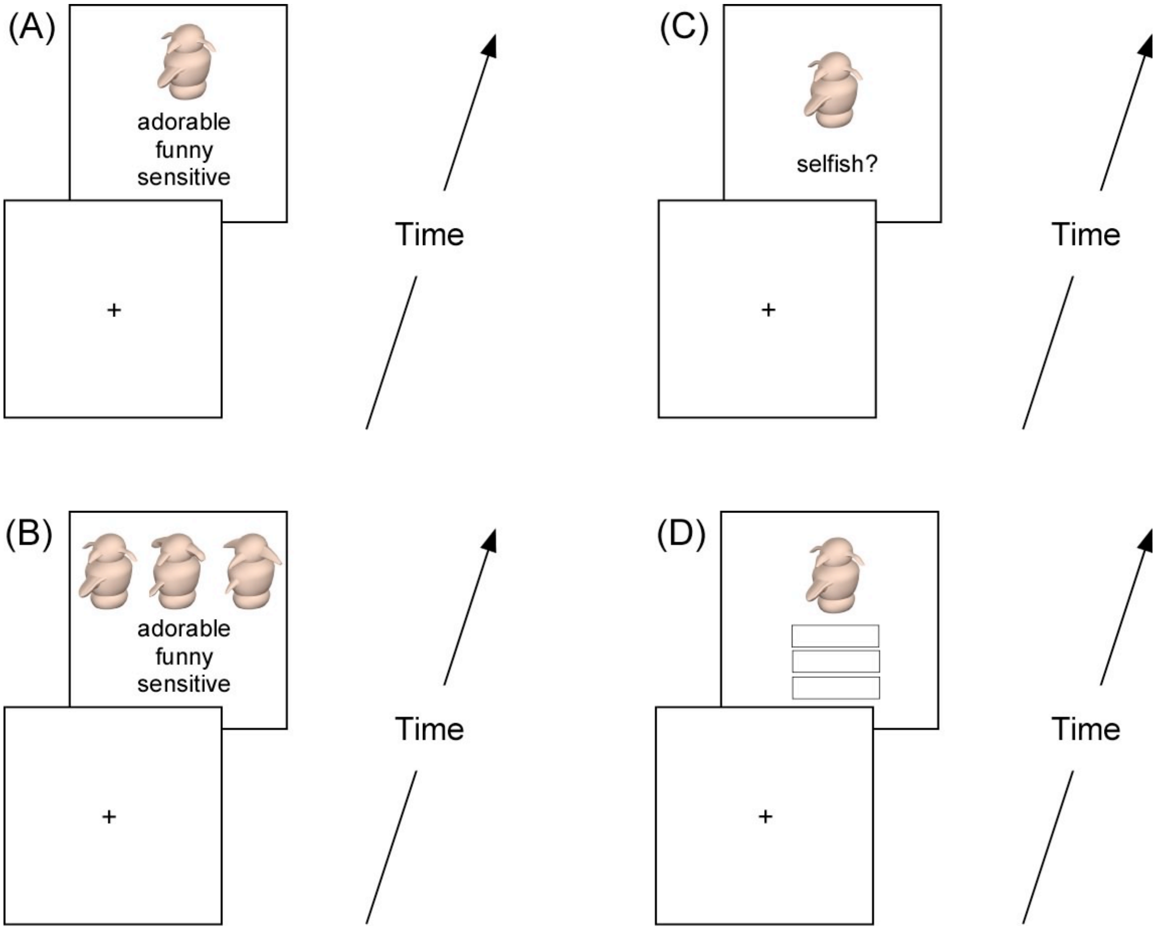


Figure 4. Semantic training tasks. *Figure 4A:* In a passive viewing task, participants viewed a Greeble with its 3 semantic features with no time limit. *Figure 4B:* In a 3-feature matching task, participants decided which one of the three Greebles matched the 3 semantic features. *Figure 4C:* In a single feature verification task, participants judged if a semantic feature matched a Greeble. *Figure 4D:* In a fill-in-the-blanks task, participants had to input the 3 semantic features for each Greeble.

Passive viewing

Because the objects and their associations were novel, the passive viewing phase allowed participants to study each Greeble and its associations for as long as they liked (Figure 4A). On each trial, a Greeble image and its 3 associated words were presented. Participants terminated each study trial when they had finished studying an object and the words. Each pair was presented twice. There were a total of 16 and 24 passive viewing trials in Sessions 1 and 2 respectively.

Three-feature matching

During the 3-feature matching task (Figure 4B), participants practiced identifying the appropriate object associated with each set of 3 semantic features. On each trial, 3 Greebles from the same category were presented along with 3 semantic features that were associated with the same target. Participants selected the target Greeble that matched the semantic associations. All choices remained on the screen until a response was made. After each incorrect trial, the computer beeped and the Greeble and its three associated features were shown on the screen, allowing participants to further study them until they pressed the space bar. Because participants were still learning the object-word associations, accuracy was emphasized at this stage, but if the average response times (RTs) for each 36-trial block exceeded 4 seconds, a warning appeared at the end of the block to encourage participants to speed up their responses. There were a total of 576 three-feature matching trials in each session.

Single-feature verification

To prevent participants from learning only 1 but not all 3 semantic features for each Greeble, each semantic feature was then presented singly in the single feature verification task (Figure 4C). On each trial, a word was presented below a Greeble image until a response was made. Participants pressed a key to indicate whether the word was associated with the Greeble. In *different* trials, the image was of one of the other Greebles from the same set. Incorrect responses were followed by a beep, and then the correct Greeble and its 3 semantic features were presented for unlimited study. Accuracy was again emphasized over speed in this phase but participants were warned when the average response time for each 36-trial block was over 4 seconds. There were a total of 576 single-feature verification trials in each session.

Fill-in-the-blanks

Participants' ability to recall the associated semantic features for each Greeble was also tested at the end of each session (Figure 4D). On each trial, a Greeble was shown and participants were asked to type its 3 associated semantic features. Each trained Greeble was shown twice during this phase. No feedback was given. There were a total of 16 and 24 fill-in-the-blank trials in Sessions 1 and 2 respectively.

Visual training procedures

The visual training (Figure 5) involved individuation training with 12 S-Greebles and 12 A-Greebles, including the 6 Greebles in each category that were already used during semantic training. Additional objects were used here because the visual training required speeded naming to a large number of exemplars within each category. Each of the 24 trained Greebles was assigned a 4-letter nonsense words as its name (e.g., Piko, Tawu, Xedo, Kica). Name assignment was randomized for the 12 participants in each group, but was matched between the two groups. Note that the association of proper names to individual objects has a minimal effect on the conceptual network in the cortex, compared to the association with more meaningful information (James & Gauthier, 2004). Based on previous visual training studies, approximately 7-10 hours of training was sufficient to produce face-like expertise performance both in behavior and in neural activity (Gauthier & Tarr, 2002; Gauthier et al., 1998; Rossion et al., 2004; Rossion et al., 2007; Wong et al., 2009a; Wong et al., 2009b). Since half of the training objects used in the individuation training had already been used during semantic training, it was assumed that a shorter training time would be sufficient. The visual training was conducted in four 1.5-hour sessions.

Each training session consisted of 3 tasks to encourage individuation but avoid task-specific effects (e.g., Gauthier et al., 1998; Wong et al., 2009a). Session 1 included 12 objects (3 S-Greebles and 3 A-Greebles learned during semantic training, and 3 unlearned objects from each category). In Session 2, the rest of 12 training objects were introduced. All objects were then used in Sessions 3 and 4.

Both speed and accuracy were emphasized in all training tasks: The mean speed and accuracy for the block were shown at the end of each block to motivate participants. All objects were presented on the screen until a response was made in Sessions 1-3. To encourage fast responses and keep the task difficult, each object was shown only briefly in Session 4. The three tasks included in each training session are described below. Half of the blocks in each task were devoted to A-Greebles, and the other half to S-Greebles (order counter-balanced across participants).

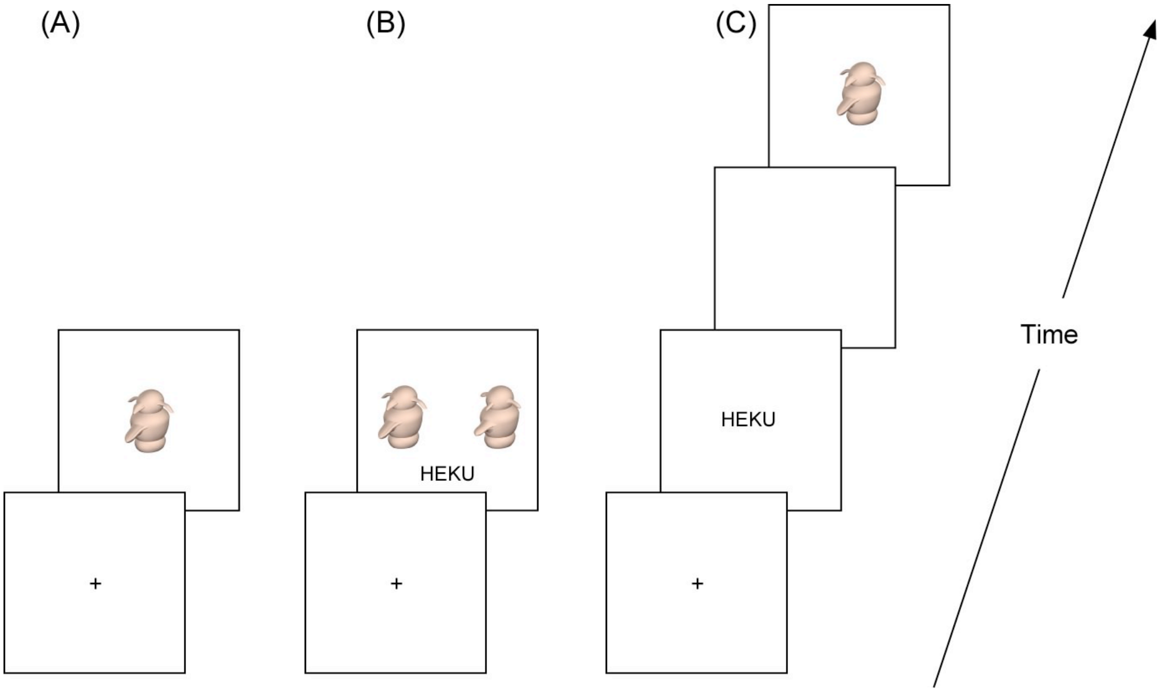


Figure 5. Visual training tasks. *Figure 5A:* In a naming task, participants pressed the key for the first letter of a Greeble's name. *Figure 5B:* In a name-matching task, participants judged which one of the two Greebles matched with a name. *Figure 5C:* In a name-verification task, participants judged whether a Greeble and a name matched.

Naming

Participants first learned names associated with each of the trained Greebles (Figure 5A). On each trial, a central fixation appeared (250ms), and was replaced by an object image. A name was presented only during the first three presentations of each object. Participants were asked to type the first letter of the names. After each incorrect response, a beep was presented and the target Greeble and its name would be shown for unlimited study.

Name-matching

Participants were then asked to choose, quickly and accurately, which of two Greebles matched a name (Figure 5B). On each trial, a name was presented below two Greeble images that were presented side by side. Participants pressed a key to choose the left or right image. The non-target was another trained Greeble from the same category. Incorrect responses were followed by a beep, and then the target Greeble and its name were presented for unlimited study.

Name-verification

Participants then judged whether a name and a Greeble matched (Figure 5C). On each trial, a name was presented (1000ms), followed by a blank screen (250ms), and then by a Greeble image. Participants pressed a key to indicate whether the name matched the image. For non-match trials, the object was another training object in the same category. Incorrect responses were followed by a beep and then the target Greeble and its name were presented for unlimited study.

Behavioral pre- and post-training tests

The lexical judgment task examined the interference from a task-irrelevant object on lexical decision. The sequential matching task measured the categorization performance at the basic- vs. subordinate-level. The part-matching task tested the emergence of holistic and configural effects. The details for the three behavioral tasks are given below.

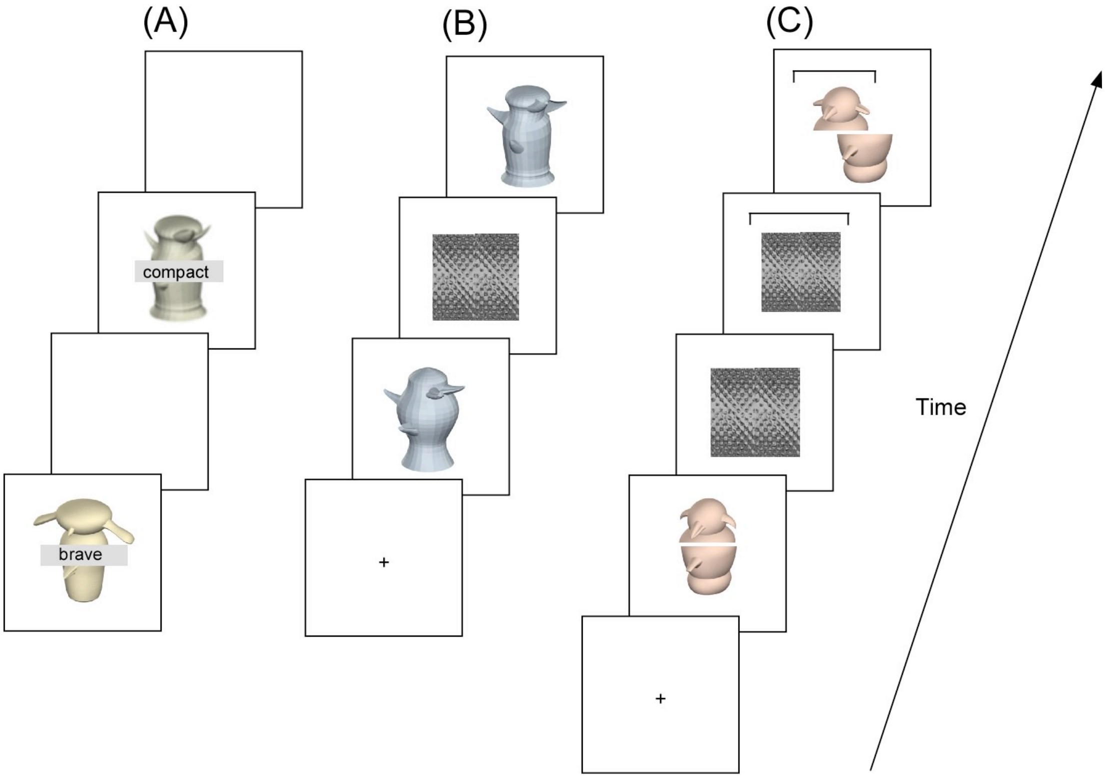


Figure 6. Behavioral tasks. *Figure 6A:* In a lexical judgment task, participants judged if a target word was better used to describe people or man-made objects. *Figure 6B:* In a sequential matching task, participants judged if two sequentially presented objects were from the same category (with category defined by the central body part; basic-level trials) or if they were identical (despite differences in viewpoint; subordinate-level trials). A sample basic-level trial is shown here. *Figure 6C:* In a part-matching task, participants judged if a cued part of a test object matched that of a study object.

Lexical judgment task

Previous studies showed that an irrelevant face presented concurrently with a name could influence RTs when participants judged whether a name belonged to a politician or a celebrity (de Fockert et al., 2001; Jenkins et al., 2003). Here, a similar task was used to evaluate the effect of an irrelevant object in a lexical judgment task (Figure 6A). On each trial, one S-Greeble, A-Greeble (trained, transfer-1 or transfer-2), or phase-scrambled image was presented behind a word. Twenty-four social words and 24 inanimate words were used. Participants judged if the word was better used to describe people or man-made objects. In this task, each Greeble was shown four times, twice for each of the word types (social or inanimate). Each word was presented three times (once with an S-Greeble, once with an A-Greeble and once with a scrambled image). There was a 1-second interval in between trials. The combination of Greeble trials included 2 (Visual appearance: S-Greebles vs. A-Greebles) × 3 (Training status: trained vs. transfer-1 vs. transfer-2) × 2 (Word type: social vs. inanimate) × 2 (Viewpoint: 6° vs. 12°) × 12 trials. In addition, there were 48 trials where scrambled images were presented, for a total of 432 trials, with all trial types randomized. No feedback was given. This task lasted about 15 minutes.

Sequential matching task

The sequential matching task (Figure 6B) was used to test for a reduction of basic-level advantage expected after training (Tanaka, 2001; Wong et al., 2009a). The basic-level advantage reveals faster recognition performance at the basic-level (e.g., “dog”) than at the subordinate-level (e.g., “golden retriever”), with basic-level categories defined as those across which the maximal diagnostic information in shape is present (Jolicoeur et al., 1984; Rosch et al., 1976). A hallmark of face-like perceptual expertise is the reduction of the basic-level advantage, since expert observers recognize objects of expertise at the subordinate-level almost as quickly as at the basic-level (Tanaka, 2001; Tanaka & Taylor, 1991; Mack et al., 2009).

There were two types of trials in the sequential matching task. Participants studied an image and judged either whether the following image showed an object from the same category (as defined by the different central body parts; basic-level trials) or showed the same object (subordinate-level trials), regardless of slight viewpoint difference (6˚). In the basic-level trials, a non-match for S-Greebles was randomly selected from the other category of S-Greebles, while a non-match for A-Greebles was randomly chosen from the other category of A-Greebles. To force participants to use shape information rather than the color differences across categories in the basic-level trials, the non-match objects in those trials were rendered to be the same color as the first (target) image of each trial. In the subordinate-level trials, each non-match was selected from within each training status subset of the same type of visual appearance (e.g., trained S-Greebles).

On each trial (Figure 6B), a fixation was shown (300ms), followed by a study object (800ms), a mask (500ms), and then by a test image (1s). The design included 2 (Visual appearance: S- vs. A-Greebles) × 4 (Training status: trained-semantically and visually vs. trained-visually vs. transfer-1 vs. transfer-2) × 2 (Categorization level: basic vs. subordinate) × 2 (Viewpoint: 6˚ vs. 12˚) × 2 (Response: same vs. different) × 12 trials. Visual appearance, training status and categorization level were blocked, with 24 trials in each block. Feedback was given only during practice trials. There were a total of 768 experimental trials. This task lasted approximately 45 minutes.

Part-matching task

The part-matching task (Figure 6C) was used in previous studies to measure holistic and configural processing. This task required selective attention to parts of an object. Participants were asked to attend to and match only one half of the study and test objects (e.g., top) and ignore the task-irrelevant half (e.g., bottom). Because the identities of the target and task-irrelevant halves varied independently from each other, the relations between the two halves were either congruent (both halves were the same or both were different between study and test) or incongruent (target halves are the same and irrelevant halves are different or vice versa). A congruency effect, revealed by better performance for congruent than incongruent trials, indicates failures of selective attention to parts. The congruency effect has been obtained for faces and after subordinate-level categorization training but typically not for objects with novices (Cheung et al., 2008; Gauthier & Tarr, 2002; Wong et al., 2009a). Moreover, the congruency effect is disrupted by spatial misalignment of halves (indicating a configural effect), since misalignment of halves breaks the familiar configuration and reduces holistic processing (e.g., Young et al., 1987; Wong et al., 2009a). Both the holistic and configural effects were tested here.

On each trial (Figure 6C), participants saw a fixation cross (300ms) and then the first composite (parts always aligned) for 800ms, followed by a pattern mask

(600s) with a response cue that indicated whether the top or bottom face part was the target appearing for 200ms before the removal of the mask, and finally followed by the second composite with the response cue (1s). The parts of the second composite were either aligned or misaligned. Participants indicated by key press whether the target parts of the two composites were the same. Feedback was only given during practice trials. The experimental design included 2 (Visual appearance) × 4 (Training status) × 2 (Alignment: aligned vs. misaligned) × 2 (Congruency: congruent vs. incongruent) × 2 (Response: same vs. different) × 16 trials. There were a total of 1024 experimental trials. This task lasted about 1 hour.

Predictions

Performance on all training tasks was expected to improve in terms of RTs and accuracy during the course of both semantic and visual training. The training paradigms were designed to ensure that both groups would learn, but it is interesting to ask whether the reversed semantic associations would be more difficult to learn than the typical semantic associations. The two types of training had different emphases and thus were predicted to have different impacts on the behavioral pre- and post-training tests.

It was expected that performance in the lexical judgment task would be facilitated by a task-irrelevant image from the same category as the target word (e.g., a S-Greeble and a word that describe people) and/or impeded by an image from the other category (e.g., an A-Greeble and a word that describe people). This effect could occur in the pre-test, but following semantic training, different patterns of results were expected in the Typical vs. Reversed Pairing groups. The semantic training effects could be restricted to the trained objects (James & Gauthier, 2003). As no explicit semantic associations were mentioned during the visual training, there was no strong prediction about how the semantic training effects would change after the visual training. The effects of the explicitly learned semantic associations might be reduced after visual training because of the lack of practice. On the other hand, it is possible that biasing participants' conceptual inferences about objects early on plants the seed for effects that grow larger with more experience. Indeed, even though there is no experimental mention of semantics during visual individuation training, the participants were free to continue elaborating.

The sequential matching task was expected to reveal changes in sensitivity to objects for basic- vs. subordinate-level categorization due to the semantic and visual training. Overall performance was expected to improve in the two training groups, compared to a control group with no training. More importantly, a basic-level advantage, as revealed by better performance for basic-level than subordinate-level trials, was expected for the control group. The subordinate-level performance was expected improve following semantic training because the semantic training required individuation. The subordinate-level performance should be further improved following the visual individuation training (e.g., Gauthier & Tarr, 2002; Tanaka et al., 2005; Scott et al., 2006; 2008; Wong et al., 2009a). So far, the reduction of basic-level advantage has been demonstrated for faces (Tanaka, 2001), animals (Tanaka & Taylor, 1991; Tanaka, 2001; Tanaka et al., 2005; Mack et al., 2009), symmetric Greebles (Gauthier & Tarr, 1997; Gauthier et al., 1998) and tool-like novel objects (Wong et al., 2009a). Thus, it was predicted that the shift would occur for both S-Greebles and A-Greebles.

For the part-matching task, a congruency effect was expected for trained S- and A-Greebles after visual training and the congruency effect should be reduced by misalignment, since holistic effects have been demonstrated with various animate and inanimate object categories when tested with experts (e.g., faces, Young et al., 1987; Cheung et al., 2008; cars, Gauthier et al., 2003b; tool-like Ziggerins, Wong et al., 2009a). It is unclear whether holistic processing might start to emerge after very few training sessions (Gauthier & Tarr, 2002). Another outstanding question was whether the introduction of semantic training would result in different effects, although conceptual knowledge was not predicted to have a strong impact on these effects, because holistic and configural processing arise from perceptual, rather than post-perceptual or response, processes (Richler et al., 2009).

Based on previous findings (James & Gauthier, 2003; 2004), the semantic training effects were expected primarily in the trained objects whereas the visual training effects were expected in both trained and transfer-1 objects (e.g., Gauthier & Tarr, 2002). However, since the transfer-2 objects differed from the trained and transfer-1 objects only in dimensions that were not critical during either stage of training (i.e., central shape and color), to what extent generalization might occur in the transfer-2 objects remained an open question.

Results

Semantic training

Participants improved in both accuracy and RTs during training (Figure 7A & 7B). Learning associations with S-Greebles was easier than learning associations with A-Greebles in both training groups (Figure 7A, 7B, & 7C).

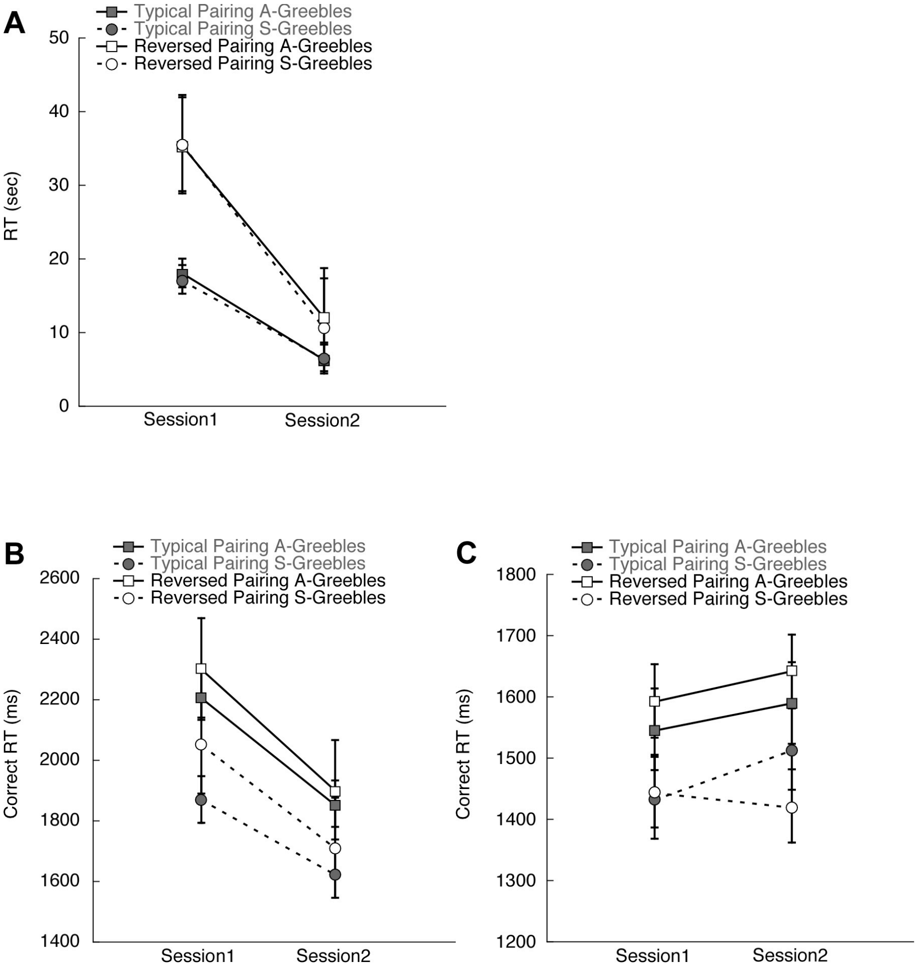


Figure 7. Semantic training results. *Figure 7A*: mean RTs in the passive viewing task. *Figure 7B*: mean correct RTs in the three-feature matching task. *Figure 7C*: mean correct RTs in the single-feature verification task.

In the passive viewing task (Figure 7A), RTs in Session 2 were faster than Session 1, *F*(1,22)=24.95, *p*<.0001. There was no significant difference of Visual appearance, *F*(1,22)=.09, *p*=.77. The main effect of Pairing and the interaction between Session and Pairing approached significance in RTs, *F*(1,22)=4.13, *p*=.054; *F*(1,22)=3.37, *p*=.08.

In the three-feature matching task (Figure 7B), performance was faster and better in Session 2 than Session 1, RTs: *F*(1,22)=74.93, *p*<.0001; Accuracy: *F*(1,22)=11.38, *p*=.003. Participants matched S-Greebles faster than A-Greebles, *F*(1,22)=34.09, *p*<.0001. The interaction between Session and Pairing was marginally significant in accuracy, *F*(1,22)=4.21, *p*=.052. There were no other significant effects or interactions, RTs: *F*s(1,22)≤1.06, *p*s>.31; Accuracy: *F*s(1,22)≤.71, *p*s>.4.

In the single-feature verification task (Figure 7C), participants were faster and better at verifying semantic features for S-Greebles than A-Greebles, RTs: *F*(1,22)=15.16, *p*<.001; Accuracy: *F*(1,22)=5, *p*<.04. The interaction between Session and Pairing approached significance in accuracy, *F*(1,22)=3.78, *p*=.065. There were no other significant effects or interactions, RTs: *F*s(1,22)≤2.14, *p*>.15; Accuracy: *F*s(1,22)≤1.96, *p*>.17.

In the fill-in-the-blanks task (Table 1), participants recalled most of the associations at the end of each training session. There was a marginally significant effect of Pairing, *F*(1,22)=3.3, *p*=.083. There were no other significant results, *F*s(1,22)≤1.17, *p*>.29. At the very end of the entire study, the number of correct recall dropped significantly compared to that at the end of the semantic training,

$F(1,22)=239.6$, $p<.0001$. There were again no significant effects within this last test, $Fs(1,22)\leq.58$, $p>.45$.

Session	Typical Pairing S-Greebles	Typical Pairing A-Greebles	Reversed Pairing S-Greebles	Reversed Pairing A-Greebles
1	98.6% (.009)	96.5% (.016)	95.1% (.024)	94.4% (.036)
2	99.5% (.004)	98.1% (.008)	95.4% (.019)	97.7% (.014)
End of study	36.6% (.057)	39.8% (.06)	38.0% (.065)	36.1% (.069)

Table 1. Mean accuracy in the fill-in-the-blank task during the two semantic training sessions and at the end of the study. Standard error of the mean was shown in the brackets.

Summary of the semantic training results

In sum, both training groups were able to learn the associations between the objects and words. Interestingly, visual appearance modulated learning, with S-Greebles matched and verified better than A-Greebles. This is likely due to the facilitation due to symmetry on encoding (Reber et al., 2004). Also, at the end of the study, both groups were still able to recall about 1/3 of the learned features to the appropriate objects, indicating that some explicit associations remained even though these associations were not practiced during the visual training.

Visual training

The data for the 4 training days were divided into 8 temporal bins. Data from one participant from each group was excluded from analyses due to an error in recording response times. In general, the results across the three training tasks were highly similar and participants improved in all three tasks during training (Figure 8A, 8B, & 8C). Also, performance was better for S- than A-Greebles.

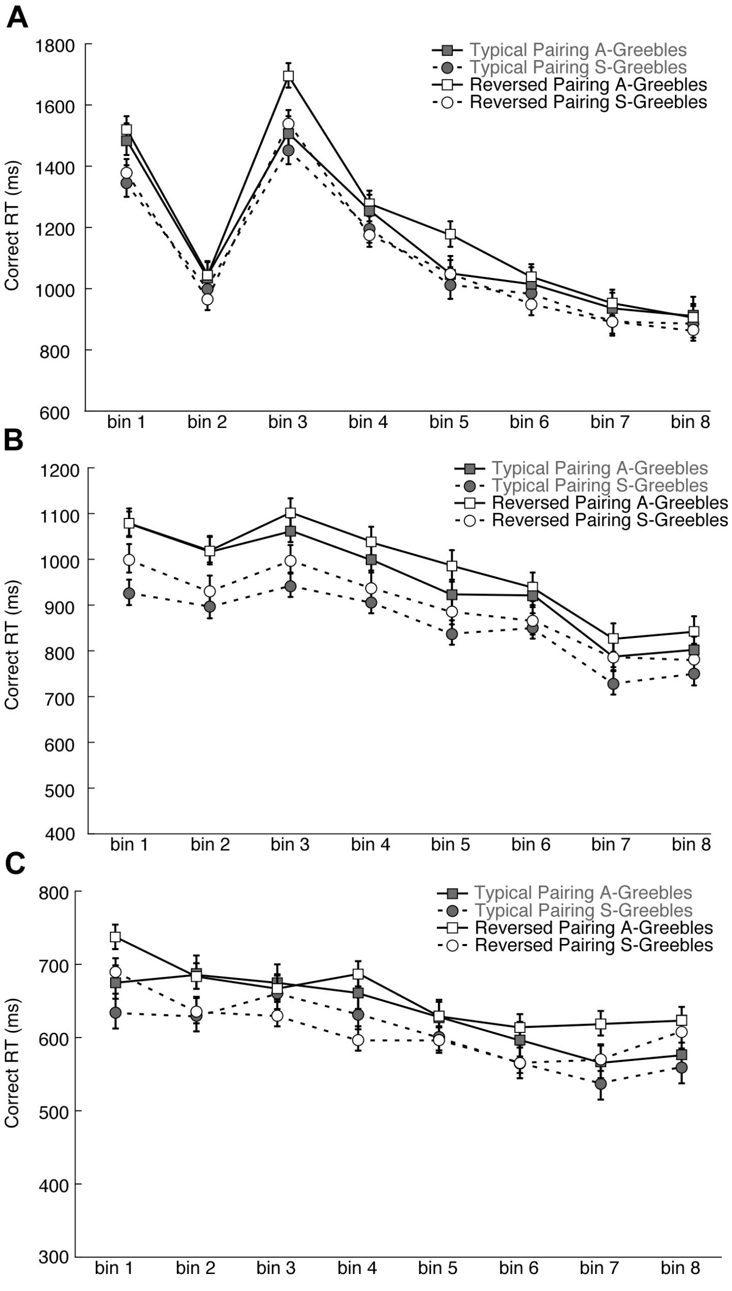


Figure 8. Visual training results. *Figure 8A*: mean correct RTs in the naming task. *Figure 8B:* mean correct RTs in the name-matching task. *Figure 8C:* mean correct RTs in the name-verification task.

In the naming task (Figure 8A), RTs were slowest and accuracy was the lowest when names were first introduced (12 names were shown during Bin 1 and an additional 12 names were shown during Bin 3), RTs: *F*(7,140)=113.5, *p*<.0001; Accuracy: *F*(7,140)=5.12, *p*<.001. Participants named S-Greebles faster than naming A-Greebles, *F*(1,20)=15.3, *p*<.001. The interaction between Bin and Visual appearance was also significant in RTs, *F*(7,140)=2.12, *p*<.05, with a larger improvement across bins for A- than S-Greebles. There were no other significant effects or interactions in RTs or accuracy, *F*s<1.5, *p*s>.2.

In the name-matching task (Figure 8B), a significant effect of Bin also revealed slower responses in Bins 1 and 3 and faster responses as more training was received, RTs: *F*(7,140)=55.87, *p*<.0001; Accuracy: *F*(7,140)=19.05, *p*<.0001. Participants matched Greebles with their names faster for S-Greebles than A-Greebles, *F*(1,20)=39.25, *p*<.0001. The interaction between Bin and Visual appearance was also significant in RTs, *F*(7,140)=2.85, *p*<.01. There were no other significant effects or interactions in RTs or accuracy, *F*s<1.02, *p*s>41.

In the name-verification task (Figure 8C), there was also a significant effect of Bin, RTs: *F*(7,140)=11.53, *p*<.0001; Accuracy: *F*(7,140)=3.09, *p*<.0001. Names for S-Greebles were faster and better verified than those for A-Greebles, RTs: *F*(1,20)=26.03, *p*<.0001; Accuracy: *F*(1,20)=15.44, *p*<.001. The interaction between Bin and Visual appearance approached significance in RTs, *F*(7,140)=1.87, *p*=.078. There were no other significant effects or interactions in RTs or accuracy, *F*s<1.44, *p*s>.19.

Summary of the visual training results

To summarize, performance improved during the course of the visual individuation training. Also, S-Greebles were individuated faster and better than A-Greebles. Such difference in performance between the two Greeble categories was largest at the beginning of the training and was reduced as training went on. It is important to note that performance of the two training groups was comparable during this training stage.

Lexical judgment task

The two training groups were tested three times in the lexical judgment task (see Figure 3 & 6A). Performance was analyzed separately at each of the training stages: pre-training, post-semantic training and post-visual training. The word "curvy" was discarded because of its ambiguity as a word more commonly used to describe people or man-made objects. RTs for correct trials were the focus of this study because overall accuracy in this task was high and the only significant effect was better performance in categorizing social than inanimate features for the trained objects after semantic training, $F(1,22)=5.56$, $p=.028$. RT outliers of 4 standard deviations from the mean were excluded from analyses.

Pre-test

During pre-test (Figure 9A), all objects were novel and data for both training groups were collapsed for analysis. A 2 (Visual appearance) × 2 (Word type) ANOVA was conducted. For RTs, performance was faster for social than inanimate features, *F*(1,22)=9.15, *p*<.01. Critically, the significant interaction between Word type and Visual appearance revealed that S-Greebles facilitated judgment for social features compared to inanimate features, whereas A-Greebles showed a smaller effect, *F*(1,22)=9.48, *p*=.0055. These results indicated an implicit interaction between visual and semantic properties of objects. No other effects or interactions were significant, *F*s(1,22)<1.45, *p*s>.24.

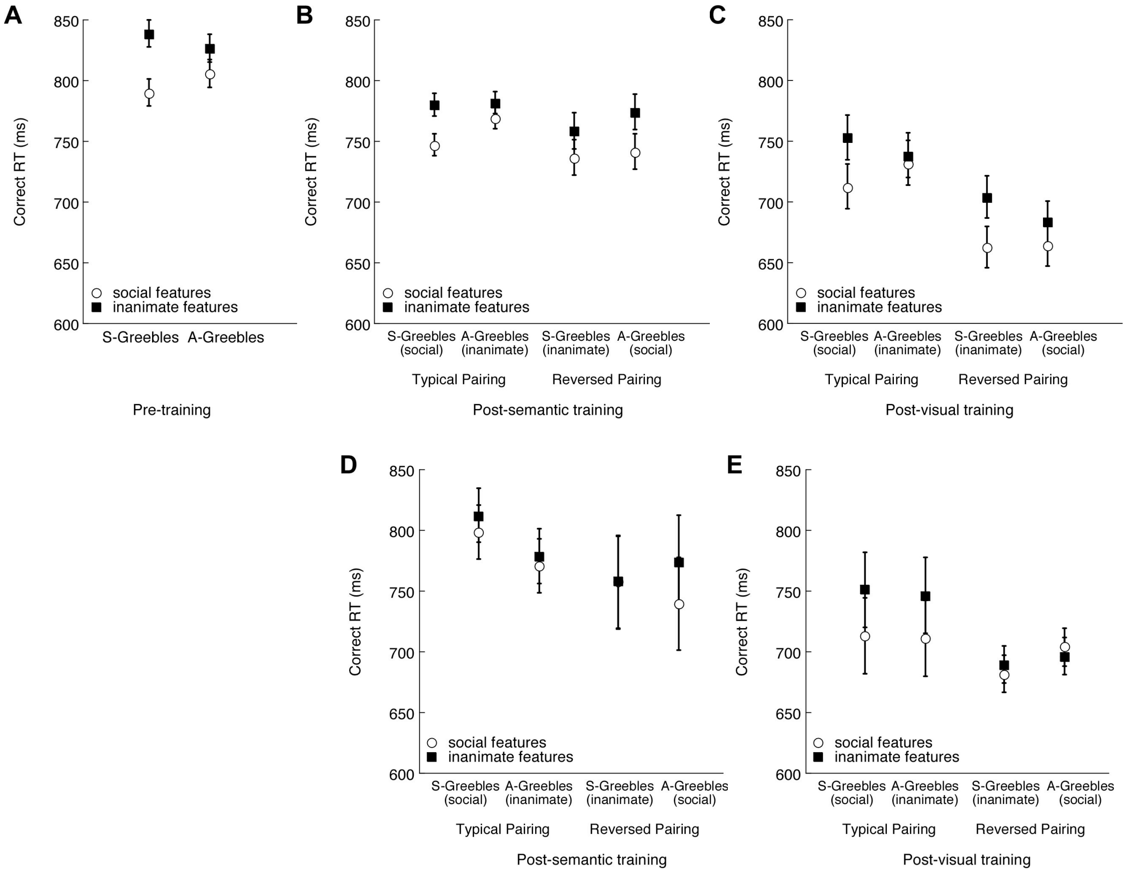


Figure 9. Mean correct RTs in the lexical judgment task. *Figure 9A*: all objects (collapsing across training status) in the pre-test. *Figure 9B:* trained and transfer-1 objects in the post-semantic training test. *Figure 9C*: trained and transfer-1 objects in the post-visual training. *Figure 9D:* transfer-2 objects in post-semantic training. *Figure 9E:* transfer-2 objects in post-visual training. Error bars represent the 95% confidence interval of the Visual appearance × Word type interaction for each analysis within each group in each testing session.

Post-training tests

For the post-semantic and post-visual training tests, a 2 (Pairing) × 2 (Visual appearance) × 2 (Word type) ANOVA was first conducted separately for the trained, transfer-1 and transfer-2 objects. But since there were no significant differences between the trained and transfer-1 objects in either post-semantic or post-visual training test, *F*s(1,22)≤2.12, *p*s≥.16, the data for these two object subtypes were collapsed for analyses.

Post-semantic training test

Trained and transfer-1 objects

The semantic training led to different patterns of results in the two groups (Figure 9B). First of all, the main effects of Word type and Visual appearance, and the three-way interaction between Pairing, Word type and Visual appearance approached significance, *F*(1,22)=3.19, *p*=.09, *F*(1,22)=3.33, *p*=.08, *F*(1,22)=3.88, *p*=.06. Because an important goal in this task was to examine the effects of different pairings of visual appearance and conceptual knowledge after the semantic training, a 2 (Visual appearance) × 2 (Word type) ANOVA was conducted separately for each group. The ANOVAs showed a significant interaction between Word type and Visual appearance for the Typical Pairing group, *F*(1,11)=6.22, *p*<.03, but not for the Reversed Pairing group, *F*(1,11)=.61, *p*=.45. These results indicated that the training led to different biases in the two groups.

Transfer-2 objects

There was no interaction between Visual appearance and Word type for the transfer-2 objects after the semantic training (Figure 9D), *F*(1,22)=.005, *p*>.94. Nonetheless, RTs were slower when the irrelevant images were S-Greebles relative to A-Greebles, *F*(1,22)=5.03, *p*=.035.

Post-visual training test

Trained and transfer-1 objects

Both groups showed similar effects for the trained and transfer-1 objects after the visual training (Figure 9C), presumably because the object-word pairings were not emphasized. There was an interaction between Word type and Visual appearance, *F*(1,22)=6.1, *p*=.02, revealing a larger effect of Word type for S-Greebles than A-Greebles (same pattern of results as in the pre-test). Also, social features were again classified faster than inanimate features, *F*(1,22)=12.22, *p*=.002. Notably, the 3-way interaction between Pairing, Word type and Visual appearance was not significant, *F*(1,22)=.27, *p*=.61.

Transfer-2 objects

After the visual training, no significant results were found for the transfer-2 objects, *Fs*(1,22)<2.48, *ps*>.12.

Summary of the lexical judgment results

There were several novel and important findings in the lexical judgment task. First, without any prior exposure to the novel objects at pre-test, there appears to be a pre-existing bias for associating symmetric, animal-like objects with social semantics rather than inanimate semantics. Second, this bias could be reduced by explicitly pairing the categories of objects and words in a reversed manner, as shown in the Reversed Pairing group after the semantic training. Third, as reversed visual-semantic pairing was not emphasized during the visual training, the bias for social semantics and symmetric, animal-like objects was found again for both groups in the post-visual training test.

Note that these effects were only found for the trained and transfer-1 objects but not for the transfer-2 objects, perhaps because generalization of semantic effects is often limited (James & Gauthier, 2003; 2004). But it was also surprising that the initial pre-test bias was not found again for these objects. One possibility is that once participants were trained to pay attention to the trained categories, attention to these objects might have competed in some way with attention to the untrained objects over the course of trials where the two were randomized.

Sequential matching task

In the sequential matching task, the results are not reported for trials with the transfer-2 objects as the target objects because the non-match exemplars in those trials were objects in the trained categories in a different color (see Figure 6B). Since the pattern of results was different between the trained and transfer-1 objects, these results are reported separately. For the control groups, analyses were conducted collapsing across all trained and transfer-1 objects and the same results were compared with the results for trained vs. transfer-1 objects in the two training groups.

Comparison between the control group (no training) and the training groups after semantic training

A 3 (Group: no training vs. typical pairing vs. reversed pairing) × 2 (Visual appearance) × 2 (Categorization level: basic vs. subordinate) ANOVA was conducted on RTs in the correct trials and d' for the trained and transfer-1 objects.

Trained objects

First, performance of the two training groups in the post-semantic training test was compared with that of the control group (no training) (Figure 10A & 10B). For the trained objects, overall performance of the three groups was not statistically different, RTs: $F(2,33)<1.25$, $p>.29$; d': $F(2,33)<2.31$, $p>.11$. Basic-level categorization was faster and better than subordinate-level categorization, RTs: $F(1,33)=122.08$, $p<.0001$; d': $F(1,33)=28.85$, $p<.0001$. This difference was reduced in both training groups after the semantic training, relative to the control group, RTs: $F(2,33)=3.9$, $p=.03$; d': $F(2,33)=.62$, $p<.001$. Notably, S-Greebles showed a reduced basic-level advantage compared to A-Greebles, as revealed by an interaction between Visual appearance and Categorization level in RTs, $F(1,33)=6.58$, $p=.015$.

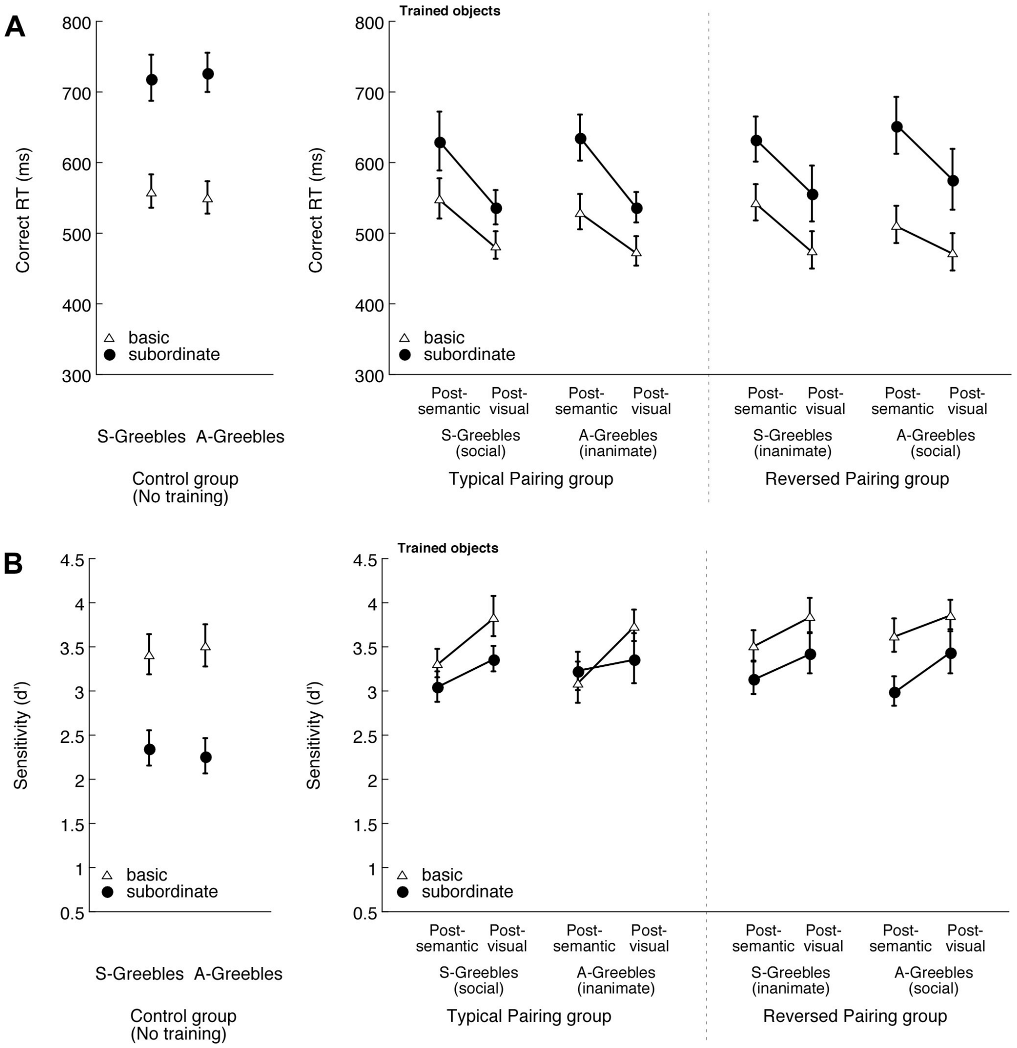


Figure 10. Results of sequential matching task for the trained objects. *Figure 10A:* mean correct RTs for the control group (left), the Typical Pairing group (middle) and the Reversed Pairing group (right). *Figure 10B:* mean d′ for the control group (left), the Typical Pairing group (middle) and the Reversed Pairing group (right). Error bars represent the standard errors of the mean.

Transfer-1 objects

For the transfer-1 objects (Figure 11A & 11B), performance was better in the two training groups compared to the control group in d′, $F(2,33)=5.27$, $p=.01$. Basic-level categorization was faster and better than subordinate-level categorization, RTs: $F(1,33)=152.92$, $p<.0001$; d′: $F(1,33)=88.49$, $p<.0001$. The basic-level advantage was reduced in the two training groups relative to the control group in RTs: *F*(2,33)=9, *p*<.001. Also, S-Greebles showed a reduced basic-level advantage compared to A-Greebles in RTs, *F*(1,33)=9.37, *p*<.005.

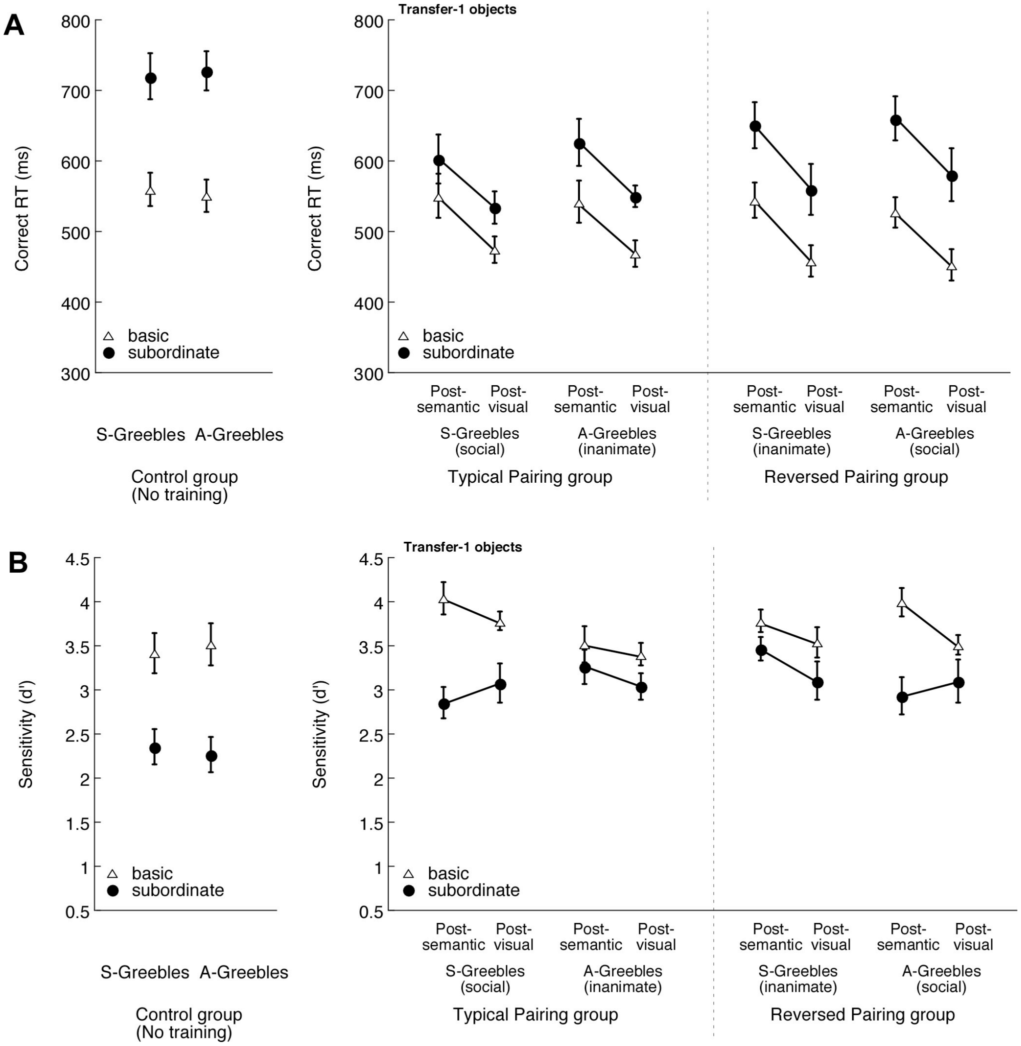


Figure 11. Results of the sequential matching task for the transfer-1 objects. *Figure 11A:* Mean correct RTs the control group (left), the Typical Pairing group (middle) and the Reversed Pairing group (right). *Figure 11B:* mean d′ for the control group (left), the Typical Pairing group (middle) and the Reversed Pairing group (right). Error bars represent the standard errors of the mean.

Apart from an effect of visual appearance in basic- vs. subordinate-level categorization, an effect of semantic associations was also found: The 3-way interaction of Pairing, Visual appearance and Categorization level was significant in d′, $F(2,33)=10.2$, $p<.0005$. Specifically, the basic-level advantage in d′ was only reduced, after the semantic training, in the training groups compared to the control group for Greeble categories with explicit inanimate semantic associations (i.e., A-Greebles for the Typical Pairing group and S-Greebles for the Reversed Pairing group, $ps \leq .02$), but the magnitude of the basic-level advantage was not statistically different between the control and training groups for Greeble categories that were associated explicitly with social semantic features ($ps>.56$).

Interim summary after semantic training

In sum, only two sessions of semantic training was sufficient to reduce the differences between basic- vs. subordinate-level categorization in both the trained and transfer-1 objects, presumably due to practice in discriminating several trained objects. Visual appearance also affected the magnitude of reduction, with S-Greebles facilitated subordinate-level categorization more than A-Greebles. A surprising and novel result was that explicit conceptual associations also had an impact, as the objects in categories that were explicitly associated with inanimate than social features led to larger reduction in the basic-level advantage.

Comparison between the two training groups at post-visual training test

To examine the changes in the basic- vs. subordinate-level advantage between the semantic vs. visual training, performance in the two post-tests was compared between the two training groups. A 2 (Pairing) × 2 (Session) × 2 (Visual appearance) × 2 (Categorization level) was conducted on correct RTs and d′.

Trained objects

For the trained objects, overall performance improved after the visual training in both training groups (Figure 10B), RTs: $F(1,22)=29.09$, $p<.0001$; d′: $F(1,22)=11.33$, $p<.005$. Performance remained better for basic- than subordinate-level categorization, RTs: $F(1,22)=62.75$, $p<.001$; d′: $F(1,22)=22.61$, $p<.0001$. Nonetheless, this difference was reduced from the post-semantic test compared to the post-visual test for both trained S- and A-Greebles, RTs: $F(1,22)=9.8$, $p=.005$. Also, S-Greebles again showed a smaller difference between basic- vs. subordinate-level categorization compared to A-Greebles, RTs: $F(1,22)=7.4$, $p\leq.01$. No other significant results were found in RTs or d′ for the trained objects in the training groups, Fs$(1,22)<2.05$, ps$>.16$.

Transfer-1 objects

For the transfer-1 objects, overall response times were faster in post-visual compared to post-semantic training tests in both training groups (Figure 11B), RTs: $F(1,22)=41.17$, $p<.0001$. The basic-level advantage remained, $F(1,22)=109.6$, $p<.0001$, d′: $F(1,22)=75.35$, $p<.0001$, and was larger for A-Greebles than S-Greebles, RTs: $F(1,22)=12.82$, $p=.002$, and was also larger in the Reversed Pairing group than the Typical Pairing group, RTs: $F(1,22)=7.12$, $p<.02$. Intriguingly, the 4-way interaction of Pairing, Session, Visual appearance and Categorization level was also significant in d′, $F(1,22)=6.69$, $p<.02$. Specifically, the basic-level advantage was comparable between the two post-tests for the object categories associated with inanimate semantics, while the basic-level advantage was reduced from the post-semantic training test compared to the post-visual training test for the object categories associated with social semantics.

Summary of the sequential matching results

There were several interesting findings in the sequential matching task. First, the reduction of the basic-level advantage was found for both the trained and transfer-1 objects after each stage of training, indicating some level of perceptual expertise. Interestingly, S-Greebles facilitated the acquisition of perceptual expertise, since the overall basic-level advantage was smaller for S- than A-Greebles. More intriguingly, semantic effects were also observed in this visual task for the transfer-1 objects. Specifically, object categories associated with inanimate features facilitated the reduction of basic-level advantage sooner than object categories associated with social features (at post-semantic training test), while object categories associated with social features also showed a reduction after both stages of training. There may be no surprise that no difference of pairing was found for the trained objects, since all those objects were trained at the subordinate-level during the visual training. However, for the transfer-1 objects, typical pairing of objects and words facilitated subordinate-level processing, compared to reversed pairing. These results expand from prior evidence that semantic processing is automatically engaged in visual tasks where semantic information is completely irrelevant (James & Gauthier, 2003; 2004), and suggest that semantic processing can influence subordinate-level categorization.

Part-matching task

Part matching was performed twice, once after the semantic training and again after the visual training (see Figure 3 & 6C). Data from one participant in each group were discarded because of a programming error in the post-semantic training session. A 2 (Pairing) × 2 (Visual appearance) × 2 (Alignment: aligned vs. misaligned composites) × 2 (Congruency: congruent vs. incongruent) ANOVA was conducted separately for the trained, transfer-1, and transfer-2 objects in the two post- tests. Correct RTs and d′ were both analyzed, with d′ as the main measure (e.g., Cheung & Gauthier, 2010; Cheung et al., 2008; Richler et al., 2008a; 2008b).

Post-semantic training test

Trained objects

For the trained objects (Figure 12), d′ was better for aligned than misaligned composites, $F(1,20)=7.93$, $p=.01$. Performance was also better for congruent than incongruent trials, $F(1,20)=8.52$, $p<.01$. Critically, there was a 4-way interaction between Pairing, Visual appearance, Alignment and Congruency, $F(1,20)=4.28$, $p=.05$. To explore the effects for each of the Greeble sub-categories (i.e., S- vs. A-

Greebles) in the two groups, 2 (Alignment) × 2 (Congruency) ANOVAs were then conducted. Only the S-Greebles associated with social semantics (in Typical Pairing group) showed a significant interaction between Alignment and Congruency, *F*(1,10)=10.99, *p*<.01, indicating holistic and configural processing. For the other three trained Greeble sub-categories, there were no significant main effect of Alignment, Congruency or their interactions, *F*s(1,10)<3.13, *ps*>.1.

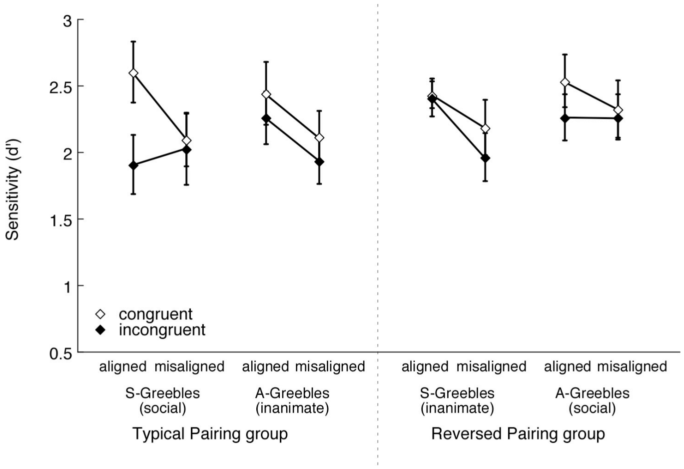


Figure 12. Mean d′ for the trained objects in the part-matching task in the two training groups. Error bars represent the standard errors of the mean for each condition.

Transfer-1 and transfer-2 objects

There was evidence of holistic processing for one case of the transfer-1 objects. For these objects (Figure 13A), overall d′ was better for aligned than misaligned composites, *F*(1,20)=13.05, *p*<.002. There was an interaction between Visual appearance and Alignment, *F*(1,20)=5.5, *p*<.05, with lower performance for misaligned A-Greebles than the other conditions (*ps*<.05). The 3-way interaction between Visual appearance, Alignment and Congruency was also significant, *F*(1,20)=4.71, *p*<.05. When a 2 (Alignment) × 2 (Congruency) ANOVA was conducted to probe the holistic and configural effects within each Greeble sub-category, a significant Alignment × Congruency interaction was found only for A-Greebles associated with social semantics (in the Reversed Pairing group), $F(1,10)=7.84$, $p<.02$. The interaction was not significant in the other sub-categories, $Fs(1,10)\leq1.53$, $ps>.24$.

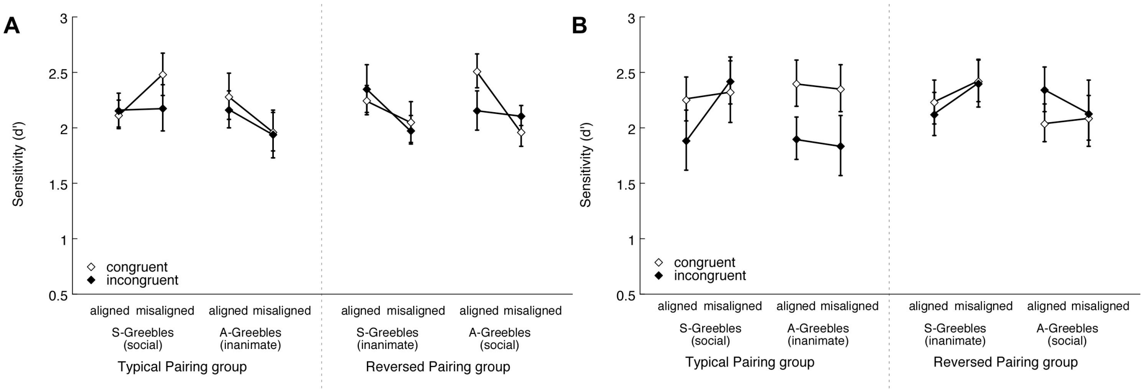


Figure 13. Mean d′ for the transfer objects in the part matching task in the two training groups after the semantic training. *Figure 13A:* Transfer-1 objects. *Figure 13B:* Transfer-2 objects. Error bars represent the standard errors of the mean.

For the transfer-2 objects (Figure 13B), there was no evidence of holistic or configural effects. There was an interaction between Pairing and Congruency, $F(1,20)=6.14$, $p=.02$, an interaction between Pairing, Visual appearance and Congruency, $F(1,20)=4.5$, $p<.05$, and an interaction between Visual appearance, Alignment and Congruency, $F(1,20)=4.67$, $p<.05$.

Interim summary after semantic training

Since holistic processing is a hallmark of perceptual expertise (typically observed after 7-10 hours of lab training or several years of learning), it is surprising that a holistic effect was found for trained S-Greebles and transfer-1 A-Greebles, both categories associated with social semantics. Note however, that the typical holistic effect of expertise is not specific to only trained objects, but generalizes to other exemplars in the expert category (e.g., Gauthier et al., 1998; Wong et al., 2009a). In contrast, the effect observed here did not consistently show generalization of the effects, suggesting that it is likely a different effect, although possibly a precursor of a more category-general perceptual strategy.

Post-visual training test

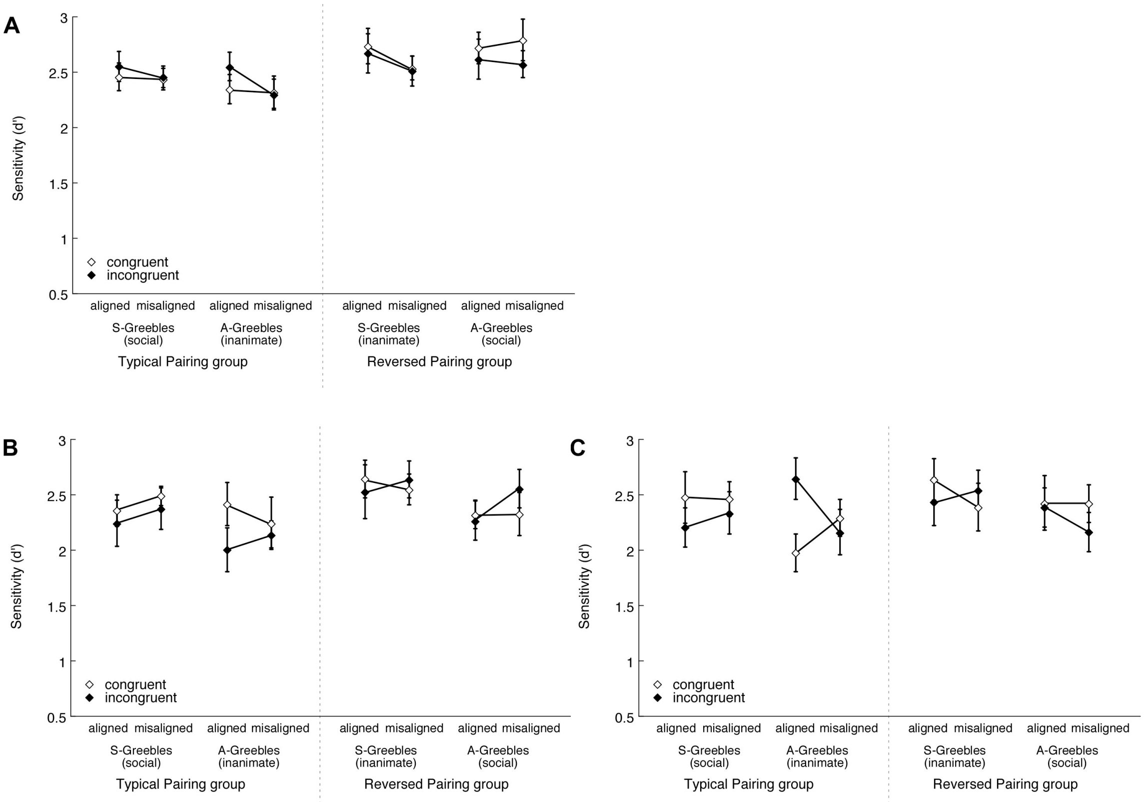


Figure 14. D′ for the trained and transfer objects in the part matching task in the two training groups after the semantic training. *Figure 14A:* Trained objects. *Figure 14B:* Transfer-1 objects. *Figure 14C:* Transfer-2 objects. Error bars represent the standard errors of the mean.

Trained, transfer-1 and transfer-2 objects

Although holistic processing was expected after the visual training, there was no evidence found for this effect (Alignment × Congruency interaction) found for any Greeble sub-categories. For the trained objects (Figure 14A), the only significant result was better performance for aligned than misaligned composites in d′, $F(1,22)=7.27$, $p<.02$. For the transfer-1 objects (Figure 14B), the only significant result was better performance for S-Greebles than A-Greebles in d′, $F(1,22)=5.44$, $p=.03$. For the transfer-2 objects (Figure 14C), there were an interaction between Visual appearance, Alignment and Congruency in d′, $F(1,22)=11.84$, $p<.005$, and a marginally significant interaction between Pairing, Visual Appearance and Congruency, $F(1,22)=3.95$, $p=.06$. None of these suggest holistic or configural processing.

Summary of the part-matching results

The results from the part-matching task were unexpected. It was surprising that an Alignment × Congruency effect was observed for trained S-Greebles and transfer-1 A-Greebles associated with social semantics, and that this effect did not consistently generalize to other exemplars in the categories. More surprisingly, this effect was minimized after the extensive visual individuation training - which would have been expected to *increase* holistic effects (Gauthier & Tarr, 2002; Wong et al., 2009a) - and no holistic effects were found for any of the object categories after the visual training.

To speculate on the reasons for failing to replicate holistic effects following individuation training, it is possible that the training (4 sessions of 1.5 hours each, plus 2 sessions of semantic training for half of the trained objects) was not sufficient for obtaining the holistic effects. Note that holistic processing (Alignment × Congruency interaction) has been found for the novel object set "Ziggerins" (Wong et al., 2009a) but such effect was only obtained for Greebles in the last training session, after participants were trained up to a predefined expertise criterion (Gauthier & Tarr, 2002). Nonetheless, it would be interesting to understand the nature of the effect for Greebles with social semantic associations, and why it disappeared following further visual training.

General Discussion

These results confirm that object and semantic processing are affected by the interactions between visual and conceptual properties of objects. In particular, the lexical judgment task and the sequential matching task provided several novel and important results.

First, the lexical judgment task revealed an implicit bias to associate social semantics with symmetric, animal-like objects, and to associate inanimate semantics with asymmetric, tool-like objects. However, this implicit bias is flexible and can be modified by training, even when new associations are contradictory to the bias (as in the Reversed Pairing group). If the reversed pairings are not emphasized, the implicit bias re-surfaces. These results are very encouraging for a few reasons: This is a rare experimental demonstration for pre-existing visual-conceptual associations for novel, complex objects. Also, the lexical judgment task, developed for this study, showed great potential for use in further investigation of the interactions between visual and conceptual information for objects.

Although I failed to observe holistic processing in the part-matching task by the end of the visual training, there is still some indication of perceptual expertise in the sequential matching task with basic- vs. subordinate-level judgments. Indeed, the basic-level advantage was reduced after the short semantic training, and further reduced after the longer visual individuation training. More importantly, the results in this task not only reflected a training effect, but also showed that matching performance is influenced by the visual and conceptual properties of the objects and their associations. For instance, S-Greebles facilitated subordinate-level processing. Note that the A- and S-Greebles were constructed based on similar principles (4 peripheral parts on a body, including 2 symmetric parts although one was rotated for A-Greebles) but it is possible that S-Greebles may be easier to individuate, as symmetry enhances perceptual fluency (Reber et al., 2004).

Nonetheless, it is novel and surprising that semantic associations can also influence this perceptual task, although only for the transfer-1 objects. For the trained objects, it is probable that semantic effects were not observed because of the extensive individuation practice with all trained objects. For the transfer-1 objects, a larger basic-level advantage was found for objects associated with social semantics than those associated with inanimate semantics, especially immediately after the semantic training. This effect occurs independent of visual effects. Further study is needed to explain why this effect occurs.

In addition, the pairing between visual and conceptual properties also has an impact on the matching task, as subordinate-level processing was facilitated by typical relative to reversed pairings. Note that the two training groups did not differ in training performance at the end of each training stage, but the Reversed Pairing group showed slower performance in matching the transfer-1 objects at the subordinate-level. Again, this is a novel finding which reveals that the pairing of visual-semantic associations can affect visual decisions about objects.

Overall, the current findings demonstrate that there are important interactions between visual and conceptual properties of objects, even when one of the dimensions is task-irrelevant.

CHAPTER III

INTERACTIONS BETWEEN VISUAL APPEARANCE AND CONCEPTUAL KNOWLEDGE IN OBJECT RECOGNITION: NEURAL MEASURES

Introduction

In Chapter 2, I found that visual appearance and conceptual knowledge interact to influence behavioral performance in object recognition. In this chapter, I focus on how these two factors interact to constrain object representations in the brain. First, I will ask whether the explicit social vs. inanimate semantic associations acquired before the first fMRI scan can be automatically engaged in a visual task where these associations are irrelevant, as demonstrated in prior work for modality specific associations (James & Gauthier, 2003). A new question not addressed in prior work is whether reinforcing vs. reversing "typical" visual-semantic associations can influence how the visual system processes the objects.

In addition, I will ask whether any neural effect of semantic associations with objects can survive after a longer period of visual training that was identical for both groups of participants and in which no semantic information was mentioned. This is a model for a situation where novices might generate some semantic information about a novel object category early in training ("these objects look like animals" or "this one looks like a duck"), information that is not diagnostic for individuation but that may nonetheless bias visual learning. To address these questions, all participants in the Typical Pairing and Reversed Pairing groups were scanned twice during the course of the study, once after the semantic training and again after the visual individuation training.

The effect of visual appearance is expected to result in differential responses in the occipital-temporal cortex (Chen et al., 2007; Op de Beeck et al., 2008a; Op de Beeck et al., 2008b; Sasaki et al., 2005). Semantic training effects are also expected. Note that the task used in the scanner only involved visual discrimination, but a main effect of having acquired explicit semantic associations is expected to replicate activity in inferior frontal areas and other semantic processing areas after semantic training (James & Gauthier, 2003; 2004). It is also possible that social vs. inanimate semantics will engage different visual areas, particularly in the lateral vs. medial fusiform gyri (Chao et al., 1999; Mahon & Caramazza, 2009; Martin, 2007). More importantly, the key question in this study is to explore whether visual areas are sensitive to typical vs. reversed pairing of visual and semantic information.

It is an open question whether the visual individuation training will reduce semantic training effects, since no semantic associations are mentioned. Instead, visual individuation training may be expected to lead to increases in activity in the right fusiform gyrus as in prior studies (Gauthier et al., 1999; Gauthier et al., 2000; Wong et al., 2009b). One original goal of the current study was to test for expertise effects as a function of visual-semantic pairing. To reveal this effect, activity associated with objects of expertise would be compared with activity for control objects from an untrained category. As will be discussed later in this chapter, despite the use of different central shapes and colors to signify differences between trained vs. untrained Greeble categories, there was a non-negligible amount of generalization of learned semantic associations to the untrained objects (transfer-2 objects), rendering them unfit to be used as a baseline. In this context, it is unfortunately impossible to localize category-specific expertise effects for trained objects because all objects used in the fMRI study showed evidence of the semantic associations learned in the very first phase of the study with some of the objects that were used during visual training. What *can* be compared is the activity for different Greebles categories, for which participants in all cases demonstrate some level of expertise, but which differ in their geometry and semantic associations. Therefore, despite these limitations, interaction effects between visual appearance and conceptual knowledge can still be examined after the visual training. Any difference between the groups that received typical vs. reversed associations would reveal such an interaction.

Whole brain analyses will first be performed to search for interaction effects by comparing activations between the Typical Pairing vs. Reversed Pairing groups. Category-selective regions for object and semantic processing will also be defined to test for any interactions in these areas. These regions include the fusiform face area (FFA), the parahippocampal place area (PPA), the lateral occipital complex (LOC), posterior cingulate gyrus (CG) and supramarginal gyrus (SMG).

Methods

Participants

All 24 participants in the training groups (12 in each group) took part in two fMRI sessions, one following semantic training and one following the visual training.

Apparatus and Stimuli

All scans were completed on a 3T Philips Intera Achieva scanner at the Vanderbilt University Institute of Imaging Science. All testing was conducted on a Power Mac computer using Matlab (MathWorks, Natick MA) with the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). The stimuli were presented on a LCD panel and back-projected on a screen. Participants viewed the stimuli through a mirror mounted on top of an RF coil above their head.

The post-semantic training and post-visual training fMRI sessions were identical except for the randomization of trials. The 6 main experimental runs (Greeble runs, Figure 15A) used Greeble images that were identical to those used in the behavioral pre-/post-training tests. Within each set of 24 S-Greebles or 24 A-Greebles, the post-semantic training scan used 6 trained objects, 12 transfer-1 objects and 6 transfer-2 objects. Remember that the transfer-1 objects were new exemplars of the same color and central body shape as the trained Greebles, whereas transfer-2 objects had a different color and central body shape. The post-visual training scan also included the same 24 S-Greebles and 24 A-Greebles, in each case including 12 trained objects (the 6 used during both the semantic and visual training and 6 additional ones used only in the visual training), 6 transfer-1 objects and 6 transfer-2 objects. All objects were presented at either 6˚ or 12˚ (the training orientations were 0˚ and 18˚).

Each scan also included two sets of localizer runs. Two Visual localizer runs (Figure 15B) used images of faces, scenes, common objects, and phase-scrambled objects, with a total of 36 images from each category. Two Semantic localizer runs used 18 social-relevant words and 18 inanimate-relevant words.

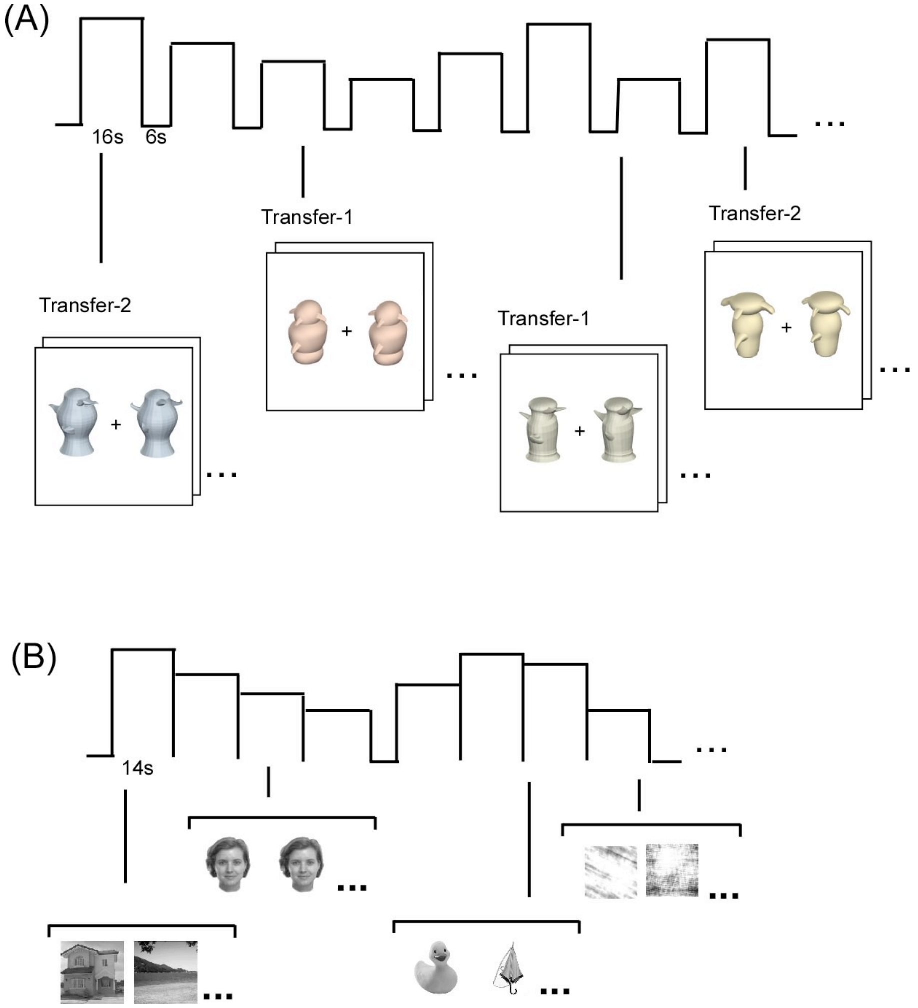


Figure 15. Sample Greeble runs and Visual localizer runs in the fMRI sessions. *Figure 15A*: A transfer Greeble run consisted of 16-second blocks of transfer-1 and transfer-2 S- and A-Greebles, with 6-second fixation in between blocks. *Figure 15B*: A Visual localizer run consisted of 14-second blocks of faces, common objects, scenes and phase-scrambled objects.

Procedures and Design

In the 6 Greeble runs (Figure 15A), a simultaneous matching task was adopted because this task was used before to demonstrate semantic training effects (e.g., James & Gauthier, 2003; 2004). There were 2 kinds of Greeble runs (3 runs for each kind). The Greebles used in the first type of run (Trained runs) were those experienced either during both semantic and visual training, or only during visual training. The Greebles used in the second type of run (Transfer runs, Figure 15A) were either transfer-1 or transfer-2 objects that were not shown during training. Each run included 2 (Visual appearance: S-Greebles vs. A-Greebles) × 2 (Training status in the Trained runs: objects trained in both stages vs. objects trained in the second stage only; or Training status in the Transfer runs: transfer-1 vs. transfer-2 objects) × 4 repetitions of blocks. The trials in these runs were blocked by visual appearance and training status. On each trial, participants judged whether two objects presented simultaneously were identical or different (half of the trials showed identical pairs, half of the trials showed different objects of the same training status). Each block lasted 16 seconds, with object pairs from each category presented for 3 seconds each, followed by a 200ms fixation. Six-second fixation periods were included between blocks, and 12-second fixation periods were included at the beginning and end of a run. Each run lasted 6.2 minutes.

In the Visual localizer runs (Figure 15B), there were four types of blocks (faces, common objects, places, phase-scrambled objects), with four blocks for each category. Each block lasted 14 seconds, with 14 images from a category presented one at a time, each for 800ms following a 200ms fixation. Participants performed a one-back identity judgment task and were asked to make a key response as fast as possible whenever they saw two identical images in a row, which happened on about 10% of trials. Six-second fixation periods were included after every cycle of the four conditions. Twelve-second fixation periods were included at the beginning and end of each run. Each run lasted 4 minutes and 26 seconds.

In the Semantic localizer runs, there were two types of blocks (social vs. inanimate semantic features). Social or inanimate semantic features (not used in semantic training or the lexical judgment task) were shown one at a time, with instructions to imagine a person or an object that matched the feature. There were 6 blocks for each type of semantic feature in each run. Each block lasted 16 seconds, with 4 words from each category presented for 3 seconds each, followed by a 500ms fixation, and a 2-second blank screen was also included at the end of each block. Twelve-second fixation blocks were added at the beginning and end of each run. Each run lasted 3 minutes and 34 seconds. No behavioral data was collected during these runs.

Imaging parameters and analyses

A 3-Tesla, whole body Philips MRI system and a birdcage head coil located at the Vanderbilt Medical Center (Nashville, USA) were used to perform the imaging study. The field of view was 22.4 × 22.4 × 11.85 cm, with an in-plane acquisition resolution of 64 × 64 pixels, a reconstruction size of 80 × 80 pixels, and 34 contiguous axial scan planes per volume, resulting in a voxel size of 2.8 × 2.8 × 3 mm. A 0.5 mm gap was added in between each volume to achieve maximal brain coverage. Images were collected using a T2*-weighted EPI acquisition (TE=35 ms, TR=2000ms, flip angle=79°) for blood oxygen-level dependent (BOLD) based imaging. High-resolution T1-weighted anatomical volumes were also acquired using a 3-D fast spoiled grass (FSPGR) acquisition (TI=218ms, TE=3.68 ms, TR=8ms, flip angle=5°). The functional data was analyzed using Brain Voyager (http://brainvoyager.com). Data preprocessing included 3D motion correction, slice scan time correction, temporal filtering (3 cycles/run high-pass), and spatial smoothing (4 mm FWHM Gaussian). A general linear model (GLM) analysis computed the correlation of predictor variables or functions with the recorded activation data (criterion variables) across scanning sessions. The predictor functions were based on the blocked stimulus presentation paradigm of the particular run being analyzed and represented an estimate of the predicted hemodynamic response during that run. To properly model the hemodynamic response, the predictors were represented as stimulus protocol boxcar functions convolved with the appropriate gamma function ($\Delta=2.5$, $\tau=1.25$) estimate of a typical hemodynamic response (Boynton et al., 1996).

Whole brain analyses

Random-effects GLM analyses were conducted for whole brain contrasts. Specifically, these contrasts were thresholded at $p<.01$ (uncorrected) after which minimum cluster size with probability $p<.01$ (corrected) was calculated using a bootstrapping procedure implemented in Brain Voyager's Cluster Threshold Estimator (Forman, Cohen et al., 1995; Goebel, Esposito et al., 2006).

ROI analyses

Regions of interest (ROIs) were identified for each participant by comparing activation for different image or semantic conditions in the localizer runs using fixed effects analyses with a threshold of q(FDR)<.05 and a cluster threshold of at least over 30 mm^3 voxels but less than 1000mm^3 around the single voxel that showed the peak activation. FDR (False Discovery Rate) is a multiple comparison correction method that controls for the expected proportion of false positive voxels among those that are above the threshold (Genovese et al., 2002).

To define object-sensitive areas along the ventral visual stream, a contrast comparing common objects vs. phase-scrambled objects was used (Grill-Spector et al., 1998; Grill-Spector et al., 2000), with areas in the left lateral occipital complex (LO) and in the left medial fusiform gyrus (FG) successfully defined in 23 participants. Category-selective areas were defined by comparing faces to common objects and scenes (Epstein & Kanwisher, 1998; Kanwisher et al., 1997), with the right fusiform face area (FFA) localized in 22 participants and the left parahippocampal place area (PPA) localized in all participants. These visual ROIs from a representative participant were shown in Figure 21.

Areas associated with social and inanimate conceptual processing were localized by contrasting social and inanimate semantics (e.g., Chao et al., 1999; Martin, 2007; Mahon & Caramazza, 2009), with a region in the left posterior cingulate gyrus (CG) revealing higher activity for social than inanimate semantics in 22 participants, and a region in the left supramaginal gyrus (SMG) revealing higher activity for inanimate than social semantics in 20 participants. The semantic ROIs from a representative participant were shown in Figure 22.

BOLD responses in the ROIs for each condition in the Greeble runs were first normalized by subtracting activity levels measured during the fixation periods. An ANOVA was then conducted for each ROI, separately for the two scans, with one between-subjects factor of Pairing (Typical vs. Reversed) and one within-subjects factor of Visual appearance (S-Greebles vs. A-Greebles). Note that by the nature of the design, the interaction between visual and semantic information is included within the Pairing factor (e.g., participants who attached social semantics to S-Greebles also attached inanimate semantics to A-Greebles). For this reason, any effect of semantics or visual appearance must be interpreted as possibly dependent on this specific pairing context. Note that in the Figures that follow, the Semantics factor that is not explicit in the analyses (the ANOVA cross Pairing Group with Visual Appearance) is indicated in color. This is to make clear which associations were made in each condition.

Predictions

First, in the post-semantic training scan, it was expected that areas recruited by semantic processing would be automatically engaged in the visual judgment task, even though the associations were task-irrelevant (James & Gauthier, 2003; 2004). It was also predicted that objects associated with social and inanimate semantics would engage different neural substrates during the visual matching task (Chao et al., 1999; Martin & Chao, 2001; Caramazza & Mahon, 2006).

The ventral and dorsal visual streams were expected to be sensitive to visual appearance (e.g., symmetry, Sasaki et al., 2005; Chen et al., 2007). One specific prediction was that S-Greebles would be more likely to engage the fusiform 'face' area (FFA) than A-Greebles, because S-Greebles appear face-like (Kanwisher, 2000). In contrast, the object- and scene-selective parahippocampal 'place' area (PPA) in the medial fusiform gyrus would be more likely to be recruited for A-Greebles because of their visual resemblance to tools (Chao et al., 1999).

Critically, any difference between the Typical Pairing and Reversed Pairing groups would reveal effects that depend on the specific associations between shapes and semantic categories, suggesting an interaction between visual appearance and conceptual knowledge in object recognition. Recall that both groups learned the same objects and the same semantic features and only differed in the manner in which they were paired. The interaction would be expected in areas sensitive to conflict (e.g., cingulate gyrus, Carter et al., 1998). It was of interest whether the interaction would also be found in visual areas, including the FFA and PPA. A recent study has shown that activity in these two areas can be modulated by language processing specifically related to the preferred category for the areas (e.g., faces for the FFA and places for the PPA, Aziz-Zadeh et al., 2008).

There was no strong prediction regarding how the Pairing effects would be modulated by the visual training. The results in the post-visual training scan can provide insights into how long lasting the Pairing effects are.

Note that semantic training effects might be specific to the trained objects, since novel exemplars from the trained object set did not show modality-specific training in a previous study (James & Gauthier, 2003). In contrast, generalization of the visual individuation skills to unlearned exemplars within the same object space is a hallmark of perceptual expertise (Gauthier et al., 1998; Gauthier & Tarr, 2002; Wong et al., 2009a; Wong et al., 2009b; Bukach et al., in press). With the combination of the training paradigms, it was of interest to test the extent to which the training effects could generalize to the two types of transfer objects in this study.

Results

Behavioral results in the Greeble runs

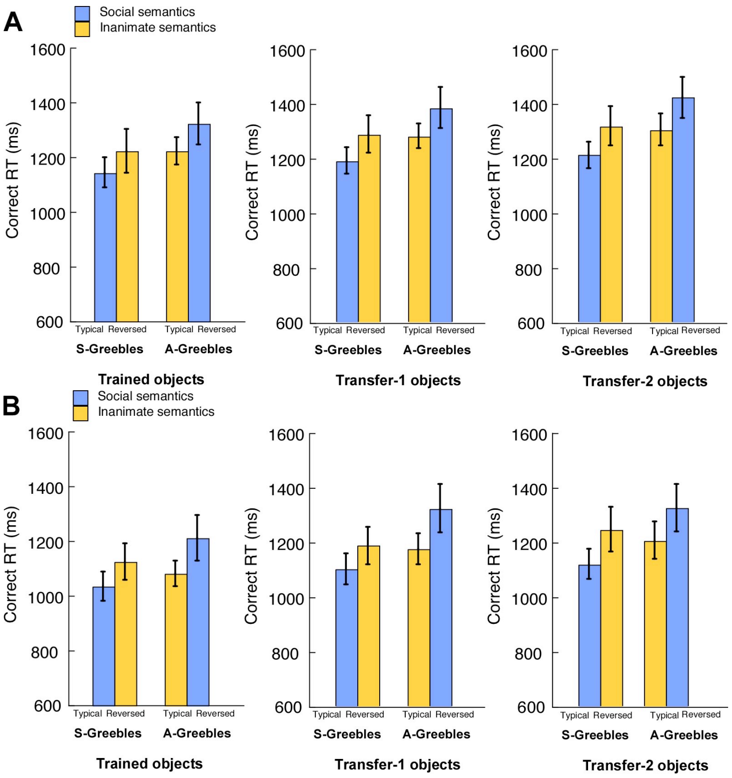


Figure 16. Mean correct RTs in the Greeble runs. *Figure 16A*: results in the post-semantic training scan for the trained objects (left), transfer-1 objects (middle) and transfer-2 objects (right). *Figure 16B*: results in the post-visual training scan for the trained objects (left), transfer-1 objects (middle) and transfer-2 objects (right). Error bars represent the standard errors of the mean.

For each type of Greebles (trained, transfer-1 and transfer-2 objects), I conducted a 2 × 2 ANOVA with the between-subjects factor Pairing (Typical vs. Reversed) and the within-subject factor Visual appearance (S- vs. A-Greebles) on d′ and RTs for correct trials (Figure 16). All ANOVAs revealed slower performance for A-Greebles than S-Greebles in RTs, Fs(1,22)>11.72, ps<.0025. RTs revealed no main effect of Pairing, Fs(1,22)≤1.57, ps≥.22, or interaction between Pairing and Visual appearance, Fs(1,22)≤3.04, ps≥.095. D′ showed no significant results, Fs(1,22)≤3, ps≥.092.

Results collapsed across the two scans

I analyzed the behavior in the scanner separately for each Greeble type because the main analyses for the imaging data were conducted separately for the different Greeble sub-categories in the two scans. But it would nonetheless be informative to know whether overall performance in the behavioral task differed across the various subsets of objects and across the two scans. Therefore, a 2 (Pairing) × 2 (Visual appearance) × 3 (Training status) × 2 (Session) ANOVA was conducted on d′ and RTs for correct trials for all trials across both scans. Performance improved in the post-visual training relative to the post-semantic training scans, as revealed by faster RTs in the second scan than in the first scan, F(1,22)=21.64, p<.0001. There was also a main effect of Training status, RTs: F(2,44)=42.07, p<.0001; d′: F(2,44)=26.61, p<.0001, with faster and better performance for the trained objects compared to the transfer-1 objects (ps<.0001 in RTs and d′), which in turn were matched faster than the transfer-2 objects (p<.013

in RTs). In d′, the interaction between Pairing and Training status approached significance, $F(2,44)=3.08$, $p=.056$. There were no other significant results, $Fs \leq 2.64$, $ps \geq .083$.

Imaging results

Two types of analyses are reported here for the imaging data (please refer to p.74-76 for the statistical criteria for both types of analysis). First, statistical parametric maps generated would reveal brain areas sensitive to the semantic training, visual appearance of objects and the pairing between visual appearance and explicit semantic associations. Second, results from several ROI involved in object or semantic processing are also reported.

Post-semantic training scan results

Whole brain contrasts

Effects of semantic training

A first analysis examined whether the semantic training was effective by comparing activity for the trained vs. transfer-1 objects in the Trained runs (Figure 17). Replicating previous findings (James & Gauthier, 2003; 2004), several areas in the left hemisphere revealed higher activity for the trained objects that had been associated with explicit semantic knowledge relative to the transfer-1 objects, which were introduced later during the visual training without any explicit semantic features. In particular, superior frontal cortex [1], superior temporal gyrus [4], posterior cingulate gyrus [5] and precuneus [6] have been implicated for semantic processing in visual discrimination tasks (James & Gauthier, 2003; 2004); whereas supramarginal gyrus [2] and angular gyrus [3] are often involved in lexical processing (Price, 2010).

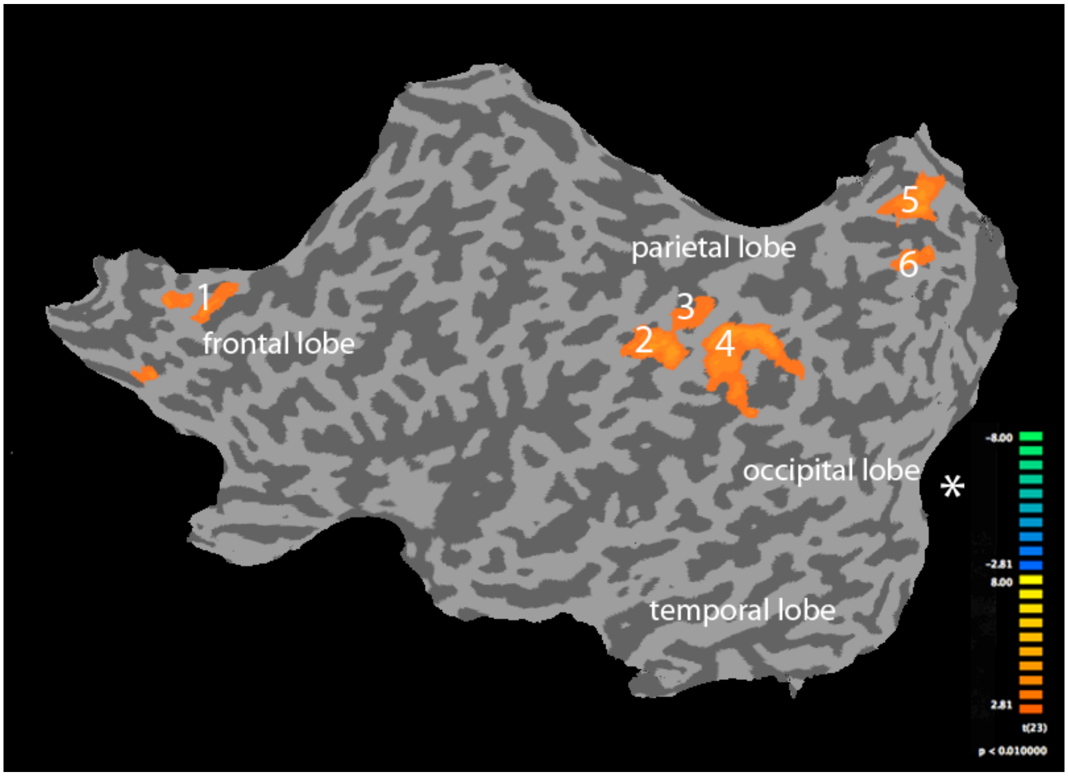


Figure 17. Semantic training effects in the post-semantic training scan, presented on a flattened left hemisphere of a representative participant. Dark grey areas represent sulci whereas light grey areas are gyri; the most posterior point on the occipital lobe is indicated by an asterisk. Orange clusters indicate regions more active for the trained objects associated with explicit semantics than the transfer-1 objects. 1: superior frontal gyrus (-12, 46, 35); 2: supramarginal gyrus (-50, -44, 37); 3: angular gyrus (-46, -53, 39); 4: superior temporal gyrus (-40, -59, -40); 5: posterior cingulate gyrus (-2, -51, 22); 6: Precuneus (-7, -61, 23). The contrast was thresholded at *p*<.01 (uncorrected) after which minimum cluster size with probability *p*<.01 (corrected) was calculated using a bootstrapping procedure.

Associations with social semantics vs. inanimate semantics also led to differential activations in the simultaneous matching task (Figure 18). Specifically, objects explicitly associated with social vs. inanimate semantics were contrasted, collapsing across the two training groups. Six areas were found to be more engaged for inanimate than social semantics, including the right medial/inferior frontal gyrus [5] which is involved in semantic processing. In addition, the left dorsal occipital area [2] in the visual system also showed such effect of semantics. No areas were more active for social than inanimate semantics at this threshold.

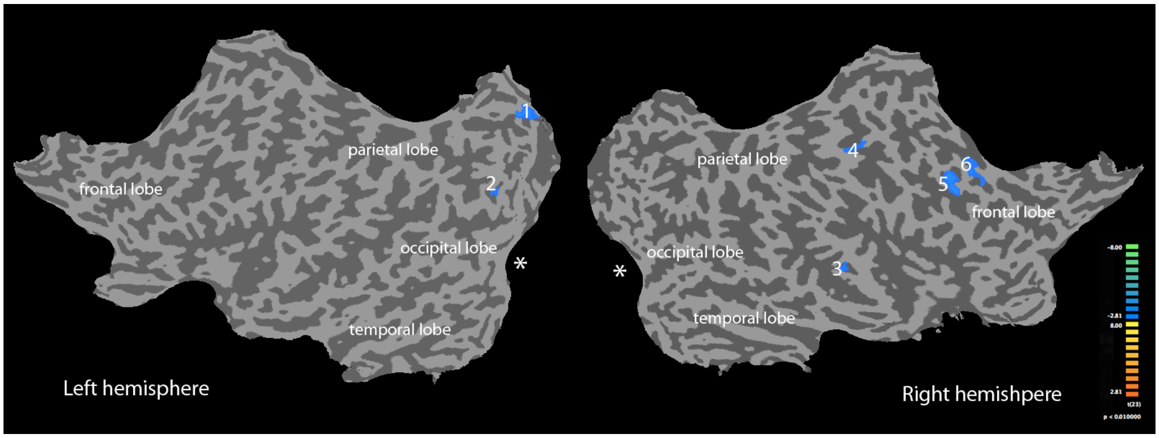


Figure 18. Differential effects for objects with explicit social vs. inanimate semantic associations in the post-semantic training scan, presented on the two flattened hemispheres of a representative participant. Blue clusters represent higher selectivity for objects associated with inanimate relative to social semantics. 1: posterior cingulate gyrus (-8, -50, 7), 2: dorsal occipital gyrus (-19, -85, 26), 3: planum temporale (51, -14, 5), 4: postcentral gyrus (34, -22, 49), 5: medial/inferior frontal gyrus (34, 14, 30), 6: medial frontal gyrus (30, 24, 40). The contrast was thresholded at $p<.01$ (uncorrected) after which minimum cluster size with probability $p<.01$ (corrected) was calculated using a bootstrapping procedure.

Effects of visual appearance

To evaluate the effect of visual appearance, the response to S- vs. A-Greebles was compared, combining across both training groups. This comparison was first conducted separately for each sub-category of the trained, transfer-1 and transfer-2 objects. All maps showed highly similar activations. Figure 19A shows the combined results for trained and transfer-1 objects in the Trained runs. Figure 19B shows the effects for the transfer-2 objects. As expected, visual appearance drove effects in many visual areas, with higher activity for A- than S-Greebles in the in bilateral medial object-selective fusiform areas [1, 2], in a widespread and bilateral region of ventral and dorsal occipital cortex [3] and in the superior parietal lobe [4]. Note that the activation pattern in the visual areas is highly similar between all Greeble subcategories (Figure 19A & 19B). However, higher activity for A- than S-Greebles in the parietal areas such as the left angular gyrus [5] and the left supramarginal gyrus [6] was only observed for the trained and transfer-1 objects (Figure 19A). While A-Greebles were more difficult to match than S-Greebles (see the Behavioral results in the scanner on p.78-80), the widespread activity for A-Greebles cannot merely be accounted for by task difficulty, since S-Greebles also showed more activity than A-Greebles in a few visual areas, including bilateral lingual gyri [7]. Instead, the more widespread activity for A- than S-Greebles might be due to an increased processing load from asymmetry (Reber et al., 2004) or the tool-like resemblance of A-Greebles activating areas including the parietal and medial fusiform areas (e.g. Chao et al., 1999).

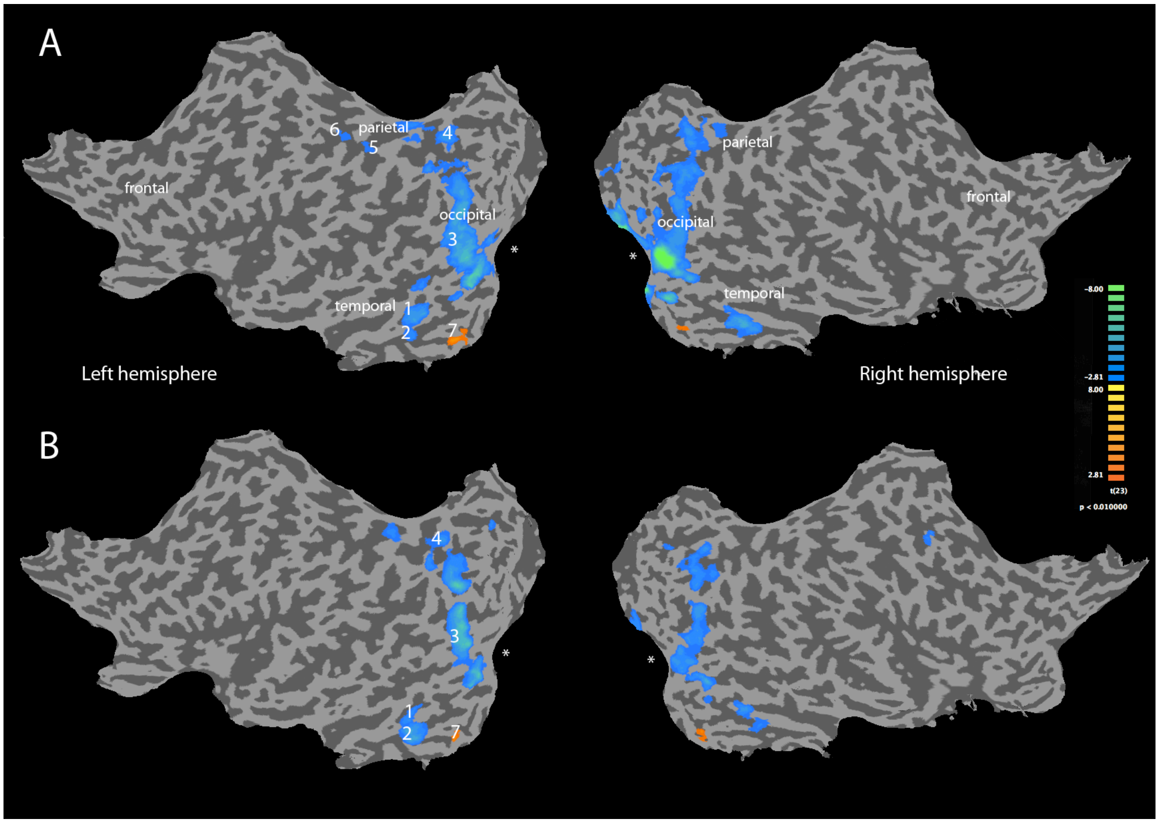


Figure 19. Differential effects for A-Greebles and S-Greebles in the post-semantic training scan, presented on the two flattened hemispheres of a representative participant. Blue clusters represent higher selectivity for A-Greebles than S-Greebles; orange clusters represent higher selectivity for S-Greebles than A-Greebles. *Figure 19A*: An averaged map for the trained and transfer-1 objects in the Trained runs. 1: medial fusiform gyrus (-30, -55, -11), 2: hippocampal gyrus (-26, -42, -8); 3: ventral and dorsal occipital areas (-31, -80, 9), 4: superior parietal lobe (-18, -42, -66); 5: angular gyrus (-33, -46, 35), 6: supramarginal gyrus (-30, -35, 36), 7: lingual gyrus (-6, -62, -4). *Figure 19B*: An averaged map for the transfer-2 objects. The contrast was thresholded at $p<.01$ (uncorrected) after which minimum cluster size with probability $p<.01$ (corrected) was calculated using a bootstrapping procedure.

Interactions between visual-semantic pairing

One of the main goals of this study was to investigate whether Typical vs. Reserved Pairings between visual appearance and conceptual knowledge would have an influence in the brain and where this would be found. A whole brain contrast was first conducted separately for the trained, transfer-1 and transfer-2 objects. Since the effects for the trained and transfer-1 objects in the Trained runs were once again highly similar, the results were combined to maximize power. The training groups showed significant differences between the two groups in several areas for these objects (Figure 20A). Specifically, the Reversed Pairing group showed more activity in a widespread occipital-parietal network including the bilateral posterior inferior fusiform and occipital gyri, dorsal occipital gyri [3-7], bilateral precentral gyri [1, 10], left supramarginal gyrus [2], and the right occipital-parietal junction [9]. An effect of difficulty would be unable to account for these results as the Typical Pairing group also showed higher activity than the Reversed Pairing group in several occipital-parietal areas, including the left precuneus [8], right lingual gyrus [11] and right superior parietal lobe [12]. In sum, the nature of the visual-semantic pairing affected which areas were more strongly recruited by objects.

Unexpectedly, the transfer-2 objects also showed effects of Pairing (Figure 20B), with more activity for the Reversed Pairing group than the Typical Pairing group in several occipital-parietal areas overlapping those found for the trained and transfer-1 objects, including the bilateral precentral gyri [1, 10], the left supramarginal gyrus [2], the ventral and dorsal occipital-temporal cortex [3-7], and the right occipital-parietal junction [9]. These results show that the effects of visual-semantic pairing, learned only for the trained objects, generalized to the transfer-2 objects in a task where semantic associations were completely irrelevant.

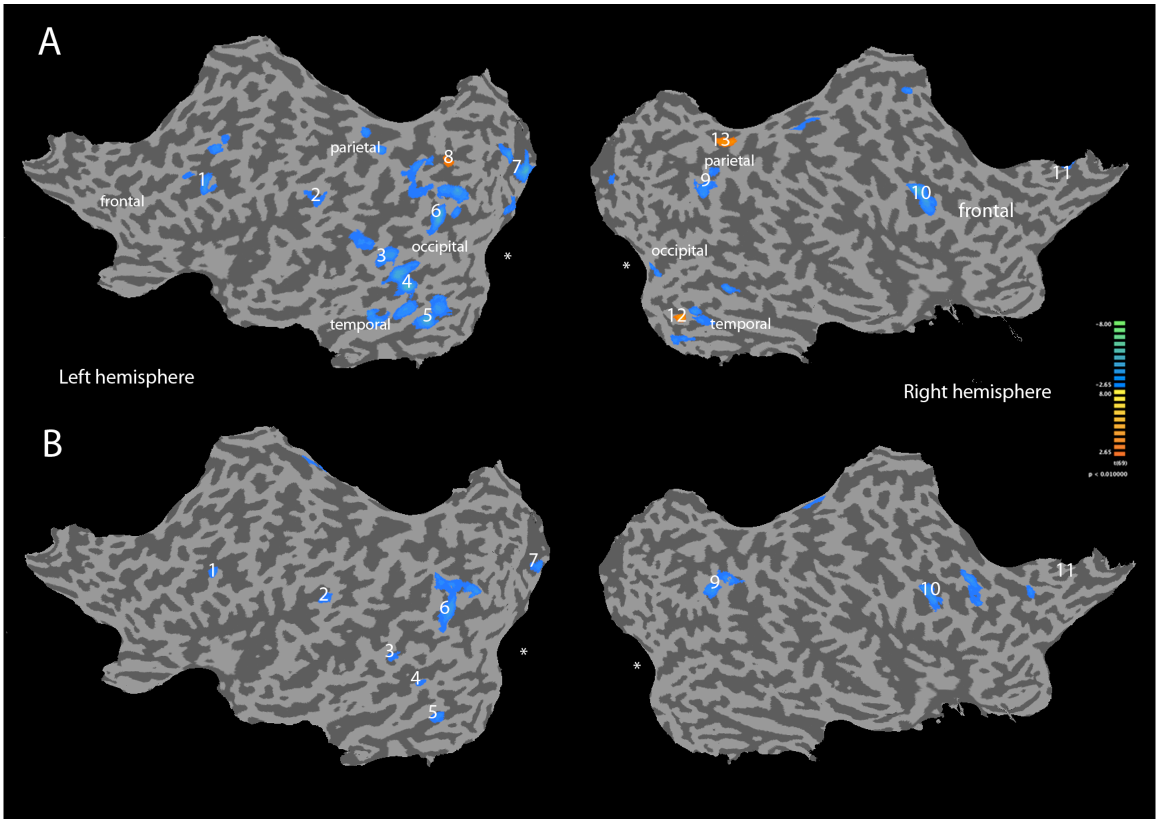


Figure 20. Differential activations for the Typical vs. Reversed Pairing groups in the post-semantic training scan, presented on the two flattened hemispheres of a representative participant. Blue clusters represent higher selectivity in the Reversed Pairing than Typical Pairing groups; orange clusters represent higher selectivity in the Typical Pairing than the Reversed Pairing groups. *Figure 20A*: An averaged map for the trained and transfer-1 objects in the Trained runs. 1: precentral gyrus: (-39, 4, 24); 2: supramarginal gyrus (-54,-40, 29); 3: middle temporal gyrus (-50, -57, 2); 4: inferior temporal gyrus (-37, -58, -2); 5: posterior fusiform area (-23, -61, -9); 6: middle temporal gyrus (-37, -68, 18); 7: striate area (-13, -76, 6); 8: precuneus (-20, -61, 26); 9: parietal-occipital junction (29, -75, 35); 10: right precentral gyrus (39, 4, 24); 11: right inferior frontal gyrus (39, 4, 22); 12: right lingual gyrus (14, -73, -15); 13: right superior parietal lobe (19, -55, 48). *Figure 20B*: An averaged map for the transfer-2 objects. The contrast was thresholded at *p*<.01 (uncorrected) after which minimum cluster size with probability *p*<.01 (corrected) was calculated using a bootstrapping procedure. The degrees of freedom in this contrast was calculated by t(total number of participants-1 × total number of conditions in each run-1) = t(24-1 × 4-1) = t(69).

Regions of interest (ROI) results

Activity for the trained objects in the visual ROIs

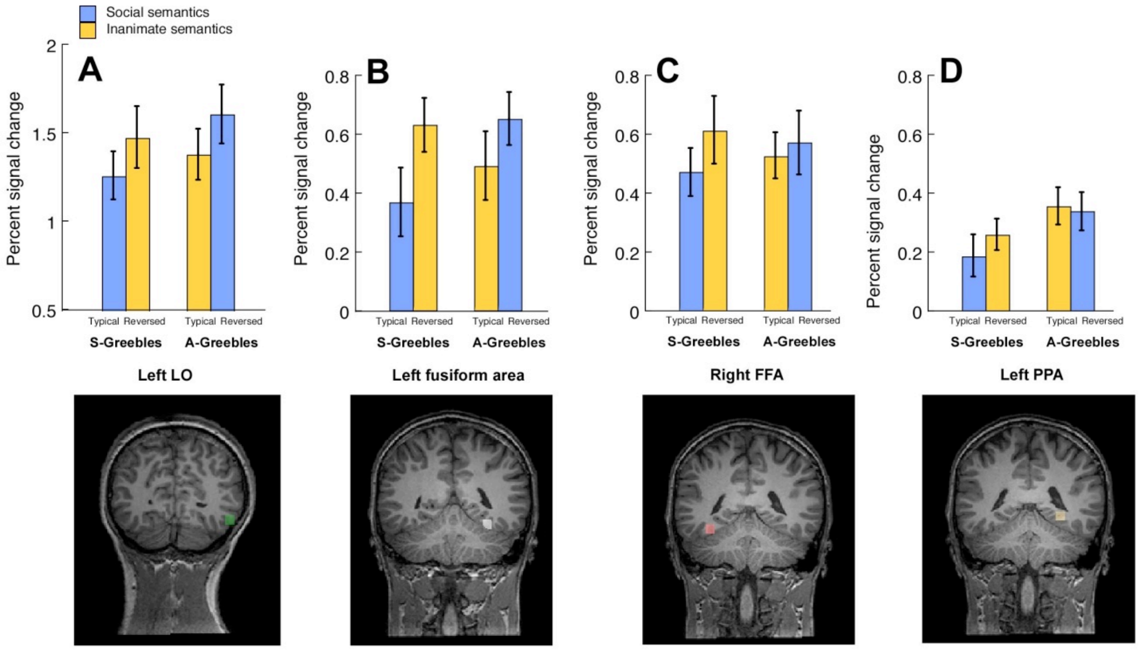


Figure 21. Mean percent signal change for the trained objects in the visual ROIs during the post-semantic training scan. The ROIs were defined for each participant and the ROIs of one representative participant are shown here. *Figure 21A*: Left lateral occipital area (LO: -45, -71, -12). *Figure 21B*: Left medial fusiform area (FG: -29, -46, -13). *Figure 21C*: Right fusiform face area (FFA: 34, -42, -14). *Figure 21D*: Left parahippocampal area (PPA: -27, -42, -6). Error bars show standard error of the mean.

Left lateral occipital complex (objects > phased-scrambled objects) (LO)

As expected, the shape-selective left LO was sensitive to visual object properties (Figure 21A): Higher activity was found for A- than S-Greebles in this region, *F*(1,21)=25.53, *p*<.0001. The main effect of Pairing and the interaction between Pairing and Visual appearance were not significant, *Fs*(1,21)≤.99, *ps*>.33.

Left medial fusiform gyrus (objects > phased-scrambled objects) (FG)

This object-selective area in the medial fusiform gyrus also revealed higher activity for A- than S-Greebles (Figure 21B), *F*(1,21)=9.14, *p*=.0065. The main effect of Pairing was not significant, *F*(1,21)=2.15, *p*>.15, but interestingly, there was an interaction between Pairing and Visual appearance, *F*(1,21)=4.52, *p*=.045. Specifically, the associations had a larger difference on the activity observed for S- compared to A-Greebles, which produced more activity if they had been paired with inanimate than social semantics.

Right fusiform face area (faces > objects and scenes) (FFA)

Researchers have sometimes postulated that the FFA would be sensitive to visual-semantic interactions, in particular if the geometry of the objects was face-like and could therefore be thought of as animate (Cox et al., 2004; Op de Beeck et al., 2006). However, only the interaction between Pairing and Visual appearance approached significance (Figure 21C), *F*(1,20)=3.7, *p*=.07, and no significant results were observed in this region, *F*s(1,20)≤.47, *p*s≥.49.

Left parahippocampal place area (objects and scenes > faces) (PPA)

Tool-like A-Greebles also showed higher activity than face-like S-Greebles in this region (Figure 21D), *F*(1,22)=19.98, *p*<.0002. An interaction between Pairing and Visual appearance was also observed, *F*(1,22)=4.85, *p*<.04, revealing lower activity for S-Greebles associated with social semantics than those associated with inanimate semantics (*p*<.02) but there was no Pairing effect for A-Greebles (*p*=.56).

Activity for the trained objects in the semantic ROIs

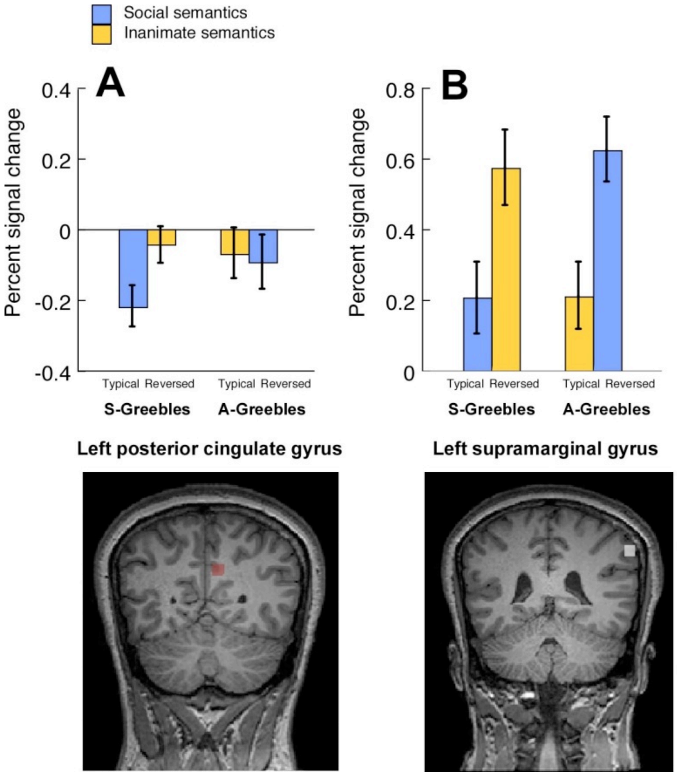


Figure 22. Mean percent signal change for the trained objects in the semantic ROIs during the post-semantic training scan. *Figure 22A*: Left posterior cingulate gyrus (CG: -7, -57, 27). *Figure 22B*: Left supramarginal gyrus (SMG: -59, -40, 39). Error bars show standard error of the mean.

Left posterior cingulate gyrus (social > inanimate semantics) (CG)

This area revealed an interaction between Pairing and Visual appearance for the trained objects (Figure 22A), *F*(1,20)=5.16, *p*=.034: More deactivation for S-Greebles associated with social semantics compared to those associated with inanimate semantics (*p*=.01), while no difference was found for A-Greebles associated with the different types of semantics (*p*>.7). There were no other significant results, *Fs*(1,18)≤1.4, *ps*≥.24.

Left supramarginal gyrus (inanimate > social semantics) (SMG)

Higher activity was observed for the trained objects in this region for the Reversed Pairing group compared to the Typical Pairing group, regardless of visual appearance (Figure 22B), *F*(1,18)=8.19, *p*=.01. This result suggests that this region may be a locus for learning different new and/or contradictory associations. No other effect was significant, *Fs*(1,18)≤.87, *ps*≥.36.

Activity for the transfer-1 and transfer-2 objects in the visual ROIs

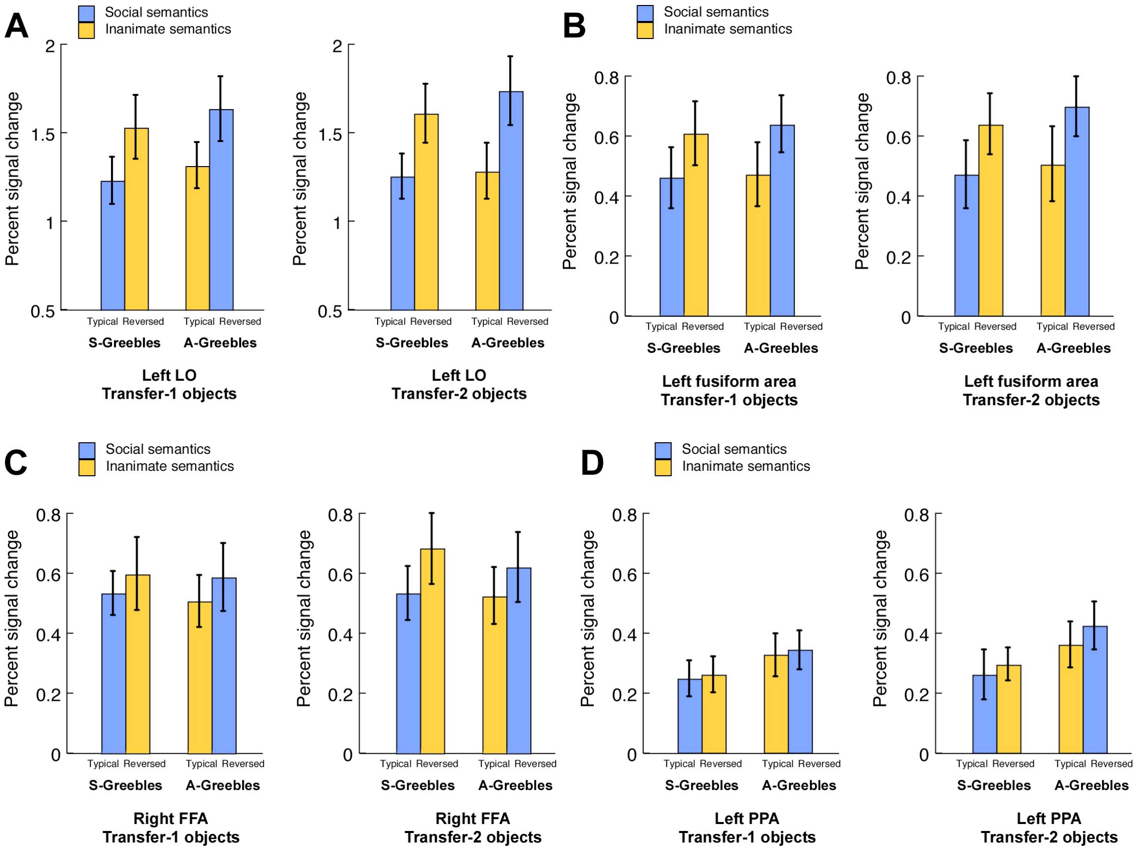


Figure 23. Mean percent signal change for the transfer-1 and transfer-2 objects in the visual ROIs during the post-semantic training scan. *Figure 23A*: Left LO. *Figure 23B*: Left FG. *Figure 23C*: Right FFA. *Figure 23D*: Left PPA. Error bars show standard error of the mean.

Activity for the transfer-1 and transfer-2 objects was highly similar in all visual ROIs. Specifically, in the left LO (Figure 23A) and left PPA (Figure 23D), higher activity was found for A- than S-Greebles, *F*s≥5.62, *p*s<.03. Note that the Reversed Pairing group showed numerically higher activity than the Typical Pairing group in several areas (e.g., left LO), but the effect of Pairing only approached significance for the transfer-2 objects, *F*=3.05, *p*=.095 and not statistically significant even when data for the trained, transfer-1 and transfer-2 objects was collapsed, *F*=2.48, *p*=.13. No other significant results were found in any of these visual areas (Figure 23).

Activity for the transfer-1 and transfer-2 objects in the semantic ROIs

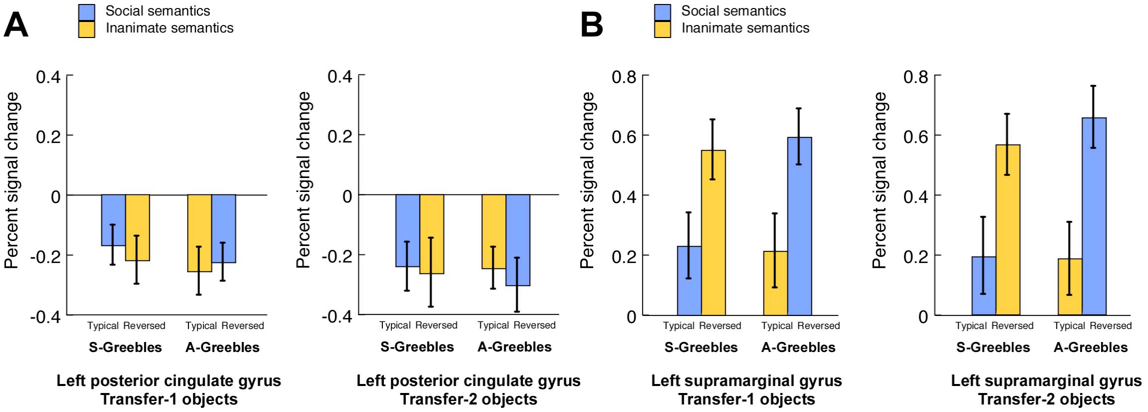


Figure 24. Mean percent signal change for the transfer-1 and transfer-2 objects in the semantic ROIs during the post-semantic training scan. *Figure 24A*: Left posterior cingulate gyrus. *Figure 24B*: Left supramarginal gyrus. Error bars show standard error of the mean.

No results were significant in the left posterior CG for the transfer objects (Figure 24A), *F*s(1,20)≤3, *p*s>.095. But critically, similar to the results with the trained objects, the Reversed Pairing group showed higher activity than the Typical

Pairing group for both transfer-1 and transfer-2 objects in the left SMG (Figure 24B), $Fs(1,18) \geq 5.43$, $ps \leq .03$, suggesting that the pairing effect for the trained objects also generalized to both transfer-1 and transfer-2 objects. No other significant results were found in this region, $Fs(1,18) < 2.23$, $ps \geq .15$.

<u>Summary of the post-semantic training scan results</u>

The whole brain analyses first replicated the main phenomenon of prior semantic training studies (James & Gauthier, 2003; 2004), with areas in the semantic processing network showing more activity for objects explicitly associated with semantic information than objects that were not seen before. Differences between social vs. inanimate semantic features and between A-Greebles and S-Greebles were also revealed in the whole brain analyses. More importantly, an effect of Pairing was also found, with widespread areas more strongly engaged by objects in the Reversed Pairing group than the Typical Pairing group, and a few smaller areas more active for the Typical Pairing group than the Reversed Pairing group. It is important to note that there was a great deal of generalization of the Pairing effects to the transfer-2 objects.

The ROI analyses also revealed a generalized Pairing effect in the left supramarginal gyrus, which was found for the trained, transfer-1 and transfer-2 objects. Moreover, an interaction between Pairing and Visual appearance was also observed in several other visual or semantic ROIs, with the interaction only observed for the trained objects. Specifically, this result was found in the left medial fusiform gyrus (the PPA and the FG) and the posterior cingulate gyrus showing a larger Pairing effect for S-Greebles than A-Greebles, with lower activity found for S-Greebles associated with social semantics than those associated with inanimate semantics. Interestingly, the right FFA also showed a similar trend.

To sum up, these results have revealed two kinds of visual-semantic pairing effects. The first is a main effect of Pairing (found regardless of visual properties of objects) that also applied to objects that were not explicitly associated with semantics (transfer-1 objects) and to new objects that were different from trained objects on several salient visual features (transfer-2 objects). The second is an interaction of Pairing and Visual appearance, which shows a difference in activity for S-Greebles but less so for A-Greebles. Note that this effect was specific to the trained objects and did not generalize to transfer-1 or transfer-2 objects.

Post-visual training scan results

I now turn to the results in the post-visual training scan, to consider the fate of the effects that resulted from the semantic training following the second phase of the experiment that consisted of training to individuate A- and S-Greebles. Half of the objects during this visual training were new and never received semantic associations and no semantic associations were ever mentioned during the visual training. Critically, the experience of the two pairing groups was identical in this phase of this experiment, which totaled 6 hours (across 4 sessions).

Note that since the visual appearance effects and the visual-semantic interactions were found for the transfer-2 objects in the first scan, it would be difficult to reveal expertise effects with a comparison between these and the trained objects here. Despite these limitations, the results after visual training remain informative. Would the previously learned visual-semantic associations in the semantic training survive the visual training where they are irrelevant and not practiced? This question was examined in both the whole brain contrasts and in the ROI analyses.

Whole brain contrasts

Effects of Visual Appearance

An effect of visual appearance persisted after the visual training for the trained (Figure 25A) and transfer-2 objects (Figure 25B). As might be expected, several of these areas overlap with those found in the first scan for the same contrast, with A-Greebles showing widespread activity in the medial fusiform gyrus, ventral and dorsal areas and in the parietal lobe [1, 2, 3, 4], and S-Greebles showing more activity in lingal gyrus [7].

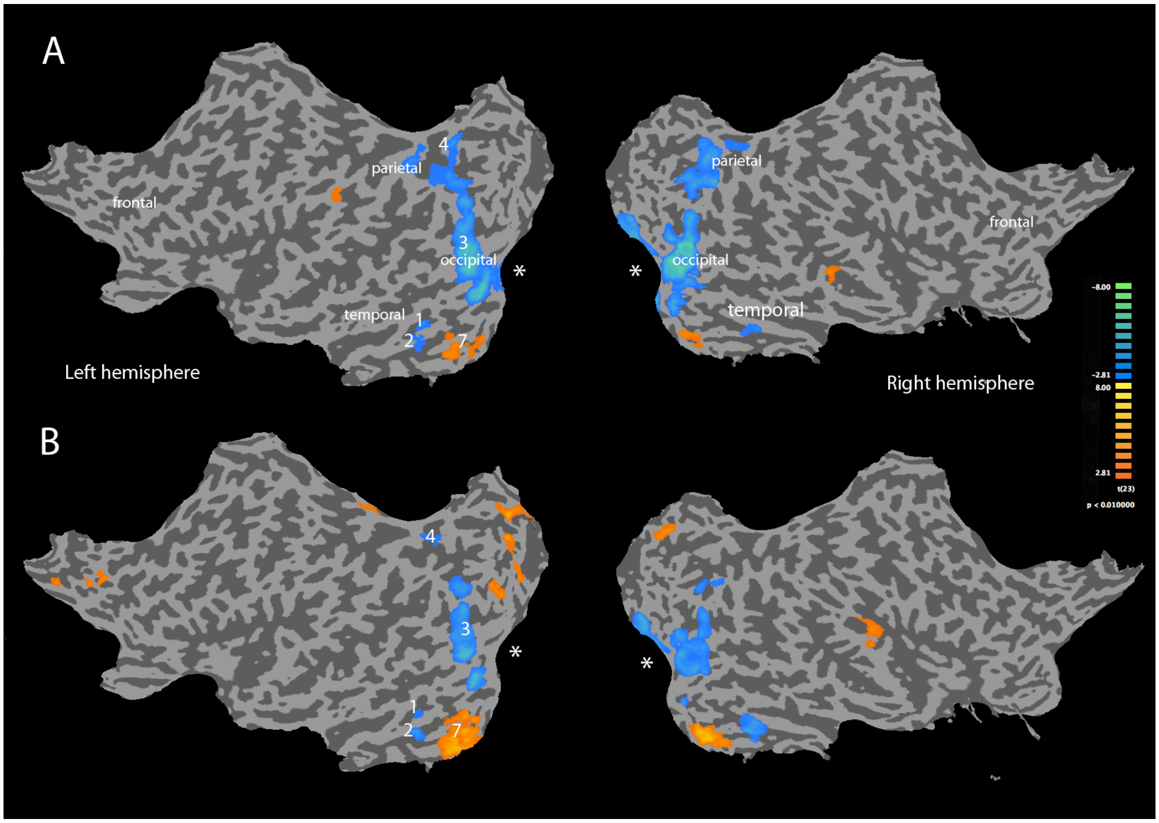


Figure 25. Differential effects for A-Greebles and S-Greebles in the post-visual training scan, presented on the two flattened hemispheres of a representative participant. Blue clusters represent higher selectivity for A-Greebles than S-Greebles; orange clusters represent higher selectivity for S-Greebles than A-Greebles. *Figure 25A*: An averaged map for the trained objects. 1: medial fusiform gyrus (-30, -55, -11), 2: hippocampal gyrus (-26, -42, -8); 3: ventral and dorsal occipitial areas (-31, -80, 9), 4: superior parietal lobe (-18, -42, -66); 7: lingual gyrus (-6, -62, -4). *Figure 25B*: An averaged map for the transfer-2 objects. The contrast was thresholded at $p<.01$ (uncorrected) after which minimum cluster size with probability $p<.01$ (corrected) was calculated using a bootstrapping procedure.

Interactions between visual appearance and semantic training

Although semantic associations were not emphasized during the visual training, some differences between the two training groups still remained in several areas, indicating that the pairings learned in the beginning of the study can still influence visual processing. For the trained objects, the Reversed Pairing group showed higher activity in several posterior visual areas [4, 5, 7], in the right precentral gyrus [10] and the right inferior frontal gyrus [11], whereas the Typical

Pairing group showed higher activity in the left superior temporal sulcus [13], left postcentral gyrus [14], right precentral gyrus [15], and right superior parietal lobe [16]. Fewer areas were found for the transfer-2 objects, but all were a subset of those found for the trained objects, including the left posterior fusiform area [5], right precentral gyrus [10] and right inferior frontal gyrus [11], indicating generalization of the pairing effects.

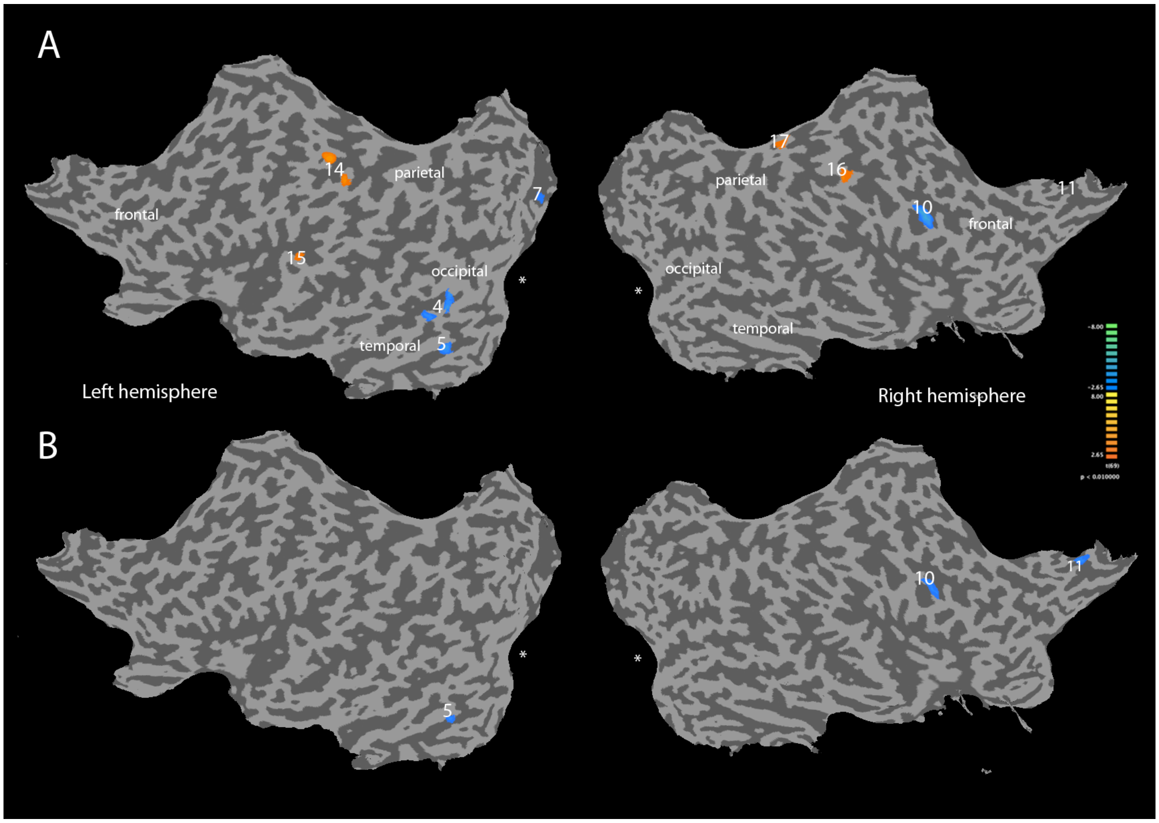


Figure 26. Differential activations in the Typical vs. Reversed Pairing groups in the post-visual training scan, presented on the two flattened hemispheres of a representative participant. Blue clusters represent higher selectivity in the Reversed Pairing than Typical Pairing groups; orange clusters represent higher selectivity in the Typical Pairing than the Reversed Pairing groups. *Figure 26A*: An averaged map for the trained objects. 4: inferior temporal gyrus (-37, -58, -2); 5: posterior fusiform area (-23, -61, -9); 7: striate area (-13, -76, 6); 10: right precentral gyrus (39, 4, 24); 11: right inferior frontal gyrus (39, 4, 22); 12: right lingual gyrus (14, -73, -15); 13: right superior parietal lobe (19, -55, 48); 14: left postcentral gyrus (-42, -34, 42; -40, -31, 46); 15: left superior temporal sulcus (-46, -27, 11); 16: right precentral gyrus (40, -16, 53); 17: right superior parietal lobe (23, -44, 60). *Figure 26B*: An averaged map for the transfer-2 objects. The contrast was thresholded at $p<.01$

(uncorrected) after which minimum cluster size with probability *p*<.01 (corrected) was calculated using a bootstrapping procedure. The degrees of freedom in this contrast was calculated by t(total number of participants-1 × total number of conditions in each run-1) = t(24-1 × 4-1) = t(69).

ROI analyses

Next, I present results from the post-visual training scan in the same ROIs explored after the semantic training. The results are again reported separately for the trained, transfer-1 and transfer-2 objects. To anticipate the results, all semantic effects were reduced in this scan compared to the post-semantic training scan.

Activity for the trained objects in the visual ROIs

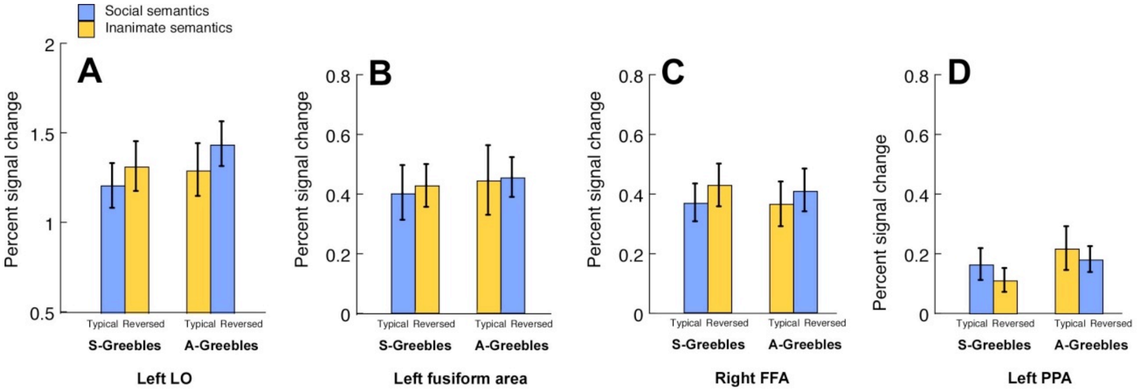


Figure 27. Mean percent signal change for the trained objects in the visual ROIs during the post-visual training scan. *Figure 27A*: Left LO. *Figure 27B*: Left fusiform area. *Figure 27C*: Right FFA. *Figure 27D*: Left PPA. Error bars show standard error of the mean.

Left lateral occipital complex (objects > phased-scrambled objects) (LO)

The shape-sensitive left LO again showed higher activity for A- than S-Greebles after the visual training (Figure 27A), *F*(1,21)=15.92, *p*=.0007. No other significant results were observed, *F*(1,21)≤.46, *p*≥.5.

Left medial fusiform gyrus (objects > phased-scrambled objects) (FG)

Although this area was sensitive to visual-semantic training effects in the post-semantic training scan, there was no effect of Pairing or interaction between Pairing and Visual appearance after the visual training (Figure 27B), $Fs(1,21) \leq .12$, $ps \geq .72$, and no significant effect of Visual appearance: $F(1,21)=2.72$, $p>.11$.

Right fusiform face area (faces > objects and scenes) (FFA)

No significant results were found in the right FFA after the visual training (Figure 27C), $Fs(1,20) \leq .3$, $ps \geq .6$.

Left parahippocampal place area (objects and scenes > faces) (PPA)

As in the post-semantic training scan, A-Greebles again revealed higher activity relative to S-Greebles in the left PPA (Figure 27D), $F(1,22)=12.66$, $p<.002$. However, the interaction between Pairing and Visual appearance were no longer significant, $F(1,22)<.17$, $p>.68$ and no Pairing effect was observed, $F(1,22)<.37$, $p>.55$.

Activity for the trained objects in the semantic ROIs

The left posterior cingulate gyrus showed no significant results, $Fs(1,20) \leq .26$, $ps>.61$ after the visual training (Figure 28A). In the supramarginal gyrus, the main effect of Pairing approached significance (Figure 28B), $F(1,18) \leq 2.54$, $p \geq .13$, with the Reversed Pairing group showed numerically higher activity than the Typical Pairing group. There were no other significant results, $Fs(1,18)<1$, $ps>.78$.

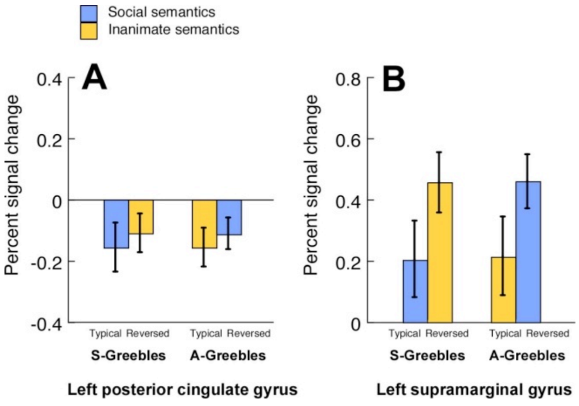


Figure 28. Mean percent signal change for the trained objects in the semantic ROIs during the post-visual training scan. *Figure 28A*: Left posterior cingulate; *Figure 28B*: Left supramarginal gyrus. Error bars show standard error of the mean.

Activity for the transfer-1 and transfer-2 objects in the visual ROIs

Left lateral occipital complex (objects > phased-scrambled objects) (LO)

In the shape-sensitive left LO, the transfer-1 A-Greebles showed higher activity than the transfer-1 S-Greebles (Figure 29A, left panel), $F(1,21)=10.89$, $p=.003$. A similar trend was also found for the transfer-2 objects (Figure 29A, right panel), $F(1,21)=3.23$, $p=.087$. No other results were significant, $Fs(1,21)\leq 2.02$, $p\geq .17$.

Left medial fusiform gyrus (objects > phased-scrambled objects) (FG)

There were no significant results in this region for the transfer objects (Figure 29B), $Fs(1,21)<3.32$, $ps>.08$.

Right fusiform face area (faces > objects and scenes) (FFA)

Interestingly, unlike in the first scan and unlike the effect for trained objects after the visual training, higher activity was found for S- than A-Greebles in the right

FFA (Figure 29C), where the effect was found for both transfer-1 and transfer-2 objects, *F*s(1,20)≥6.22, *p*s≤.02. No other results were significant, *F*s(1,20)≤.67, *p*s≥.42.

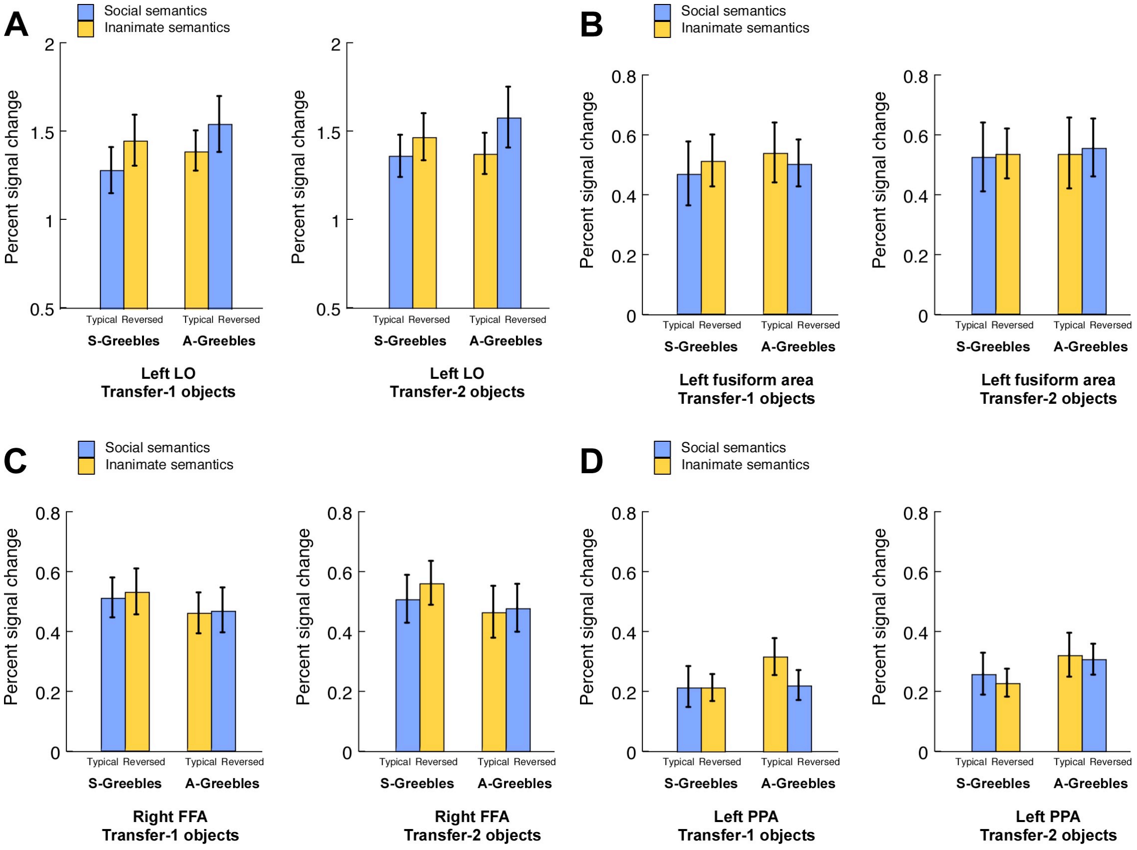


Figure 29. Mean percent signal change for the transfer-1 and transfer-2 objects in the visual ROIs during the post-visual training scan. *Figure 29A*: Left LO. *Figure 29B*: Left FG. *Figure 29C*: Right FFA. *Figure 29D*: Left PPA. Error bars show standard error of the mean.

Left parahippocampal place area (objects and scenes > faces) (PPA)

A-Greebles from both the transfer-1 and transfer-2 sets revealed higher activity relative to S-Greebles in the left PPA, *F*s(1,22)=9.43, *p*≤.006. Also, there was an interaction between Pairing and Visual appearance for the transfer-1 objects, *F*(1,22)=7, *p*=.015, with higher activity for A-Greebles associated with inanimate semantics than other Greeble types (*ps*<.001). Note that this interaction showed a different pattern from that observed in the post-semantic training scan for the trained objects, which did not show an effect for A-Greebles while there was an effect of Pairing for S-Greebles. This was confirmed by an interaction between Session, Pairing and Visual appearance in a 3-way ANOVA, *F*(1,22)=4.61, *p*=.043.

Activity for the transfer-1 and transfer-2 objects in the semantic ROIs

As it was the case for the trained objects, no significant result was observed after visual training for the transfer objects in these semantic processing areas (Figure 30), although the main effect of Pairing approached significance in the left supramarginal gyrus for both the transfer-1 and transfer-2 objects (Figure 30B), *Fs*≤2.93, *ps*≥.1.

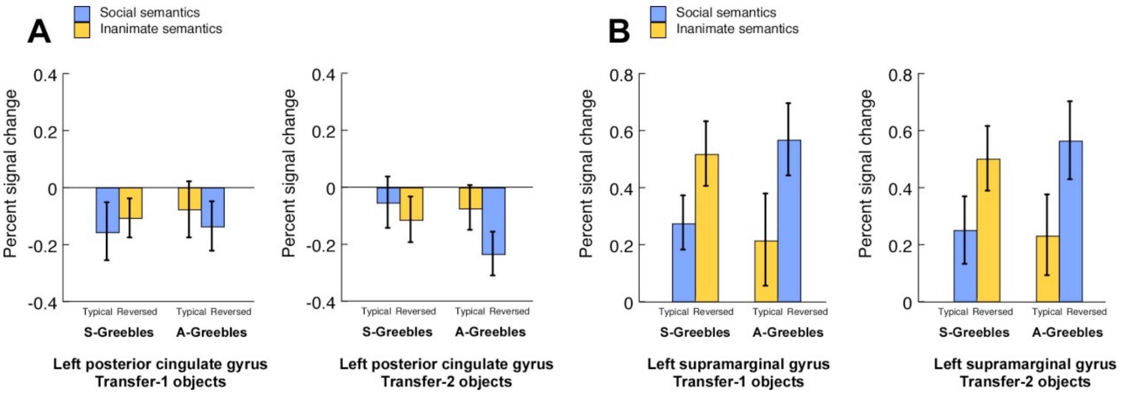


Figure 30. Mean percent signal change for the transfer-1 and transfer-2 objects in the semantic ROIs during the post-visual training scan. *Figure 30A*: Left posterior cingulate gyrus; *Figure 30B*: Left supramarginal gyrus. Error bars show standard error of the mean.

Summary for the post-visual training scan results

Since semantic associations were never mentioned in the visual training, it might not be surprising that the effects of visual-semantic pairing were reduced in the second scan compared to the first scan. Indeed, although a Pairing effect was observed for the trained objects in the medial fusiform area and in the left cingulate gyrus after the semantic training, there was no evidence of this effect in any of the ROIs after the visual training. However, it is nonetheless striking that effects of visual-semantic pairing still remained in several brain areas as revealed in the whole brain contrasts, suggesting that the visual-semantic associations learned during the semantic training continued to influence neural object representations. In particular, the whole brain contrasts revealed a Pairing effect that was observed consistently in three areas across both scans and for all object sub-categories. These areas include the inferior frontal gyrus, posterior fusiform gyrus and right postcentral gyrus, which all revealed higher activity for the Reversed Pairing group than the Typical Pairing group.

While the topography of differential activations in the ventral and dorsal visual system for S-Greebles and A-Greebles was stable across the two scans, it is interesting to note that the right FFA showed a preference for the animal-like S-Greebles compared to the tool-like A-Greebles only after the visual training. However, this preference in the FFA was only found for the transfer objects, not for the extensively trained objects.

General Discussion

These results provide strong evidence for interactions between visual and conceptual object properties in the brain, including the visual system. Clearly, learning new associations between visual and conceptual features can have lasting effects in the neural representations of objects. Recall that the two training groups learned the same objects and semantic features and underwent identical semantic and visual training procedures. The two groups only differed in learning the pairing between social vs. inanimate categories of objects and words. Nonetheless, vast differences in the activations for the Typical and Reversed Pairing groups were observed for both trained and transfer objects after the semantic training. Critically, some of these effects were still observed even after the visual training. Of particular interest is that some of these are visual areas (e.g., posterior fusiform area, striate area), providing evidence that visual responses we measure with fMRI are sensitive to conceptual associations we have with objects and even to the history of conceptual associations with visually similar objects. In that sense, these results demonstrate that even with novel objects and using a visual task, we may never be able to study object recognition outside the context of conceptual associations evoked by the objects. I will discuss this at more length in Chapter 4.

Note also that generalization of the Pairing effect to the transfer-2 objects was not predicted but is quite informative. On the one hand, participants learned unique semantics for each object in the study. But on the other hand, the visual features fell into clear social vs. inanimate categories that were likely salient to them, as suggested by the neural Pairing effects. The transfer-2 objects differed from the other objects in salient visual features (color and body shape), but generalization of the semantic associations may be driven by one or both of these aspects of visual information: the configuration of their parts (symmetric or asymmetric) or the shape of the parts themselves (smooth pointing down for S-Greebles, sharp and pointing up for A-Greebles). These two factors could be distinguished in future work.

CHAPTER IV

CONCLUDING REMARKS

Implications for existing theories of object recognition

In this dissertation I asked whether and how interactions between visual and conceptual properties constrain object recognition. As discussed in Chapter 1, several influential theories in the object recognition literature build on the assumption that object recognition depends almost entirely on visual attributes of objects (e.g., Marr, 1982; Biederman, 1987; Perrett & Oram, 1993; Riesenhuber & Poggio, 1999; Jiang et al., 2006) and that semantic associations should have no influence on object recognition (e.g., Pylyshyn, 1999; but see Goldstone & Baraslou, 1998). Additionally, researchers who are interested in the role of shape in object processing often use novel objects or shapes to prevent influences from non-visual information, such as object names, familiarity and conceptual content (e.g., Op de Beeck et al., 2008). Here I gathered behavioral and neural evidence that seriously questions these assumptions. In Chapter 2, I showed that the shape of novel objects can influence conceptual learning and thus visual object properties can implicitly convey meaning. I designed this study in part based on an intuition, which was that symmetric, animal-like novel objects would be more easily associated with social semantics (because of their resemblance to faces) and that asymmetric, tool-like novel objects would be more easily associated with inanimate semantics (because of their resemblance to man-made objects or tools). This intuition was confirmed by the results obtained in the lexical judgment task, which was just one of the many examples from this study to suggest that meaning is implicitly evoked by visual appearance. This makes sense if concepts are rooted in perception and action (Barsalou, 1999; 2008): sensory and motor systems can be engaged by the presentation of words that evoke action (e.g., Buccino et al., 2005); these systems can also be engaged by the presentation of objects associated with these actions through experience (James & Gauthier, 2006; Wong & Gauthier, 2010), but also by the presentation of novel objects that have a history of association with such words, even if those associations are arbitrary (James & Gauthier, 2003; the semantic effects in this dissertation). In this study, the fact that even relatively unexpected semantic associations (social words with tool-like objects) generalized to objects that shared only some of the properties of the trained objects suggests a mechanism to explain the implicit effect observed for the novel objects prior to any training. In other words, the default meaning of these objects may simply be generalized on the basis of visual similarity with familiar object categories. If relatively novel and arbitrary associations that run contrary to much of our experience can generalize in this manner, a lifetime's history of conceptual learning is likely a very powerful influence in how we represent any new object that we encounter.

Some authors have emphasized the distinction between common and novel objects to test theories of object recognition. In one case, Curby, Hayward, and Gauthier (2004) examined how semantic associations modulate viewpoint invariant vs. viewpoint dependent effects (see e.g., Biederman, 1987; Tarr; 1995). They pointed out that studies supporting different representations in the two hemispheres had mainly used common objects, which likely carry a lot of semantic associations and that could have a differential effect within each hemisphere. In that study, one group of participants learned semantic associations with novel objects ("Yufos", see Figure 1), while the other group did not receive semantic training. Curby et al. found that for the untrained group, there was no difference in viewpoint dependent performance between hemispheres. However, for the trained group, semantic associations led to a reduced effect of viewpoint only when objects were presented to the left hemisphere. The authors interpreted this result as evidence that semantic associations can influence visual performance and that common objects may not be suited to test perceptual theories. However, based on my findings, I can question this assumption that novel objects do *not* to some degree implicitly engage semantic processing even upon their first presentation.

While semantics with novel objects is unlikely to be eliminated, there is still an advantage in using novel objects over common objects to study the effect of shape or semantics because these effects can be manipulated, as illustrated in the current study. Although visual properties (e.g., symmetry, curvature, color) appear to elicit semantic associations on their own, the current study is consistent with previous work showing that semantics can be relatively quickly attached to novel objects through explicit associations. Such semantic associations influenced performance in a visual matching task. For instance, objects attached to inanimate semantics led to a larger reduction of the basic-level advantage compared to objects attached with social semantics. It is unclear why inanimate features produce larger effects than social semantics in this task, but this result demonstrates that object representations can be affected by explicit semantic learning (Dixon et al., 1997; Gauthier et al., 2003a). One possible explanation for this difference between inanimate vs. social semantics is that inanimate concepts typically possess lower featural overlap than social concepts (Mechelli et al., 2006). Indeed, in this study, I did not have a way to compare the relative average distance between the various social vs. inanimate features. Furthermore, the size of each conceptual space depends on the similarity among the groupings of three features that are associated with each object. Two objects that are "elastic, shiny and antique" vs. "eco-friendly, plastic and durable" may seem to be quite different in function and likely to belong to different basic-level categories. In contrast, two objects that are "adorable, funny and sensitive" and "cheerful, talented and forgiving" are more likely two exemplars in the same basic-level category. Inanimate associations may thus result in concepts that are more dissimilar than social associations, a factor that can facilitate visual discrimination (Gauthier et al., 2003a).

One novel manipulation in this dissertation was whether the semantic associations learned by participants were those that were typical for objects of a certain shape, as opposed to associations that were relatively unexpected. What was gained by manipulating the typicality of the pairings between shape and semantics? First, reversed pairings of objects and semantics can be learned, and the associations appear to be retrieved at least as automatically as typical pairings during a visual task. More importantly, reversed pairings showed higher activity than typical pairings in several brain areas, in some cases without any regards for the visual object features or semantics (e.g., the supramarginal gyrus). It is tempting to speculate that there are regions of the brain that may be engaged when any new kind of associations must be learned, particularly when the association is not easily assimilated within our existing conceptual networks, such as the when asymmetry predicts that an object belongs in an animate category. Such an abstract function evokes the concept of “convergence zones”, areas that are thought to bind information distributed in the brain to form any given concept, but which by themselves do not play any representational role (Damasio, 1989; Damasio & Damasio, 1994).

I have shown that on top of the influence of semantic category, a mismatch in the expected and actual object-semantic associations (e.g., asymmetric objects with social semantics) can also affect object recognition performance. Specifically, the two training groups learned and perceived identical objects sets and semantic features, with the only difference between the two groups being the object-word pairing during the semantic training. According to the shape-based object recognition theories, object recognition should not be influenced by differences in semantics or in object-semantic pairings. But contrary to this prediction, behavioral and neural differences were observed between the two groups here in various visual judgment tasks where semantic features were irrelevant (e.g., sequential and simultaneous matching). Thus, my findings call for a revision of these theories, as a pure shape-based theory of object recognition cannot account for these results.

A note on perceptual expertise

This study also demonstrates that the acquisition of perceptual expertise can be modulated by visual object features, explicit conceptual knowledge and the pairing of visual-conceptual properties. Specifically, the reduction of the basic-level advantage was largest for symmetric objects associated with social semantics. Holistic processing also emerged for this type of object category after the semantic training, although this holistic effect was diminished following visual training.

It has been postulated by some that the FFA, a locus of perceptual expertise, should show higher activity for a combination of a face-like shape and social semantics (e.g., a smooth-edged object that "looks like a woman wearing a hat"; Op de Beeck et al., 2006). However, my results instead show a trend for the opposite effect following semantic training, with lower activity for symmetric objects associated with social semantics than in the other combinations. Note also that higher activity was observed for symmetric compared to asymmetric objects after visual individuation training, although this effect was only observed for transfer but not trained objects. The interactions between visual appearance, conceptual knowledge and perceptual experience could be further examined in future studies.

Remaining questions and future directions

Some might argue that the effects observed in this study were not due to semantic processing *per se*. Specifically, it is possible that the inanimate semantic category facilitated encoding of visual features, and that it was the enhanced visual encoding but not the access to the semantics during the visual task that led to improvement after the semantic training. However, this seems unlikely to be the only explanation because neural semantic effects have been observed during a simultaneous task in the fMRI scanner, suggesting that conceptual knowledge was also retrieved during a visual task. However, the current methods cannot determine when and how conceptual knowledge influences visual processing. There are at least two possibilities. First, it may be that conceptual knowledge is represented in the frontal-parietal semantic processing areas. Once this information is retrieved, it feeds backward to the occipital-temporal areas and influences the visual representation of an object. Second, visual and conceptual properties of objects may be both stored and represented in the visual areas, and can be accessed at the same time. Future research can distinguish these possibilities and ask how early the interactions between visual and conceptual processing occur.

This question can be addressed using either behavioral or neural measures. Behaviorally, the information available or processed at an early processing stage can be revealed using a *signal-to-response* technique (Corbett & Wickelgren, 1978; Dosher, 1981; Hintzman et al., 1994; Reed, 1973). This technique systematically varies the amount of time a participant is given to process an object and measures how recognition performance changes as a function of processing time. Applying this technique to the lexical judgment task or the basic- vs. subordinate-level categorization task will be useful to reveal how early the semantic effects occur. Apart from this behavioral measure, another approach to examine the time course of the visual-semantic effects is to adopt an event-related design and/or functional connectivity analysis methods in an fMRI study. With an event-related design, the time course of activity in various areas in the visual-semantic processing network can be unraveled. Moreover, functional connectivity analysis methods can be used to evaluate the nature of the connections among these areas (Assaf et al., 2009; Werner & Noppeney, 2010). Specifically, the coherence in fMRI signals among these regions when matching objects with various types semantic associations can be examined. Additionally, recording of event-related potentials (ERPs) has a high temporal resolution and thus will also be a useful technique to study the time course of these effects (Ritter et al., 1983; see also Nyhus & Curran, 2009).

It is also important to further address why certain types of semantic information facilitate object recognition more than others. While the nature of the different semantic categories may cause the effect (e.g., inanimate vs. social semantics), an alternative possibility that needs to be ruled out is the degree of featural overlap within each semantic category (Mechelli et al., 2006). To elucidate the source of semantic effects, future research should attempt to equate or *manipulate* the distinctiveness of features within different semantic categories. Moreover, the distinctiveness of visual features may also be manipulated in the same context. In this study, it is possible that explicit semantic associations enhanced the diagnosticity of visual features because the objects used in each set were homogeneous in shape and configuration of parts. It will be interesting to investigate to what extent object recognition is affected by the distinctiveness of both visual and semantic features (see Brown & Lloyd-Jones, 2006), and to test the claim of a "perceptual-semantic continuum" in object representations (Sergent & Poncet, 1990; Sergent & Signoret, 1992; Young et al., 1989).

The current study can also be compared with the research of verbal descriptions on visual memory. In a seminal study, Schooler and Engstler-Schooler (1990) demonstrated that verbal descriptions of facial features impair subsequent face recognition. While such "verbal shadowing" effect is a result of describing visual appearance of a face, adding non-visual information (e.g., global verbal descriptions such as personality trait or occupation information) instead facilitate later face recognition in other cases (Bower & Karlin, 1974; Kerr & Winograd, 1982; McKelvie, 1985; Dixon et al., 1997; Brown & Lloyd-Jones, 2005). The verbal facilitation on recognition memory has been attributed to the *level-of-processing* account (Bower & Karlin, 1974), in which verbal associations may enhance the number of features attended and stored during encoding (e.g., Winograd, 1981), improve global and featural encoding of visual information (e.g., Wells & Hryciw, 1984), and/or form richer semantic associations with a face that benefit retrieval (e.g., Bruce & Young, 1986). Similarly, the semantics associations added to the objects in this study were words that describe personalities or quality of man-made objects, and I showed that these associations enhanced performance in a perceptual task. It will be important to examine whether the same mechanisms of visual-semantic processing mediate both object perception and memory.

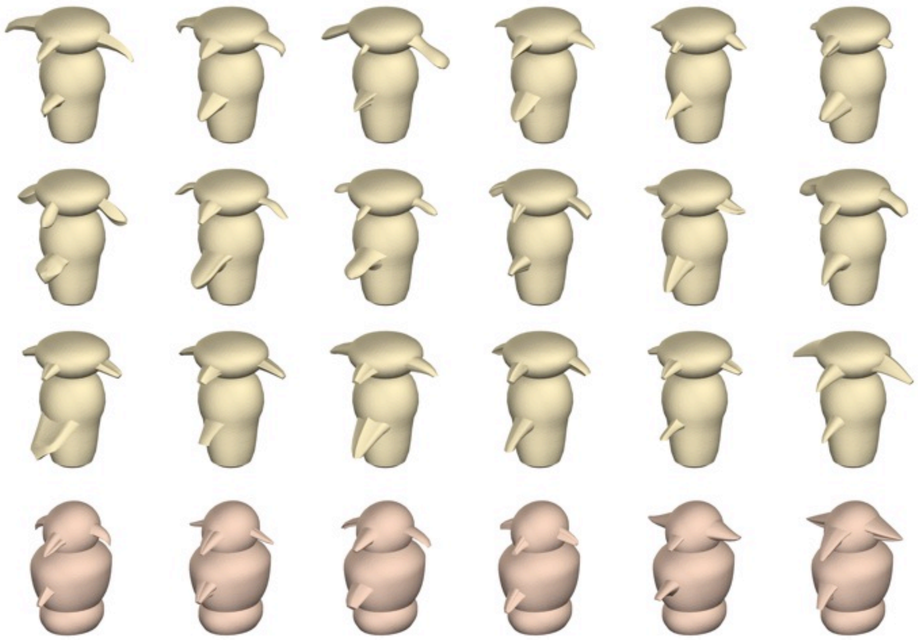
Conclusion

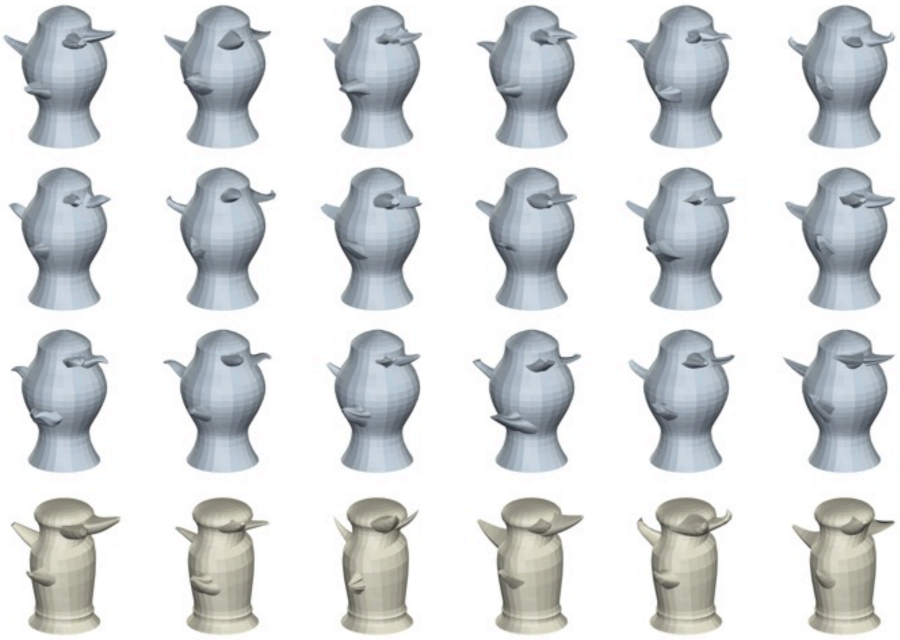
This dissertation reveals the interaction of visual and conceptual properties in object recognition and its impact on object recognition and perceptual expertise. The current work demonstrates the power of manipulating both visual and conceptual factors with artificial objects and novel concepts created out of lists of words. It opens the way for further experimentation and theoretical development with regards to how different types of information interact to determine object percepts and concepts.

APPENDIX A

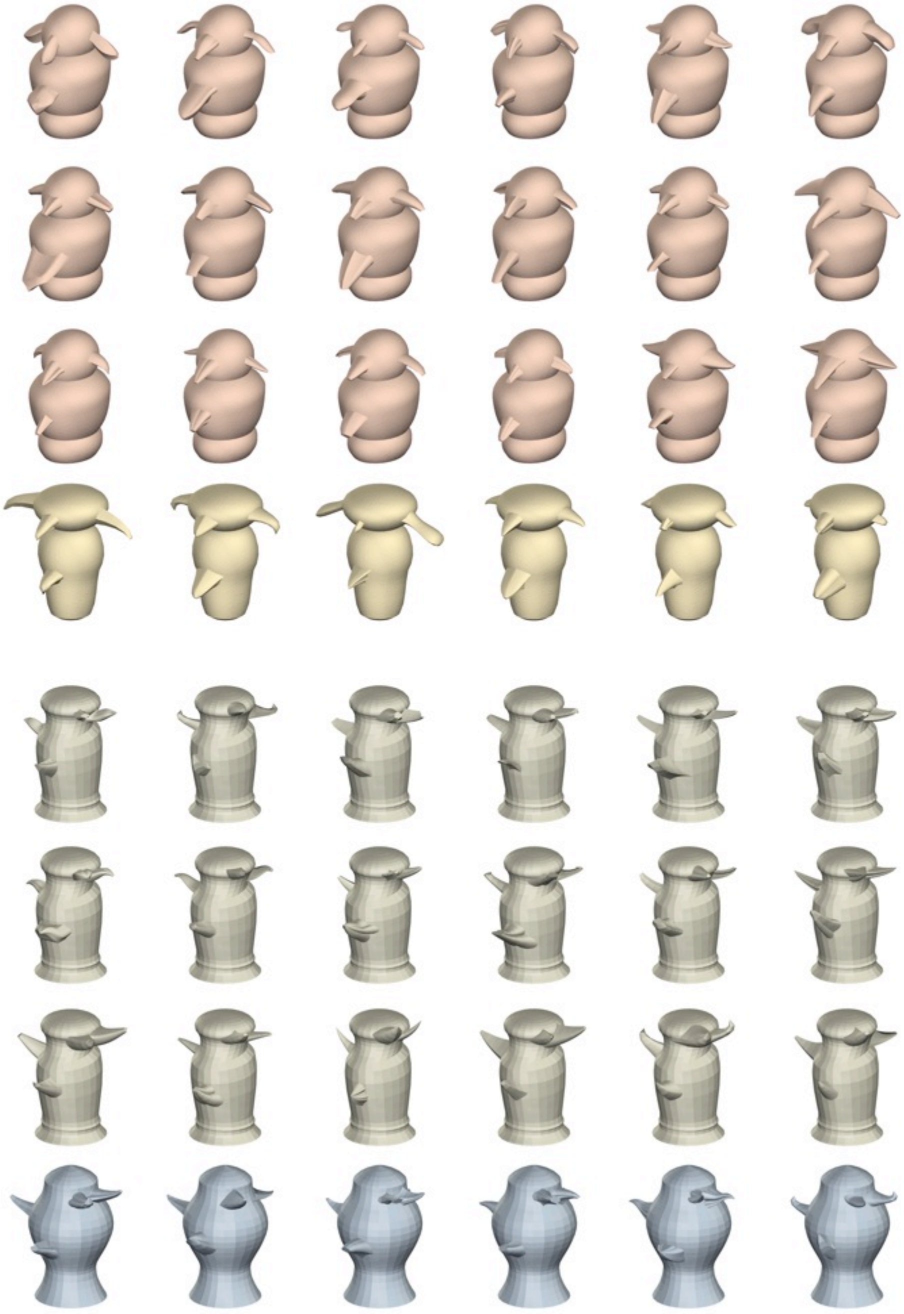
The two versions of Greebles used (presented at 6˚).

Version 1.





Version 2.



APPENDIX B

List of social and inanimate features used.

Social	Social (cont.)	Inanimate	Inanimate (cont.)
Adorable	Hopeful	Absorbent	High-tech
Agreeable	Humorous	Accessible	Imported
Alert	Ignorant	Acidic	Jagged
Aloof	Innocent	Adhesive	Lopsided
Apathetic	Introverted	Affordable	Lumpy
Appreciative	Jealous	Antique	Magnetic
Boastful	Lazy	Assembled	Matte
Bored	Lonely	Bent	Metallic
Bossy	Mature	Bouncy	Multipurpose
Brave	Motherly	Boxy	Plastic
Cheerful	Nervous	Breakable	Pointy
Clumsy	Pleasant	Bumpy	Portable
Concerned	Poised	Clunky	Prickly
Confused	Polite	Compact	Recyclable
Curious	Proud	Cumbersome	Rubbery
Cute	Rebellious	Curvy	Scratchy
Demanding	Resentful	Decorative	Second-hand
Embarrassed	Sad	Drab	Shiny
Energetic	Selfish	Durable	Slippery
Envious	Sensitive	Eco-friendly	Solid
Excited	Shy	Elastic	Sparkling
Extroverted	Spontaneous	Electrical	Spotless
Forgiving	Stupid	Expensive	Stainless
Friendly	Sweet	Flimsy	Stretchy
Frightened	Sympathetic	Functional	Synthetic
Funny	Talented	Geometric	Textured
Gifted	Thankful	Glossy	Twisted
Grateful	Thoughtful	Hand-held	Uneven
Happy	Warm	Hard-to-find	Useful
Helpful	Worried	Hardwearing	Well-made

REFERENCES

Assaf, M., Jagannathan, K., Calhoun, V., Kraut, M., Hart, J. Jr., & Pearlson, G. (2009). Temporal sequence of hemispheric network activation during semantic processing: A functional network connectivity analysis. *Brain and Cognition, 70*(2), 238-46.

Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J. & Iacoboni, M. (2008). Lateralizationof the human mirror neuron system. *Journal of Cognitive Neuroscience*, 26, 2964-70.

Baker, C.I., Liu, J. Wald, L.L., Kwong, K.K., Benner, T., Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the Nation Academy of Science, 104*(21), 9087-92.

Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science, 17*(8), 645-8.

Bar, M., & Neta, M. (2007). Visual elements of subjective preference modulate amygdala activation. *Neuropsychologia , 45*(10), 2191-200.

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577-609.

Behrmann, M., Marotta, J., Gauthier, I., Tarr, M. J., & McKeeff, T. J. (2005). Behavioral change and its neural correlates in visual agnosia after expertise training. *Journal of Cognitive Neuroscience, 17*, 554-68.

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: the influence of perceptual experience on early visual processing of faces. *Cognition , 86*(1), B1-14.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115–47.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for 3D viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 1162-82.

Booth, A. E., & Waxman, S. R. (2002). Word learning is ‘smart’: Evidence that conceptual information affects preschoolers’ extension of novel words. *Cognition, 84*(1), B11-B22.

Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology, 103,* 751-7.

Boynton, G.M., Engel, S.A., Glover, G.H., Heeger, D.J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *Journal of Neuroscience, 16*, 4207–21.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-6.

Brown, C., & Lloyd-Jones, T. J. (2005). Verbal facilitation of face recognition. *Memory & Cognition, 33*(8), 1442-56.

Brown, C., & Lloyd-Jones, T. J. (2006). Beneficial effects of verbalization and visual distinctiveness on remembering and knowing faces. *Memory & Cognition, 34*(2), 277-86.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305-27.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977-90.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V. & Rizzolattie, G. (2005). Listening to action-related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*, 355-363.

Bukach, C. M., Phillips, W. S. & Gauthier, I. (in press). Limits of generalization beteen categories and implications for theories of category specificity. *Psychonomic Bulletin & Review.*

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Science*, *10*, 159-66.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science*, *89*, 60-4.

Caramazza, A., & Mahon, B. Z. (2006). The organization of conceptual knowledge in the brain: the future's past & some future directions. *Cognitive Neuropsychology*, *23*, 13-38.

Carey, S. (2000). The origin of concepts. *Journal of Cognition and Development, 1*(1)*,* 37-41.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747–9.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience, 2*, 913-9.

Chen, C., Kao, K., & Tyler, C. (2007). Face configuration processing in the human brain: the role of symmetry. *Cerebral Cortex, 17*(6), 1423-32.

Cheung, O. S., & Gauthier, I. (2010). Selective interference on the holistic processing of faces in working memory. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 448-61.

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance, 34*(6), 1327-36.

Concar, D. (1995). Sex and the symmetrical body. *New Scientist, 146,* 40-4.

Corbett, A. T., & Wickelgren, W. A. (1978). Semantic memory retrieval: Analysis by speed accuracy tradeoff functions. *Quarterly Journal of Experimental Psychology, 30*(1), 1-15.

Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*(5667), 115-7.

Curby, K. M., Hayward, W. G., & Gauthier, I. (2004). Laterality effects in the recognition of depth rotated novel objects. *Cognitive, Affective & Behavioral Neuroscience*, *4*(1), 100-11.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition*, *33*, 25-62.

Damasio, A. R., & Damasio, H. (1995). Cortical systems for retrieval of concrete knowledge: The convergence zone framework. In C. Koch, & J. L. Davis (Eds.), *Large-scale neuronal theories of the brain.* The MIT Press.

de Fockert, J., Rees, G., Frith, C., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*(5509), 1803-6.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107-17.

Dixon, M., Bub, D., & Arguin, M. (1997). The interaction of object form and object meaning in the identification performance of a patient with category-specific visual agnosia. *Cognitive Neuropsychology, 14*(8), 1085-130.

Dixon, M., Bub, D., & Arguin, M. (1998). Semantic and visual determinants of face recognition in a prosopagnosic patient. *Journal of Cognitive Neuroscience, 10*(3), 362-76.

Dosher, B. A. (1981). The effects of delay and interference: A speed-accuracy study. *Cognitive Psychology, 13*(4), 551-82.

Dux, P., & Coltheart, V. (2005). The meaning of the mask matters: evidence of conceptual interference in the attentional blink. *Psychological Science, 16*(10), 775-9.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598-601.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine. 33*, 636–647.

Freedman D. J., Riesenhuber M., Poggio T., & Miller E. K. (2003). A comparison of primate prefrontal and inferior temporal cortex during visual categorization. *Journal of Neuroscience*, 23, 5235-46.

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Tradeoffs and strategic pluralism. *Behavior and Brain Science, 23,* 573-87.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology, 59*(2), 154-79.

Gauthier, I. (2000). What constrains the organization of the ventral temporal cortex? *Trends in Cognitive Science, 4*, 1-2.

Gauthier, I. & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 37, 1682-82.

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, 431-46.

Gauthier, I., James, T. W., Curby, K. M., & Tarr, M. J. (2003a). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology*, *20*, 507-23.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003b). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*, 428-32.

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*, 191-7.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation in the middle fusiform 'face area' increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568-73.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. W. (1998). Training "Greeble" experts: A framework for studying expert object recognition processes. *Vision Research*, *38*, 2401-28.

Gelman, R., & Spelke, E. (1981). The development of thoughts about animate and inanimate objects: Implications for research on social cognition. In J. H. Flavell, & L. Ross (Eds.). *Social cognitive development: Frontiers and possible futures.* New York: Cambridge University Press.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15*, 870-8.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology, 39,* 1-41.

Gibson, E. J., & Walker, A. S. (1984). Development of knowledge of visual-tactual affordances of substance. *Child Development*, *55*, 453-60.

Gibson, J. J. (1978). The ecological approach to the visual perception of pictures. *Leonardo, 11*(3), 227-35.

Glaser, W. R. (1982). Picture naming. *Cognition, 42,* 61-105.

Goebel, R., Esposito, F., & Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with brainvoyager QX: From single-subject cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Human Brain Mapping, 27*, 392–401.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123,* 178-200.

Goldstone, R. L., & Barsalou, L W. (1998). Reuniting perception and conception. *Cognition, 65*(2-3), 231-62.

Goldstone, R. L. & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, *130*, 116-39.

Graham, S. A., Poulin-Dubois, D. (1999). Infants' reliance on shape to generalize novel labels to animate and inanimate objects. *Journal of Child Language, 26,* 295-320.

Grelotti, D. J., Gauthier, I., & Schultz, R. T. (2002). Social interest and the development of cortical face specialization: what autism teaches us about face processing. *Developmental Psychobiology , 40*(3), 213-25.

Grelotti, D. J., Klin, A. J., Gauthier, I., Skudlarski, P., Cohen, D. J., Gore, J. C., Volkmar, F. R., & Schultz, R. T. (2005). FMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia, 43*, 373-85.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within category identification. *Nature Neuroscience,* 7, 555- 562.

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience*, *3*(8), 837-43.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the occipital lobe. *Human Brain Mapping*, *6*(4), 316-28.

Grill-Spector, K., Sayres, R., & Ress, D. (2006). High-resolution imaging reveals highly selective nonface clusters in the fusiform face area. *Nature Neuroscience, 9*(9), 1177-85.

Harley, E. M., Pope, W. B., Villablanca, P., Mumford, J., Suh, R., Mazziotta, J. C., Enzmann, D. & Engel, S. A. (2009). Engagement of fusiform cortex and disengagement of lateral occipital cortex in the acquisition of radiological expertise. *Cerebral Cortex, 19,* 2746-54.

Hilliar, K., & Kemp, R. (2008). Barack Obama or Barry Dunham? The appearance of multiracial faces is affected by the names assigned to them. *Perception, 37*(10), 1605-8.

Hintzman, D. L., Caulton, D. A., & Curran, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 275-89.

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), 453-76.

James, T. W., & Gauthier, I. (2003). Auditory and action semantic features activate sensory-specific perceptual brain regions. *Current Biology, 13*(20), 1792-6.

James, T. W., & Gauthier, I. (2004). Brain areas engaged during visual judgments by involuntary access to novel semantic information. *Vision Research*, *44*(5), 429-39.

James, K. H., & Gauthier, I. (2006). Letter processing automatically recruits a multimodal brain network. *Neuropsychologia*, *44*(14), 2937-49.

Jenkins, R., Lavie, N., & Driver, J. (2003). Ignoring famous faces: category-specific dilution of distractor interference. *Perception & Psychophysics*, *65*(2), 298-309.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., Vanmeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron, 53*, 891-903.

Jolicoeur, P., Gluck, M., & Kosslyn, S. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*(2), 243-75.

Jones, B. C., Little, A. C., Feinberg, D. R., Penton-Voak, I. S., Tiddeman, B. P., & Perrett, I. (2004). The relationship between shape symmetry and perceived skin condition in male facial attractiveness. *Evolution and Human Behavior, 25,* 24-30.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8,* 113-39.

Jones, S. S., & Smith, L. B. (1998). How children name objects with shoes. *Child Development, 13*(3), 323-34.

Jones, S. S., Smith, L. B., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*(3), 499-516.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759-63.

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*(11), 4302-11.

Keil, F. C. (1991). The emergence of theoretical beliefs as constraints on concepts. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and congition.* New Jersey: Lawrence Erlbaum Associates.

Kerr, N. H., & Windgrad, E. (1982). Effects of contextual elaboration on face recognition. *Memory & Cognition, 10,* 603-9.

Kim, J. G., Biederman, I., Lescroart, M. D., & Hayworth, K. J. (2009). Adaptation to objects in the lateral occipital complex (LOC): Shape or semantics? *Vision Research*, *49*(18), 2297-305.

Kung, C.-C., Peissig, J.J., & Tarr, M.J. (2007). Is region-of-interest overlap comparison a reliable measure of category specificity? *Journal of Cognitive Neuroscience*, *19*, 2019-34.

Lin, E. L., & Murphy, G. L. (2009). Effects of background knowledge on object categorization and part detection. *Journal of Experimental Psychology: Human Perception and Performance, 23*(4), 1153-69.

Livshits, G., & Kobyliansky, E. (1991). Fluctuating asymmetry as a possible measure of developmental homeostasis in humans: A review. *Human Biology, 63*(4), 441-66.

Lockman, J. J. (2000). A perception-action perspective on tool use development. *Child Development, 71*(1), 137-44.

Lupyan, G., & Spivey, M. J. (2008). Perceptual processing is facilitated by ascribing meaning to novel stimuli. *Current Biology, 18*(10), R410-2.

Mack, M. L., Wong, A. C.-N., Gauthier, I., Tanaka, J. W., & Palmeri, T. J. (2009). Time course of visual object categorization: Fastest does not necessarily mean first. *Vision Research, 49*(15), 1961-8.

Mahon, B., & Caramazza, A. (2009). Concepts and categories: A cognitive neuropsychological perspective. *Annual Review of Psychology*, *60*, 27-51.

Mandler, J. M. (2000). Perceptual and conceptual processes in infancy. *Journal of Cognition and Development, 1*(1), 3-36.

Marr. D. (1982). *Vision: A computation investigation into the human representation and processing of visual information.* San Francisco: Freeman.

Martin, A. & Chao, L.L. (2001). Semantic memory and the brain: structure and processes. *Current Opinion in Neurobiology*, *11*(2), 194-201.

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology, 58,* 25-45.

Massey, C. M., & Gelman, R. (1988). Preschooler's ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology, 24*(3), 307-17.

McCarty, M. E., Clifton, R. K., & Collard, R. R. (2001). The beginnings of tool use by infants and toddlers. *Infancy, 2*(2), 233-56.

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (submitted). Race-specific preceptual discrimination improvement following short individuation training with faces.

McKelvie, S. J. (1985). Effect of depth of processing in recognition memory for normal and inverted photographs of faces. *Perceptual & Motor Skills, 60,* 503-8.

Mechelli, A., Sartori, G., Orlandi, P., & Price, C. J. (2006). Semantic relevance explains category effects in medial fusiform gyri. *NeuroImage, 30*(3), 992-1002.

Nelson, D. G. K., O'Neil, K. A., & Asher, Y. M. (2008). A mututally facilitative relationship between learning names and learning concepts in preschool children: The case of artifaces. *Journal of Cognition and Development, 9*(2), 171-93.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. *Psychological Review, 81*(4), 267-85.

Newcombe, F., Young, A. W., & De Haan, E. H. (1989). Prosopagnosia and object agnosia without covert recognition. *Neuropsychologia, 27*(2), 179-91.

Nyhus, E., & Curran, T. (2009). Semantic and perceptual effects on recognition memory: Evidence from ERP. *Brain Research, 1283,* 102-14.

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience, 26*(50), 13025-36.

Op de Beeck, H., Deutsch, J., Vanduffel, W., Kanwisher, N., & DiCarlo, J. (2008a). A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. *Cerebral Cortex, 18*(7), 1676-94.

Op de Beeck, H. P., Torfs, K., & Wagemans, J. (2008b). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, *28*(40), 10111-23.

Op de Beeck, H. P., Wagemans, J., & Vogels, R. (2003). The effect of category learning on the representation of shape: Dimensions can be biased but not differentiated. *Journal of Experimental Psychology: General, 132*(4)*,* 491-511.

Osiurak, F., Jarry, C., & Le Gall, D. (2010). Grasping the affordances, understanding the reasoning. Towards a dialectical theory of human tool use. *Psychological Review, 117*(2), 517-40.

Palmeri, T. J., & Flanery, M. A. (1999). Learning about categories in the absence of training: Profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychological Science*, *10*(6), 526-30.

Parsons, P. A. (1990). Fluctuating asymmetry: An epigenetic measure of stress. *Biological Reviews, 65*(2), 131-45.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437-42.

Perrett, D. I., Burt, M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior, 20,* 298-307.

Perrett, D. I. and Oram, M. W. (1993) Neurophysiology of shape processing. *Image and Visual Computing,* 11, 317-333.

Polak, M. (2003). *Developmental instability: Causes and consequences.* New York: Oxford University Press.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: A functional magnetic resonance imaging study. *Journal of Neuroscience, 16*(16), 5205-15.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impentrability of visual perception. *Behavioral and Brain Sciences, 22*341-423.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology, 8*(4), 364-82.

Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin & Review, 5*(4), 659-69.

Richler, J. J., Cheung, O. S., Wong, A. C.-N., & Gauthier, I. (2009). Does response interference contribute to face composite effects? *Psychonomic Bulletin & Review, 16*(2), 258-63.

Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008a). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 328-42.

Richler, J. J., Tanaka, J. W., Brown, D. D., & Gauthier, I. (2008b). Why does selective attention to parts fail in face processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34* (6), 1356-68.

Riesenhuber, M. & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019-25.

Riesenhuber, M. & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience*, *3*, 1199-204.

Reed, A. V. (1973). Speed-accuracy tradeoff in recognition memory. *Science, 181,* 574-6.

Rhodes, G., Simmons, L., & Peters, M. (2005). Attractiveness and sexual behavior: Does attractiveness enhance mating success? *Evolution and Human Behavior, 26,* 186-201.

Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., & McKay, R. (2001). Do facial averageness and symmetry signal health? *Evolution and Human Behavior, 22,* 31-46.

Ritter, W., Simson, R., & Vaughan, H. G., Jr. (1983). Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. *Psychophysiology, 20*(2), 168-79.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience, 19*(3), 543-55.

Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 101*(40), 14521-6.

Sasaki, Y., Vanduffel, W., Knutsen, T., Tyler, C., & Tootell, R. (2005). Symmetry activates extrastriate visual cortex in human and nonhuman primates. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 3159-63.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*(8), 299-309.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cogntiive Psychology, 22,* 36-71.

Schultz, R. (2005). Developmental deficits in social perception in autism: The role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23*(2-3), 125-41.

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., & Skudlarski, P. (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, Biological Sciences, 358*(1430), 415-27.

Schweizer, T. A., & Dixon, M. J. (2006). The influence of visual and nonvisual attributes in visual object identification. *Journal of the International Neuropsychological Society, 12*(2), 176-83.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience*, *18*(9), 1453-65.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research*, *1210*, 204-15.

Sergent, J., & Poncet, M. (1990). From covert to overt recognition of faces in a prosopagnosic patient. *Brain*, *113*, 989-1004.

Sergent, J., & Signoret, J. L. (1992). Implicit access to knowledge derived from unrecognized faces in prosopagnosia. *Cerebral Cortex, 2*(5), 389-400.

Telling, A. L., Kumar, S., Meyer, A. S., & Humphreys, G. W. (2009). Electrophysiological evidence of semantic interference in visual search. *Journal of Cognitive Neuroscience*, doi:10.1162/jocn.2009.21348.

Tanaka, J. W. (2001). The entry point of face recognition: evidence for face expertise. *Journal of Experimental Psychology: General*, *130*(3), 534-43.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, 23, 457-82.

Tanaka, J. W., Curran, T., & Sheinberg, D. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science, 16*(2), 145-51.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*, 55-82.

Tarr, M. J. & Gauthier, I. (2000). FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nature Neuroscience*, *3*, 764-9.

Tarr, M. J., Kersten, D., & Bulthoff, H. H. (1998). Why the visual recognition system might encode the effects of illumination. *Vision Research*, *38*, 2259-75.

Tarr, M. J. & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233-82.

Thornhill, R., & Gangestad, S. W. (1994). Human fluctuating asymmetry and sexual behavior. *Psychological Science, 5*(5)*,* 297-302.

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*(12)*,* 452-60.

Tremoulet, P. D., & Feldman, J. (2000). Perception of animacy from the montion of a single object. *Perception, 29*(8), 943-51.

van der Linden, M., van Turennout, M., & Indefrey, P. (2009). Formation of category representations in superior temporal sulcus. *Journal of Cognitive Neuroscience, 6*, 1270-82.

Wells, G. L., & Hryciw, B. (1984). Memory for faces: Encoding and retrieval operations. *Memory & Cogntiion, 12,* 338-44.

Werner, S., & Noppeney, U. (2010). Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *Journal of Neuroscience, 30*(7), 2662-75.

Wheatley, T., Milleville, S. C., & Martin, A. (2007). Understanding animate agents: distinct roles for the social network and mirror system. *Psychological Science, 18*(6), 469-74.

Winograd, E. (1981). Elaboration and distinctiveness in memory for faces. *Journal of Experimental Psychology: Human Learning & Memory, 7,* 181-90.

Wisniewski, E., & Medin, D. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18*(2), 221-281.

Wong, Y. K., Gauthier, I. (2010). A multimodal neural network recruited by expertise with musical notation. *Journal of Cognitive Neuroscience, 22*(4), 695-713.

Wong, A. C.-N., Palmeri, T. J., & Gauthier, I. (2009a). Conditions for facelike expertise with objects: becoming a Ziggerin expert--but which type? *Psychological Science, 20*(9), 1108-17.

Wong, A. C.-N., Palmeri, T. J., Rogers, B., Gore, J. C., & Gauthier, I. (2009b). Beyond shape: How you learn about objects affects how they are represented in the visual cortex. *PLoS ONE, 4*(12), e8405.

Wong, A. C.-N., Jobard, G., James, K. H., James, T. W., & Gauthier, I. (2009c). Expertise with characters in alphabetic and nonalphabetic writing systems engage overlapping occipito-temporal areas. *Cognitive Neuropsychology*, *26*(1), 111-27.

Xu, Y. (2005). Revisiting the role of the fusiform face area in visual expertise. *Cerebral Cortex*, *15*, 1234-42.

Young, A., Hellawell, D., & Hay, D. (1987). Configurational information in face perception. *Perception*, *16*(6), 747-59.

Yovel, G., & Kanwisher, N. (2004). Face perception: domain specific, not process specific. *Neuron, 44*(5), 889-98.

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research*, *46*, 3802-11.