5-2020

Assessing Task-dependent Flexibility and the Temporal Dynamics of Object Categorization

Alyssa Magill Rohan
Bates College, arohan@bates.edu

Assessing Task-dependent Flexibility and the Temporal Dynamics of Object Categorization

An Honors Thesis Presented to

The Program of Neuroscience

Bates College

In partial fulfillment of the requirements for the

Degree of Bachelor of Arts

By

Alyssa Rohan

Lewiston, Maine

May 15, 2020

Acknowledgements

I would like to thank my incredible advisor, Professor Michelle Greene, for her guidance, patience and endless support. Michelle, you have introduced me to the complexities of scientific research and taught me skills I will draw on forever. From the very beginning, you were there for every setback and every step forward. Your constant support and dedication is truly inspiring. You have not only helped me develop as a scientist, but also as a student and person.

I would also like to thank all of the Bates Computational Vision Lab staff, namely Alexia Perugini, Jamie Siegart, Caitlyn McGlashan and Munashe Machoko. Their help running and organizing the experiments was paramount to the data collection process.

I would like to thank the neuroscience department and faculty for their support throughout my time at Bates, and to acknowledge the Bates Student Research Fund for helping fund this project.

Finally, I would like to thank my amazing family and friends for their unending support throughout this process. Thank you to my family for always believing in me and encouraging me to pursue my ambitions, no matter how crazy they seemed.

Table of Contents

List of Figures

Abstract

When we are presented with everyday objects, we inherently and effortlessly sort them into different categories. Nonetheless, the mechanisms by which this process occurs are not fully understood. There are three main tiers of categories into which the taxonomic tree can be separated: superordinate (e.g. *mammal*), basic (e.g. *cat*) and subordinate (e.g. *Siamese cat*). It is widely accepted that we most readily label objects at the basic level (Rosch et al., 1976). However, would this change if we were forced to attend to the features of a subordinate category? This study investigates the flexibility of assigning objects to the basic and subordinate categories as well as the neural time course of object classification. While recording EEG, participants indicated whether two successive images belong to either the same basic or subordinate category. They performed these tasks in two separate sessions alternating the taxonomic level. The EEG data obtained was analyzed through representational dissimilarity matrices (RDMs) creating a method of comparison for the neural responses of the electrodes at each time point to determine which category levels are most similar in terms of brain activity. We also used a multivariate regression approach to assess the magnitude of the relationship between neural data and the taxonomic level. We observed that the earliest neural responses seem optimized for basic-level categorization even when participants performed a different task. Therefore, these results emphasize the primacy of the basic level and provide further evidence for the automatic nature of object processing.

Introduction

When we encounter objects for the first time, we identify what they are and how they may be used. If we were not able to approach objects with some pre-existing knowledge about them, we would need to examine their every aspect before determining how they may be relevant to us. Given the incredibly large number of situations and items we encounter on a continuous basis, it would not be possible to navigate the world we live in today without pre-formed notions about objects.

Categories are not exclusively for objects; there are also categories about people, ideas and locations. Such categories help guide our behavior whenever we find ourselves in a new situation or encounter (Murphy, 2004). Object concepts are ubiquitous across all populations and are crucial for human functioning (Murphy, 2004).

Categorization helps reduce the infinite differences between entities into proportions that we are able to manage behaviorally and cognitively. Moreover, categories allow us to interact with objects that we have never seen before with ease, due to assumptions that we have about these objects (Murphy, 2004). For instance, I may have never seen a particular apple before, but based on my previous experience with other apples, I can assume that it is edible and that I will like it. These assumptions prevent us from wasting time and energy examining every new object (Murphy, 2004).

What is categorization?

While the idea of categorization may seem simple, it is necessary to clarify how categories will be defined throughout this research. Categories are often contrasted with concepts. Indeed, concepts are the mental representations of categories that allow us to appropriately respond to new objects we encounter, while categories represent the set of items that belong to that concept

(Murphy, 2004). For instance, the concept of *cat* is our core mental knowledge of cats, and the category *cat* represents all of the objects that belong to the cat category.

The question of how humans categorize objects is one that philosophers have pondered for centuries. Indeed, Plato proposed that categories are God-given and absolute; in which an object either possesses a certain feature and thus belongs to a category, or lacks that criterial property and therefore is excluded from that category (Murphy, 2004). This means that all categories can be defined by a set of necessary and sufficient features.

Aristotle then developed his text of the *Categories* in which he delineated ten main categories that could be used to separate all objects and concepts humans could express linguistically (Porphyry & Strange, 1992). In *Categories*, Aristotle proposed that in order to categorize an object, one must be able to distinguish between qualities, relations, substances, sorts and kinds (Ryle, 1940). This delineation was rather ambiguous and imprecise, and heavily relied on linguistics.

Later, Wittgenstein proposed that some words can have multiple "essences" that are relevant to the word's definition (Wittgenstein, 2001; Yeung et al., 2012). Wittgenstein illustrated these concepts through the example of the concept "game". One cannot think of a single feature that is necessary and sufficient to define all games, therefore the classical view of categories from Aristotle cannot be fully correct. Roger Brown and Eleanor Rosch proposed additional insights on the uses of category names in 1958 and 1976 respectively (Brown, 1958; Rosch et al., 1976). Brown proposed that objects can have many equally correct names, yet adults use some names more than others (Brown, 1958). He questioned the factors at play that determine why one label for a particular object is used more often than others. Later, Rosch expanded on these questions and introduced the basic level effect (Rosch et al., 1976).

Rosch defined categories as being characterized by a number of objects that are "considered equivalent" (Rosch et al., 1976). Furthermore, categories are not present in the natural environment. Instead, they are formed as a result of one's interactions with the world. Categories do not neatly follow low level features. In fact, many very different categories share surprisingly similar features. For instance, in Figure 1 the similar low-level features between a chihuahua and a blueberry muffin could lead to the unfortunate mistake of confusing the two (Folkens, 2017).

Figure 1. Images highlighting the overlap of low-level features between chihuahuas and blueberry muffins.

Hierarchical Membership and Basic Level Effect

Each object can be identified through many different labels, yet some object names are used more frequently than others (Brown, 1958). Indeed, concepts are known to have a hierarchical structure ranging from broad to specific (Miller et al., 1990). Naturally, categories also have different levels of specificity. Terms that are more abstract in nature are more inclusive in terms of categorization (Rosch et al., 1976). For instance, the term *mammal* is more abstract and therefore more inclusive than the term *Siamese cat*. An object is also usually labeled according to its utility (Miller et al., 1990). For instance, referring to an object as a "chair" implies that one can sit on it. Even though two different kinds of chairs, such as an armchair and a desk chair, may look very different, their overall function of being "somewhere to sit" is the same.

For the purposes of this research, we will define three levels of categories: superordinate (e.g. *clothing*), basic (e.g. *coat*) and subordinate (e.g. *parka*), as outlined by Rosch (Rosch et al., 1976). The superordinate category is the most abstract level, and members of this group share few features with each other (Rosch et al., 1976). The basic level is less abstract than the superordinate category and is thought to best mirror the structure of one's environment by forming simple categories with obvious attributes (Rosch et al., 1976). Plato referred to this idea as "carving nature by its joints" (Campbell et al., 2011). In essence, although nature does not present clear cut categories, there is some basic natural "anatomy" that will guide categorizations. Subordinate categories are the least abstract level, and members of this level will share many attributes with other subordinate categories (Rosch et al., 1976).

Research Supporting the Basic Level Effect

Many studies conducted by Rosch and ensuing researchers have established the salience of the basic level. Numerous researchers have turned to understanding the importance of basic level words in language acquisition. Indeed, children's words are mostly comprised of basic level names (Johnson et al., 1997; Mervis, 1983; Rosch et al., 1976). It is harder for young children to learn more abstract superordinate words, or more specific subordinate ones (Rosch et al., 1976). Furthermore, when sorting images, children are much more inclined to sort by basic level taxonomy (Rosch et al., 1976). Sorting at the superordinate level is a more complex task that children develop as they grow older (Rosch et al., 1976; Singer-Freeman & Bauer, 1997). Three year old children are able to sort images at the basic level with 99% accuracy contrasted to 55% accuracy at the superordinate level (Rosch et al., 1976). Moreover, even when young children already know a basic level label for an object, they are more inclined to learn a second basic level label than a subordinate or superordinate label (Mervis et al., 1994). Analysis of diary entries reveal that children are much less likely to be familiar with superordinate or subordinate level words (Mervis, 1983). Once children are 7-8 years old, their vocabulary is prevalent with mostly basic level words followed by superordinate and then subordinate level words (Mervis, 1983). Thus, not only is it easier for children to sort and label at the basic level, this level is also easier to understand and interact with.

In addition to experiments with participants, analysis of the shapes and overlap of pixels for objects have shown a basic level effect (Rosch et al., 1976). Outlines of images belonging to different hierarchical categories (such as clothing, furniture, modes of transportation and animals) were superimposed for the three levels of categories to find ratios of overlap between images (Rosch et al., 1976). The basic level objects were found to have a large and consistent increase in similarity of visual appearance as measured by pixel overlap, compared to the images from the superordinate level (Rosch et al., 1976). The subordinate level had a comparatively smaller advantage over the basic level (Rosch et al., 1976).

In another experiment, participants were asked to identify the outlines of two images belonging to the same hierarchical level that had been averaged together (Rosch et al., 1976). Participants were not able to recognize superordinate objects' shapes better than by chance (Rosch et al., 1976). Furthermore, the subordinate images were not more identifiable than the images from the basic level. Accordingly, the basic level is the most inclusive level at which objects were identifiable (Rosch et al., 1976).

Importantly, more simplistic models lead to increasingly reliable predictions on the categorization of a certain object, and thus how one can expect to interact with the object (Pothos & Chater, 2002). For instance, when encountering a new chair for the first time, certain salient features such as the legs and the back of the chair will classify the chair as such, thus providing information on how the chair can be used.

Indeed, at a behavioral level, objects are categorized into the basic level faster than at the superordinate or subordinate due to perceptual and distinctive attributes (Murphy & Brownell, 1985; Murphy & Smith, 1982; Rosch et al., 1976). For instance, a chair has distinctive features such as its legs and a flat surface to sit on. These features are not present in all superordinate objects, and while the subordinate level may have a few more common attributes, there will be a lot of overlap between subordinate categories (Murphy & Smith, 1982). The basic level provides the perfect balance between providing sufficient information and being cognitively economical. In contrast, the superordinate level does not provide much information, because members of that category possess few common attributes, and the subordinate category entails quantities of information that compromise cognitive efficiency (Rosch & Mervis, 1975). The basic level therefore provides the most relevant information for the least amount of cognitive effort.

The salience of the basic level can also be illustrated through the use of language. Shorter words, as measured by syllables or phonemes, tend to be used more frequently than longer ones (Zipf, 1935). However, there are some exceptions to this pattern. For instance, one is more likely to say “orange” or “banana” than “fruit” (even though the word fruit has fewer syllables), suggesting that there are other factors that influence the frequency at which words are used (Brown, 1958; Jolicoeur et al., 1984). Basic level words are often used to describe subordinate level objects (Murphy & Brownell, 1985). For instance, “park bench” would describe an object at the subordinate level, yet one is most likely to use the word “bench” to label this object. Basic level words are usually shorter than subordinate or superordinate words, and are more frequently used to describe objects (Brown, 1958; Rosch et al., 1976; Tanaka & Taylor, 1991; Wisniewski & Murphy, 1989). It is important to note that adults are more likely to use the shorter basic level words when speaking with children, as this label often also is the most “useful” (Brown, 1958).

Theories of the Basic Level

There is a plethora of research supporting the basic level effect, but what makes this level of categorization so “special”? The basic level maximizes cue validity, which is the conditional probability that an object will belong to a certain category given its possession of some feature (Murphy, 1982; Rosch et al., 1976). A cue is a property of a category, for instance, “*wings*” may be a cue for belonging to the category of *birds*. It is possible to sum up all of the individual cue validities for each feature within an abject in order to come up with a measure of total cue validity (Rosch et al., 1976; Tversky, 1977). Categories with high cue validity are differentiated from other categories yet have sufficient common attributes to be considered a self-contained category. Superordinate categories have lower cue validity, because they do not share many overlapping features, while the subordinate categories have low cue validity because they share many overlapping features with contrasting subordinate objects (Rosch et al., 1976). Indeed, objects with similar features naturally bundle together, and because the basic level has the highest cue validity and is the most abstract level that mirrors the attributes perceived in the world, objects we encounter are naturally classified at this level. This highlights the importance of abstraction, because it helps one make generalizations, as well as the importance of pragmatics, when further bundling does not help. Rosch proposed that the basic level category is the level of abstraction that contains the most information, while being the most differentiated from one another. Therefore, in everyday life, this is the level of categorization that is the most commonly referenced (Rosch et al., 1976).

According to the differentiation hypothesis proposed by Mervis and Crisafi, the order of acquisition in childhood of categorization schemes is determined by the degree of differentiation (measured by the relationship of within-category and between-category similarity) associated with each level (Mervis & Crisafi, 1982). In the basic level category, the similarity is high (objects within the same category look similar) and between these categories, the similarity is low (objects from different categories do not look similar to each other) (Mervis & Crisafi, 1982). Thus, the basic level is the most differentiated and should be acquired first.

In addition to being the most cognitively efficient category, Rosch also proposed that the motor movements associated with concrete objects reinforce the salience of the basic level effect (Rosch et al., 1976). For instance, sitting down on an office chair and a folding chair both entail similar specific sets of movements. According to Piagetian theory, a child’s first concepts of concrete objects stem from the actions related to that object (Nelson, 1974; Piaget & Cook, 1952). Associating an object with certain motor skills can inform one that objects within that category can be acted upon in similar ways (Nelson, 1974). Conversely, according to Gibsonian theory the visual system does not present the world “as it is”, but rather as the parts that enable action, thus the actions drive vision (Gibson, 1979). To illustrate these two theories, in agreement with Piaget, even though I may never have encountered a folding chair before, if I can categorize it at the basic level of “chair”, then I can reliably assume that I can interact with it similarly to how I would interact with any other chair. According to Gibson, it is the perception of being able to sit on the chair that creates the perception of “chair”, and my understanding of the basic level overcomes my lack of experience with the folding chair. In the basic level, the similarity of movements associated with objects within a category and the differences in movements between categories should both be maximized for that category to be as valid as possible (de la Rosa et al., 2014). Indeed, the basic level is the most general level of categorization that encompasses similar motor movements related to objects (de la Rosa et al., 2014; Mervis, 1983; Rosch & Lloyd, 1978).

Pothos and Chater proposed that the basic level provides the most compression of information (because this level has the most redundancy), making this level the most useful and easiest to encode (Pothos & Chater, 1998). When less information is provided, it is easier to categorize objects faster (Gosselin & Schyns, 2001). The interactions between the information demands of a categorization task and the availability of the object information both influence the speed of access (Gosselin & Schyns, 2001; Schyns, 1998). Thus, the basic level has the most efficient search through internal memory space, and is easier to access (Gosselin & Schyns, 2001).

One of the main purposes of categorization is to gain the maximum amount of information possible with the least cognitive effort (Rosch et al., 1976). The importance of simplicity in perception and cognition was first introduced by the Gestalt theory and the “law of Pragnanz”, which states that every stimulus is perceived in its most simple form (Chater, 1996; Nelson, 1974; Piaget & Cook, 1952). Later, Pomerantz and Kubovy (1986) developed the simplicity principle, which explains why one may adopt a certain “organization” of stimuli (Pomerantz & Kubovy, 1986). This principle proposes that the perceptual system’s preferred organization is the simplest one, meaning that it requires the least encoding of data and that it is also consistent with sensory input (Pomerantz & Kubovy, 1986). According to the cognitive economy theory, the basic level provides the most clear-cut information for the least amount of effort (Corter & Gluck, 1992; Rosch & Lloyd, 1978). This theory aligns with the notion that the basic level loses some of its primacy when one is an expert in that category (as discussed below), because in those situations, the cognitive load is subjectively more manageable.

Exceptions to the Basic Level Effect

One important factor that influences the process of categorization is the amount of knowledge one holds about the objects being categorized (Rosch et al., 1976; Tanaka & Taylor, 1991). In domains with which we have little experience, when asked to list category features, we list most features for categories at the basic level (Rosch et al., 1976). However, when we gain more experience, we list features of objects belonging to the basic and subordinate categories with equal frequency, and are as fast as identifying category membership between subordinate and basic level categories (Johnson & Mervis, 1997; Tanaka & Taylor, 1991). Thus, although the basic level never diminishes, the subordinate level is treated as equivalent to the basic level (Tanaka & Taylor, 1991).

A similar shift in categorization occurs when treating atypical items. Objects are considered atypical when they are distinctive and informative (Murphy & Brownell, 1985). For instance, at the subordinate level, *penguins* are considered atypical because they do not share many attributes with other *birds*. In these cases, atypical exemplars are usually categorized at the subordinate level, because they are specific, distinctive and differentiated from other subordinates (Jolicoeur et al., 1984; Murphy & Brownell, 1985). As categories become more differentiated, they adopt more qualities of the basic level, suggesting that categories fall on a hierarchal continuum rather than a dichotomous separation of basic and non-basic categories (Murphy & Brownell, 1985). To explain the discrepancy in categorization, Jolicoeur et al. suggest the notion of "entry point level". According to this theory, every object has a particular level at which it first connects with semantic memory, which often, but not exclusively, corresponds to the basic level (Jolicoeur et al., 1984). The entry point of an object is also influenced by the typicality of an object respective to the basic level (Jolicoeur et al., 1984).

Object Recognition and Categorization

It is hard to disentangle the processes of object recognition and object categorization. Indeed, recognition itself is often considered an act of categorization. To recognize an object is to recognize is as a member of a particular category (Bruner, 1957). Therefore, understanding the basic mechanisms of object categorization and recognition is important for assessing the flexibility of this process. The Recognition by Components theory (RBC), introduced by Biederman, suggests that objects are represented by the composition of viewpoint invariant geometric primitives that form the object (Biederman, 1987). Furthermore, object recognition occurs independently of the position from which the object is viewed (Biederman, 1987; DiCarlo et al., 2012). Some research suggests that objects automatically activate their conceptual representations, even when processing the object is not required for the task at hand (Greene & Fei-Fei, 2014a; Mathis, 2002). Moreover, compared to words, pictures take less time to process than written stimuli (Pellegrino et al., 1977; Rosch, 1975). Likewise, categorizing images may require little to no focal attention, and semantic representations, may be accessed without focal attention (Li et al., 2002; Mathis, 2002).

Object Categorization in the Brain

With respect to object categorization and recognition, the Lateral Occipital Complex (LOC) has been identified as the main brain area responsible for these processes (Amedi et al., 2001; Grill-Spector et al., 2001; Iordan et al., 2015; Malach et al., 1995). The LOC responds to both familiar and unfamiliar objects, and it is the earliest level at which there is evidence of basic level categorization (Grill-Spector et al., 2001; Iordan et al., 2015; Kanwisher et al., 1996). Furthermore, the representations of basic level categories seem to emerge as visual information is processed (Iordan et al., 2015). Indeed, neural activity is more self-similar to items that share a subordinate level category than to items that did not share a subordinate level category (Iordan et al., 2015). This is likely explained by the large overlap of low level features in subordinate levels (i.e. the visual resemblance of objects at this level) (Iordan et al., 2015). After the early visual areas, the fine grained representations are replaced by more general basic level ones (Iordan et al., 2015). The LOC tends to prioritize the representations at the basic level compared to other hierarchal levels (Iordan et al., 2015). Thus, the basic level advantage emerges through visual processing and processing important for subordinate level categorization likely occurs in the earlier stages.

Furthermore, the occipitotemporal cortex responds to objects based on their size in the real world (Julian et al., 2017). More specifically, larger objects tend to activate the medial parts of the ventral occipitotemporal cortex, whereas smaller objects activate the lateral parts (Konkle & Oliva, 2012). In addition to size, the hierarchy of animacy (or which animals most closely resemble humans) across animals appears to be reflected in the patterns of neural activity across cortical regions (Connolly et al., 2012). Both size and animacy are important to consider when generating stimuli for visual categorization tasks.

The Time Course of Object Categorization

While it is important to understand the neural localization of object categorization, the time course of these processes is crucial for drawing comparisons between neural responses of different hierarchal categorizations. The speed of processing visual information depends largely on the nature of the stimulus being processed and the task at hand. Multivariate pattern analysis (decoding) refers to a set of machine learning techniques through which one can assess the extent to which neural activity can be used to predict category membership. Early image decoding emerges in V1, responsible for low level processing, around 60 ms after stimulus onset (Cichy et al., 2016; Isik et al., 2014). However, it remains unclear whether the signal at 60 ms is due to object identification or low-level features of the object as a whole. The object representations required for behavioral judgements as peak at around 160 ms after stimulus onset (Bankson et al., 2018), implying that the earlier signals correspond more to low level feature analysis. In a go/no-go categorization task with ERP analysis, Thorpe et al. recorded brain activity that differentiated the target class (animals) from distractor images at 150 ms, suggesting that the lower level processing necessary to perform categorization judgements occurs in less time (Thorpe et al., 1996).

The speed of visual processing in humans depends largely on the content of what is being processed and the context in which this occurs. For instance, neural representations of scene layout (an abstract perceptual property) emerge at around 250 ms after stimulus onset, while earlier lower level visual analysis in V1 occurs within 100 ms (Cichy et al., 2014, 2015, 2017). Categorizing images as either cars or faces can be decoded within 135 ms of stimuli onset, regardless of the position of the object (Carlson et al., 2011). This means that the neural response to an object does not change as a function of the angle or position through which the object was viewed.

Other researchers have used identification tasks to observe differences in processing times depending on the location and position of the object (Isik et al., 2014). Interestingly, this result has not been completely reproduced, as Carlson et al. did not observe differences in latency between position invariant and non-position invariant tasks in their categorization task (Carlson et al., 2011). This suggests that the task of categorization requires generalization, and during this task neural correlates may already be somewhat invariant when the process of generalization occurs in the brain, explaining this discrepancy (Isik et al., 2014). Furthermore, categorizing images into abstract categories takes longer compared to less abstract categorizations (Carlson et al., 2013). This suggests that the earlier neural responses may in fact be due to low-level features that are associated with the categories, rather than the process of categorization itself (Carlson et al., 2013). While measuring neural responses to object categorization, it is often difficult to disentangle the categorization responses from the low-level features associated with the categories.

According to the theory that the basic level of categorization is the entry level of abstraction, basic level categorization occurs the fastest, but this does not necessarily mean it occurs first (Mack & Palmeri, 2011). Behaviorally, it does not take longer to detect the category of an object than to simply detect its presence, implying that similar mechanisms underlie identifying objects from a background and categorizing them at the basic level (Grill-Spector & Kanwisher, 2005). This segues into the question of whether the conscious recognition of an object and the neural processing of that object occur parallelly. In fact, the Reverse Hierarchy theory asserts that although vision progresses from low to high level features through the ventral stream, our conscious access to those features follow the opposite order (Ahissar & Hochstein, 2004). Based on this theory, basic level categorization may be faster because it is the first level at which the conscious mind recognizes objects.

Examining the Flexibility of Categorization

Objects can belong to multiple sets of categories even beyond the taxonomy outlined earlier. For instance, the object “pillow” can belong to the category of “objects necessary for sleep” or the category of “soft and fluffy items”.

Another facet of visual cognition is the flexibility of categorization. Factors such as expertise and atypicality both influence the salience of the basic level effect, as previously discussed. In reality, the hierarchical category boundaries are artificially constructed and therefore do not mirror the ways objects exist in our everyday lives. Indeed, categorization is goal and context dependent (Barsalou, 1983; Goldstone, 1994). The features one defines as representing a category are determined by one’s perception and cognition (Goldstone, 1994). Rosch defined categories as having overlapping features (Rosch et al., 1976). However, the judgement of similarity may be influenced by various factors, some of which may not be relevant to the task at hand (Goldstone, 1994). In fact, Goldstone proposes that similarity alone is not sufficient to form categories, yet it plays a crucial role in defining them.

Due to the issues raised above, research has investigated the flexibility of categorization. One example of this is the use of “ad hoc” categories. Barsalou described these categories as often used in “specialized contexts” that group otherwise seemingly unrelated objects together (e.g., “things to take on a camping trip”) (Barsalou, 1983). The ad hoc categories are different from “typical categories” in that they do not follow the general structure of the environment and are not well established in memory (Barsalou, 1983). Moreover, the ad hoc categories provide the opportunity for objects to belong to many categories. While natural categories are usually defined by certain characteristics, ad hoc categories are more goal-driven and, depending on the goal of the category, may not have definitive boundaries (Sandberg et al., 2012). Ad hoc categories, like traditional categories, are usually formed based on personal experiences (Vallée-Tourangeau et al., 1998).

Rosch and colleagues argued that individuals prefer to initially label objects according to their basic level name, yet in reality objects can be classified into many different subsequent categories (Barsalou, 1983; Rosch et al., 1976). For instance, although one may initially label a chair as such, this object may also fit into categories such as “objects in the dining room” or “things that can be used for emergency firewood” (Barsalou, 1983). The fact that humans are capable of assigning objects to different categories demonstrates that there is some flexibility in this process. Furthermore, in lieu of consciously adhering to rule-based categorization systems, one can classify objects based on task relevant goals. Schyns proposed a diagnostic recognition framework that posits the idea that object recognition and categorization are expressed through the interaction of the information required to perform the categorization task, and the visual information available (Schyns, 1998).

The type of task can influence the strategies employed for categorization. Indeed, when making subordinate level decisions, participants require more time viewing the stimulus to gain the needed diagnostic information for the categorization, implying that this more specific level of categorization requires more time and information (Malcolm et al., 2014). Furthermore, during subordinate scene categorizations, participants direct their fixation to the periphery of the image more often than when performing basic level categorizations, suggesting that subordinate decisions require a greater amount of diagnostic information (Malcolm et al., 2014). This demonstrates that the different categorization tasks require different amounts of information.

Not only does categorization task influence one’s eye movements, it also impacts one’s perception of the visual stimuli (Schyns & Oliva, 1999). Schyns and Oliva studied this by creating stimuli that were hybrids of male and female faces and different expressions, where one low spatial frequency image was superimposed on the high spatial frequencies of another image. Depending on the task participants were completing, they focused on different aspects of the images (Schyns & Oliva, 1999). Although facial processing recruits disparate neural processing, it is plausible that subordinate and basic categorization may elicit different image perceptions (Schyns & Oliva, 1999). One's goals influence the process of categorization, however, to what extent can this shift the basic level effect?

Overview

Overwhelming evidence supports the basic level theory proposed by Rosch et al., however, little is known about the extent to which the neural temporal dynamics may differ based on hierarchical categorizations (Rosch et al., 1976). The goal of this project was to compare the neural temporal dynamics for each of the three levels of categorization. This experiment used the same stimuli from Iordan et al., however instead of asking participants to perform a one-back task for identical images, we asked participants to respond to category membership (Iordan et al., 2015). Furthermore, participants completed two sessions wherein they categorized objects at either the subordinate or basic level.

In order to understand the time course of object categorization at different hierarchical levels, we leveraged the excellent temporal resolution of EEG. To avoid influences from extraneous factors such as animacy and object size, and focus the nature of the categorization task to be hierarchical, all of the objects featured were large and inanimate (Connolly et al., 2012; Julian et al., 2017). We hypothesized that subordinate categorization may have a temporal advantage over the basic level, given that the subordinate images are very similar in appearance. Additionally, we also predicted that the basic level of categorization would have more pronounced brain activity, due to the basic level effect.

We observed that the earliest neural responses seem optimized for basic-level categorization, and this remained true even when participants performed a subordinate level task. Therefore, these results emphasize the primacy of the basic level and provide further evidence for the automatic nature of object categorization.

Materials and Methods

Participants

16 participants (ages 18 -22, mean age: 19.54, 9 female) were recruited from the pool of undergraduate students at Bates College (due to data recording issues, two participants were excluded from the study). Each participant completed two separate experimental sessions following a within-participants design. For their time, they received financial compensation or course credit. Upon arrival, participants provided written, informed consent in compliance with the approval of the Bates College Institutional Review Board (IRB). They were then asked to identify their age, gender identity, and handedness. Before the beginning of the experiment, we ensured that all participants had at least 20/40 vision by administering the ETDRS eye chart test to account for discrepancies in visual acuity. In addition to the ETDRS test, we conducted a color blindness test (Ishihara, 1936). While recruiting participants, we specified that volunteers should not have a history of diagnosed neurological conditions or brain trauma.

Image Selection

Visual stimuli consisted of 1080 object photographs previously used by (Iordan et al., 2015), and were collected from the ImageNet online database (Deng et al., 2009). All of the images have been tested and verified to have a basic level effect (Iordan et al., 2015). For more details on the process of image and category selection, please refer to (Iordan et al., 2015). The image set only contains images of large (real-world size) and inanimate objects. Indeed, larger objects tend to activate the medial parts of the ventral occipitotemporal lobe, while small objects activate the more lateral areas (Konkle & Oliva, 2012). By solely using large objects, we limited the influence of size as a principle of categorization (Julian et al., 2017). Neural responses have also been shown to differ according to object animacy (Connolly et al., 2012). We controlled for this by only using inanimate objects. The images follow a three-tiered taxonomic hierarchy pattern with three superordinate levels (transportation, furniture, and musical instruments), nine basic levels (three for each superordinate), and 27 subordinate categories (three for each basic level). Each of the 27 subordinate-level categories contain 40 image exemplars, for a total of 1080 images. The three superordinate categories are vehicles, furniture, and musical instruments. The basic level categories are cars, airplanes, ships, chairs, beds, tables, drums, guitars, and pianos. Each basic level is composed of 3 subordinate categories (i.e., cars = sports car, sedan, station wagon; airplanes = airliner, biplane, fighter; ships = icebreaker, cargo ship, cruise ship; chairs = folding chair, armchair, straight chair; beds = canopy bed, sleigh bed, platform bed; tables = dining table, coffee table, pedestal table; drums = bass drum, snare drum, timpani; guitars = flamenco, Stratocaster, dreadnaught; pianos = Grand piano, Hammond organ, upright piano) (Iordan et al., 2015) (please see Figures 2 and 3 for a visual representation of the categories). To emphasize the object of interest and ensure that all neural activation is due to the object rather than the background of the image, all background originally present in the image was replaced with full-color 1/f noise, see Figure 4 for examples of stimuli. To ensure consistent retinal area activation, all images subtended approximately 8.5 x 8.5 degrees of visual angle.

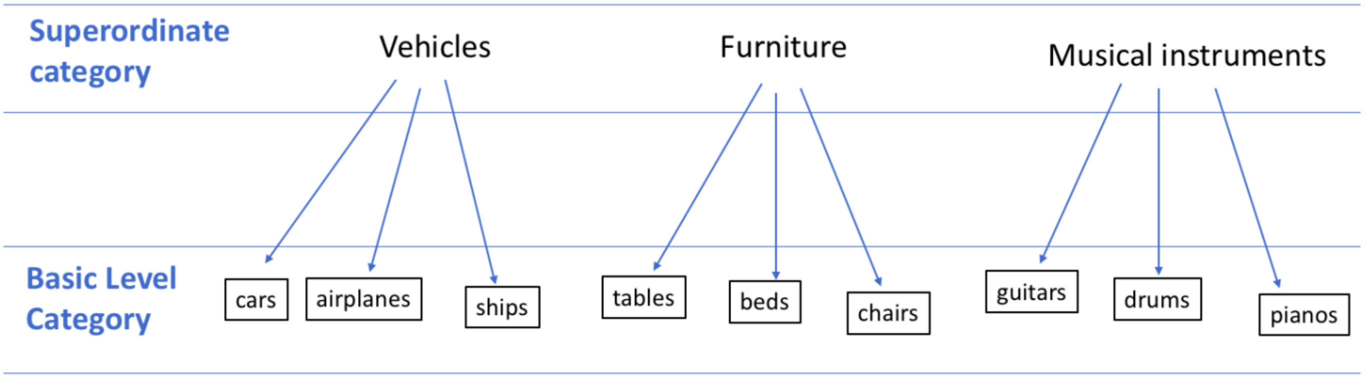

Figure 2. Relationships between the superordinate and basic level for the modes of transportation.

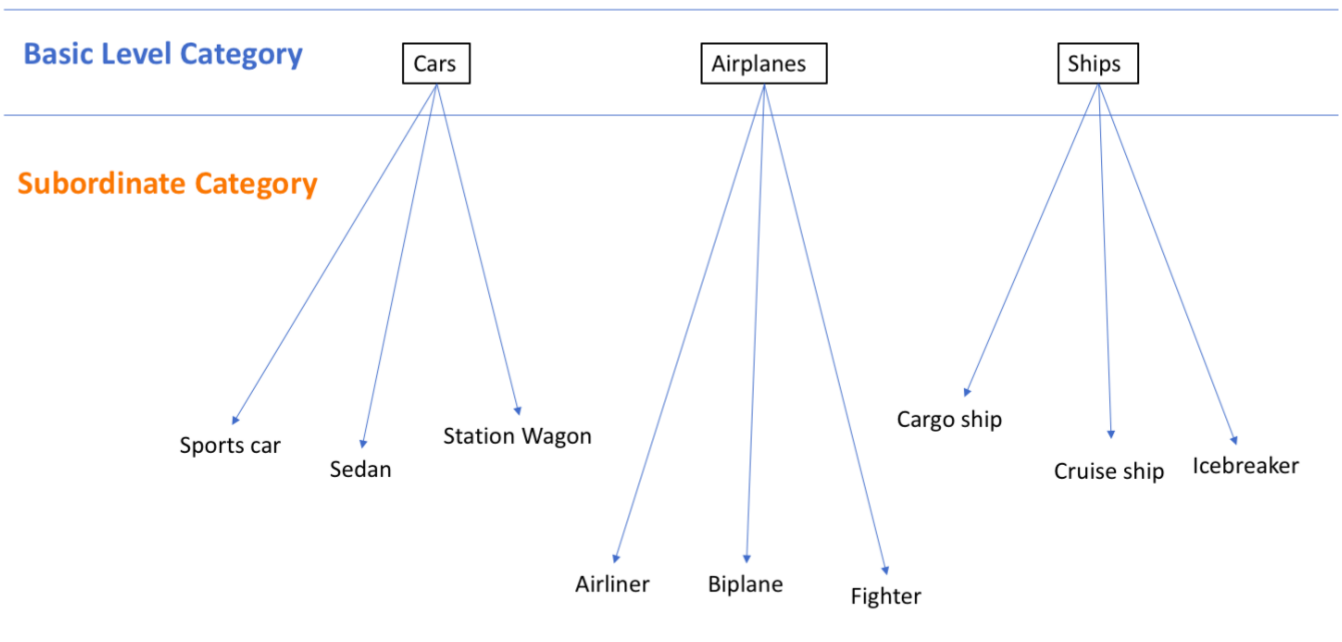

Figure 3. Relationships between the basic and subordinate levels for the modes of transportation.

Figure 4. Example Images. A sample image for each subordinate category is shown. Superordinate categories are separated by row with transportation on top, furniture in the middle and musical instruments on the bottom. Basic categories are organized together by three consecutive images in a row that belong to the same basic category.

Experimental Procedure

Upon arrival, participants were assigned either an odd or even number. Odd numbered participants performed the subordinate level category task first, while even numbered participants started with the basic level task. All participants performed the other condition during their second session of recording. Both tasks were counterbalanced to contain the same number of back-to-back category repeats. Participants were then fitted with an appropriately sized EEG cap and sat in front of a 27" LCD monitor (ASUS VG248Q). This monitor's resolution was 1920x1080 pixels and had a refresh rate of 100 Hz. While the EEG cap was being set up to record data, participants viewed a short Powerpoint presentation explaining the hierarchical structure of the categories and the visual distinctions of each. They were asked to minimize head and eye movements that could interfere with the EEG recording. Participants were shown the 1080 images one at a time for 250 ms each. Between each image, a fixation point appeared for a varying amount of time (250 ms $\pm$ 50 ms). Participants were instructed to press the space bar if two successive images belong to the same target category (1-back task). Depending on the condition, the target type was either subordinate or basic. When participants correctly identified two successive images as belonging to the same category, they received positive feedback (“Correct!” printed on the screen for 500 ms). They also received feedback when they missed or incorrectly identified two images as belonging to the same category (“Incorrect.”). No feedback was provided for correct rejections. Participants were presented with the opportunity to take a self-paced break after each set of 40 images (i.e. every 1-2 minutes).

EEG Recording

All EEG data was recorded using the ActiChamp acquisition system from a 64-channel sensor net with Ag/AgCl electrodes. A 24-bit analog-to-digital converter amplified and digitized all EEG signals at 1000 Hz. The impedances for each electrode were kept below optimal values defined by the system (15 kΩ). All EEG signals between 0.1 and 50 Hz were retained through band-pass filtering, which also removed DC offset and 60 Hz line noise. The vertex (Cz electrode) was chosen as an online reference. To detect eye movements and blinks, two electrodes were placed at the bottom and outer canthi of the right eye. The right mastoid reference was linked to the eye electrodes.

EEG Pre-processing

Raw EEG data was high-pass filtered at 1 Hz and re-referenced to the average of all electrodes. The data was epoched to preserve signals 100 ms before and 500 ms after the presentation of target stimuli. Independent component analysis (ICA) was run, and the experimenters visually identified and removed all components that reflected eye movements from the data.

Analysis

Behavioral

Categorization accuracy was identified as the percentage of correct responses (either hits or correct rejections) to either the basic of subordinate category objects as a function of the total number of trials.

Data Analysis

Creating Model RDMs

Model Representational Dissimilarity Matrices (RDMs) represent the theoretical constructs of category membership (Kriegeskorte et al., 2008). An RDM is a square matrix symmetrical about the diagonal that represents idealized category membership. Each RDM is a model of idealized category membership at one of the three levels of categorization. Each row and column of the matrix represents the distance (or inverse similarity) between any two items. Therefore, because the distance between item *j* and item *k* cannot be different from the distance between item *k* and item *j*, the matrix is symmetrical, and the diagonal of the matrix is undefined because it represents the distance between any item and itself. The model RDMs were organized based on the design and categorical nature of the images such that each subordinate category will be next to a similar subordinate category. For instance, *sports cars*, *sedans* and *station wagons* were all grouped together in the RDM alongside other transportation vehicles. There were three models that represented each hierarchical level of categorization (see Figure 5). Each model assumes a distance of 0 between images of the same category (blue squares in Figure 5) and maximal distances of 1 if the images did not belong to the same category (white areas in Figure 5).

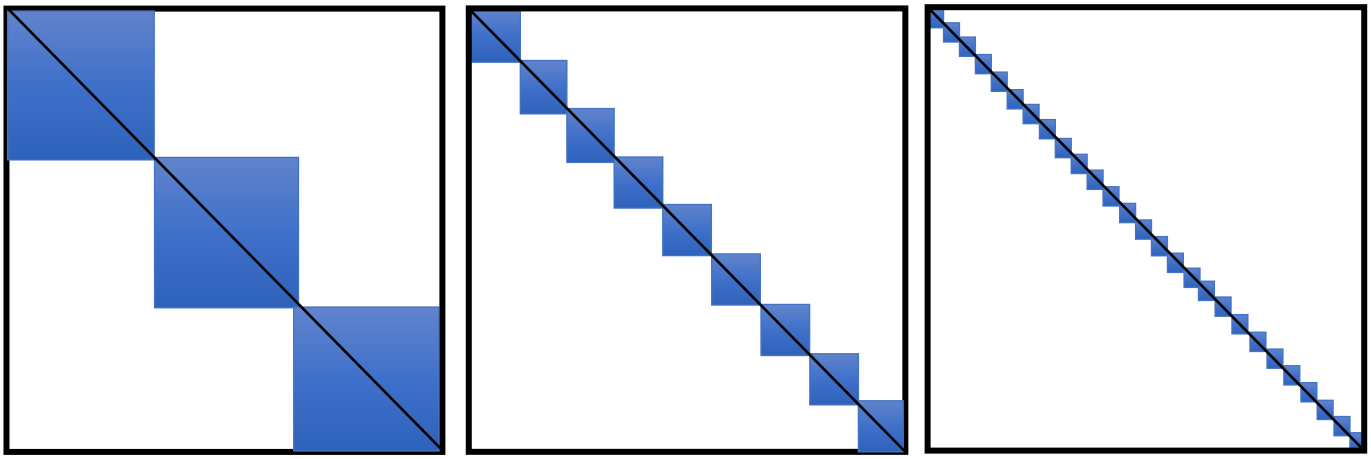

Figure 5. Model RDMs for each category from left to right: superordinate, basic and subordinate. Each box represents a category where the superordinate level has three types of objects, the basic has 9 and the subordinate has 27. Should there be no difference within categories, no matter which level, we would expect to see zeroes inside the boxes and ones around them. The diagonal line represents the 0 values, upon which the RDM would be symmetrical.

Creating Neural RDMs

In order to understand neural representation of all objects at each time point, we organized our EEG data into representational dissimilarity matrices (RDMs). The neural RDMs reflect the differences in how the brain processes all of the different images at a particular point in time. This method was chosen as it can help us distinguish between the neural responses for the same stimuli across different category levels.

In order to create a snapshot of visually evoked brain activity over time, we created independent RDMs from a set of occipitotemporal electrodes at each time point. Specifically, for each participant and each experimental session, we ordered trials by the same category order used to create the model RDMs. Next, we isolated data from all trials each time point across a pre-determined set of occipitotemporal electrodes (Pz, P3, P7, O1, Oz, O2, P4, P8, P1, P5, PO3, PO7, POz, PO4, PO8, P6, P2, Iz). This array is 1080 trials x 18 electrodes at each time point. From this array, we created an RDM by computing the distance between each trial using Euclidian distance.

Multiple Regression Analysis

In order to assess the amount of neural activity reflecting category membership at each of the three category levels, we employed multiple regression analysis. Specifically, we took the lower triangle of each of the three model RDMs to create a matrix of predictors. Similarly, we took the lower triangle of each neural RDM to serve as our dependent variable. Independent multiple regression analyses were performed at each of the 600 time points. For each analysis, we saved both the regression coefficients (β weight) as well as R^2 as a measure of effect size. Thus, at each time point, a β weight regression coefficient for each level of categorization can be used to assess the magnitude of the relation between models and neural activity. The β weight indicates the extent to which neural responses resembled the different two different categorization level conditions.

Feature Extraction

Because the regression analyses were run independently on each time point, we wanted to extract three summary statistics from the time series of β weight and R^2 for further analysis: onset time, maximum, and latency of maximum. The values of onset time for the β weights represented the first time at which the β weights were significantly different from 0. This corresponds to the time at which the category–specific neural responses were first observed. For the R^2 values, the pre-stimulus baseline was slightly above zero, therefore we tested for differences against the maximum value observed during the pre-stimulus baseline.

The maximum values for both R^2 and β weights were simply the maximum values observed across the epoch. For R2, the maximum value represents the extent to which the three category models explain the given neural RDM. For β weights, the maximum value represents the extent of similarity between a given model RDM and then given neural RDM.

Finally, the latency of the maximum β weights and R^2 values represent the time after stimulus onset when the maximum values were observed. These allow us to see *when* each model resembles neural activity.

Statistical Analysis

In order to obtain a unique onset value for each participant, we used a jackknife procedure. In this method, we iteratively computed onsets for 13 of the 14 participants (i.e. leaving one participant out each iteration). In order to control for the false discovery rate across time point, we considered only time points with at least ten sequential significant results as onsets. To correct for multiple comparisons across models in the β weight analysis, we employed Bonferroni correction with a corrected α to be 0.05/3 = 0.017.

Based on previous research, we expected to see a higher R^2 value during the subordinate task for the initial stages of neural processing, meaning that the brain response is likely due to the higher level of image similarity between the subordinate images. However, due to the basic level effect, we hypothesized that this initial subordinate activity should transition to the basic level. Should participants have the ability to shift categorization from basic back to subordinate based on task demands, we expected to see more subsequent activity in the brain as well as high levels of categorization accuracy.

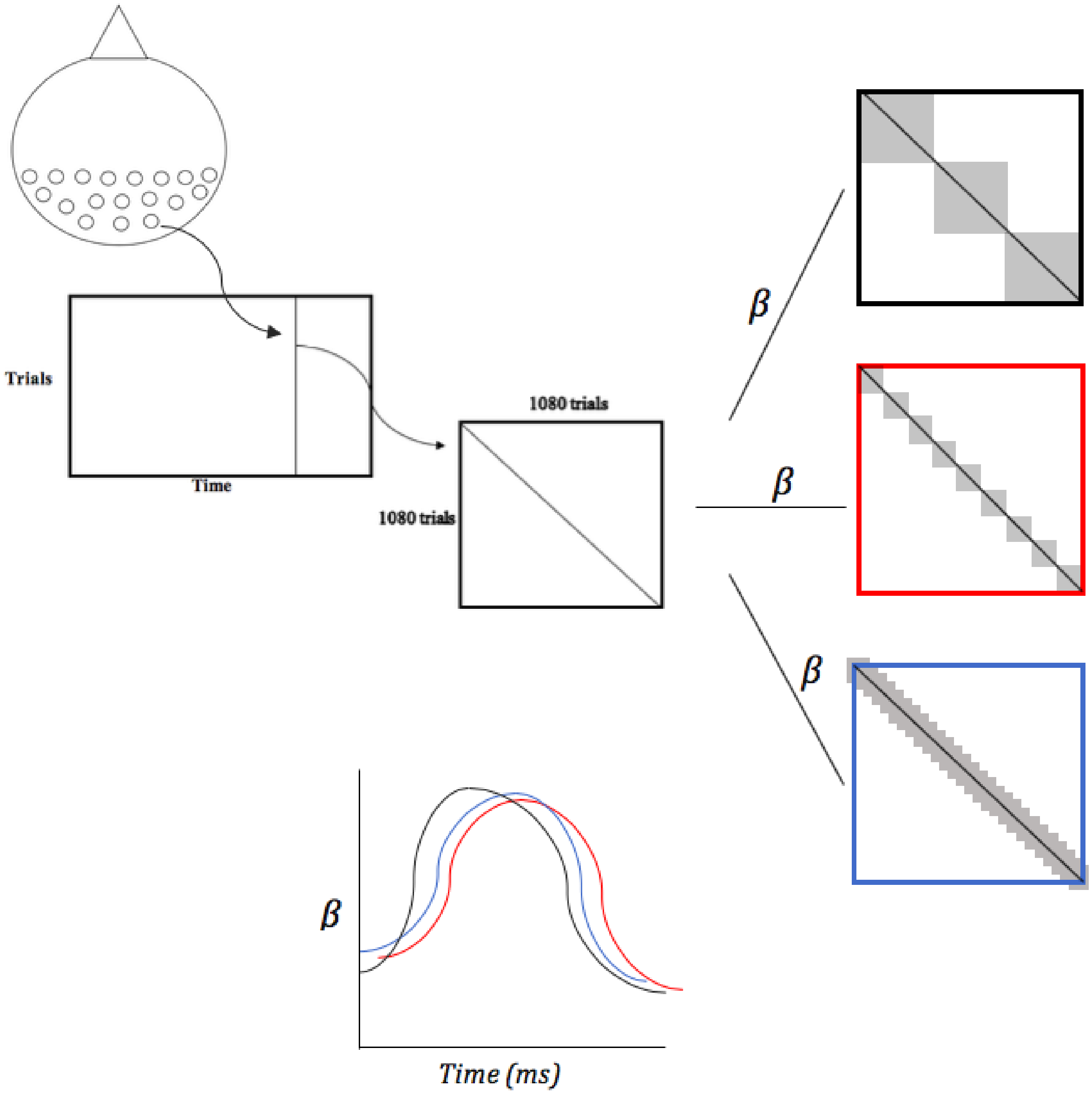

Figure 6. Visual representation of the statistical analysis with β weights.

Results

Behavioral Results

A paired sample t-test was conducted to determine whether one-back accuracy differed as a function of task. The mean accuracy for the basic condition (M= 94.6%, SD = 2.33%) was not significantly different from the one-back accuracy of the subordinate condition (M= 93.6%, SD = 3.81%), t (13) = 1.67, p = 0.118. Both categories had categorization accuracies well above 90%, which indicates that this may be a ceiling effect.

β weight Results

Repeated measures ANOVA were run to determine possible main effects and interactions for task (i.e. subordinate or basic) and model (i.e. subordinate, basic and superordinate) on onset, maximum and maximum latency.

Onset Time

In this analysis, the onset time represents the time at which the β weights for each of the three models were significantly different from zero and thus serves as a measure of when the neural activity starts to reflect each category level. The main effect of task for onset time approached statistical significance (F (1, 13) = 3.62, p = 0.0793). There was a main effect of model for onset time (F (2,26) = 11.2, p = 0.0003). Post-hoc paired t-tests were conducted with Bonferroni correction (new α = 0.05/3 = 0.017). The onset time in the subordinate model (M = 75.4 ms) was significantly higher than for the basic model (M = 67.2 ms), t (13) = 2.98, p = 0.0107. Furthermore, the onset time for the basic model was significantly earlier than the superordinate model (M = 91.1 ms), t (13) = 3.89, p = 0.00186. There was no significant difference between the onset values for the subordinate and superordinate models t (13) = 2.70, p = 0.0180. There was no interaction between task and model for onset time (F (2, 26) = 0.750, p = 0.482). This suggests that the basic level model had an earlier onset than the subordinate and superordinate models. The time of onset between the superordinate and subordinate models was not significantly different.

Maximum β weights

In addition, for each model, we measured the maximum β weight observed across time. These values represent the strength of the relationship between neural activity and each category model. There was no main effect of task (F (1,13) = 2.05, p = 0.175), nor an interaction of task and model (F (2,26) = 0.949, p = 0.400). As we observed with onset time, there was a main effect of model (F (2,26) = 30.6, p = 1.47x10^-7). Paired post-hoc t-tests (Bonferroni corrected) were conducted and revealed that both the maximum β weight in the subordinate model (M = 0.124) and the superordinate model (M = 0.0579), were significantly lower than the β weights in the basic model (M = 0.225), t (13) = 4.05, p = 0.00137); t (13) = 6.36, p = 1.65x10^-5, respectively. Finally, the β weights for the subordinate model were significantly higher than the β weights for the superordinate model t (13) = 5.78, p = 6.40x10^-5. Altogether, this suggests that the neural responses more closely resembled the basic model compared to the superordinate and subordinate models. Furthermore, the maximum neural response for the subordinate model was higher than that of the superordinate model.

Latency of Maximum β Weights

Finally, we examined the latency where the maximum β weight values were observed for each category model. We observed no main effect of task (F (1, 13) = 0.393, p = 0.542), nor model (F (2,26) = 0.837, p = 0.444), nor any interaction between task and model (F (2,26) = 0.800, p = 0.460). Therefore, there was no difference between the time of maximum between the three levels.

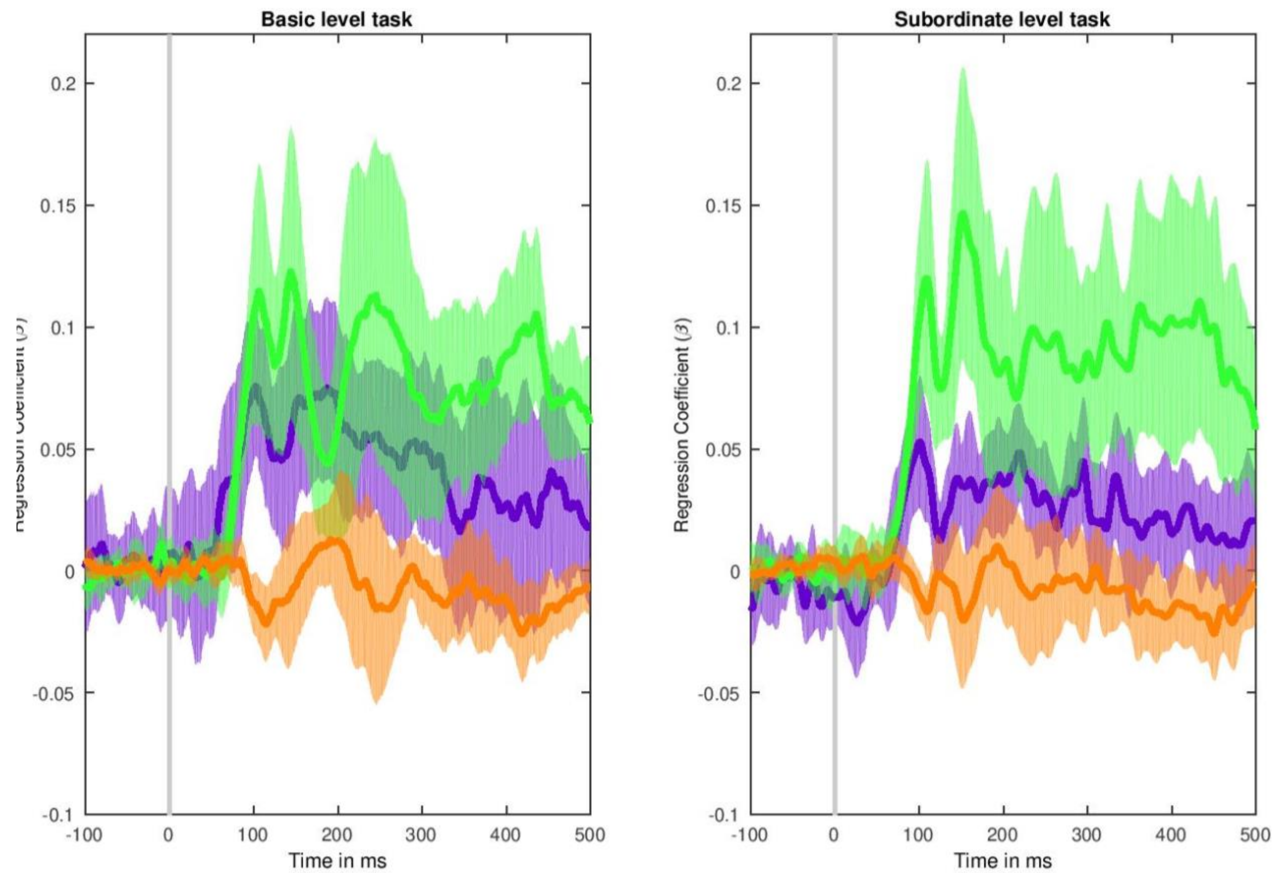

Figure 7. β weight values as a function of time for the basic and subordinate level tasks. Repeated measures ANOVA revealed that the basic model (green) had an earlier onset than the subordinate model (purple) and the superordinate model (orange) which did not vary as a function of task. The transparent curves represent error bars at each time point. The max of the basic model was higher in both the basic level task and the subordinate level task.

R^2 Results

While the β weights tell us the strength of the relationship between neural activity and each category model, the R^2 values indicate how much the neural activity resembles all three of our models. Paired t-tests were conducted to determine the potential differences in the R^2 values for onset time, maximum R^2 and latency of maximum R^2.

Onset Time

The mean R^2onset time for the basic condition (M= 20.9ms, SD = 8.44ms) was significantly earlier from the R^2 onset time of the subordinate condition (M= 52.3ms, SD = 15.6ms), t (13) = 5.74, p = 6.81x10^-5. The difference in the R^2 onset indicates that when the participants are engaging in the basic-level categorization, all three models can explain early brain activity better than when they are doing subordinate-level categorization.

Maximum R_2

The maximum R^2 values for the basic condition (M=1.92 x 10^-4, SD =1.46 x10^-4) was not significantly different from the maximum R^2 values of the subordinate condition (M=1.94 x 10^-4, SD =1.34 x10^-4), t (13) <1. These results imply that there was no reliable difference in the extent to which the models predicted neural activity across task.

Latency of Maximum R_2

The latency of maximum R^2 values in the basic condition (M=170ms, SD =70.1ms) was not significantly different from the latency of the maximum R^2 values in the subordinate condition

(M=171ms, SD =94.9ms), $t\,(13)$) <1, p = 0.945. Based on this, neither the basic nor the subordinate task has an advantage when it comes to model for the latency of the maxima.

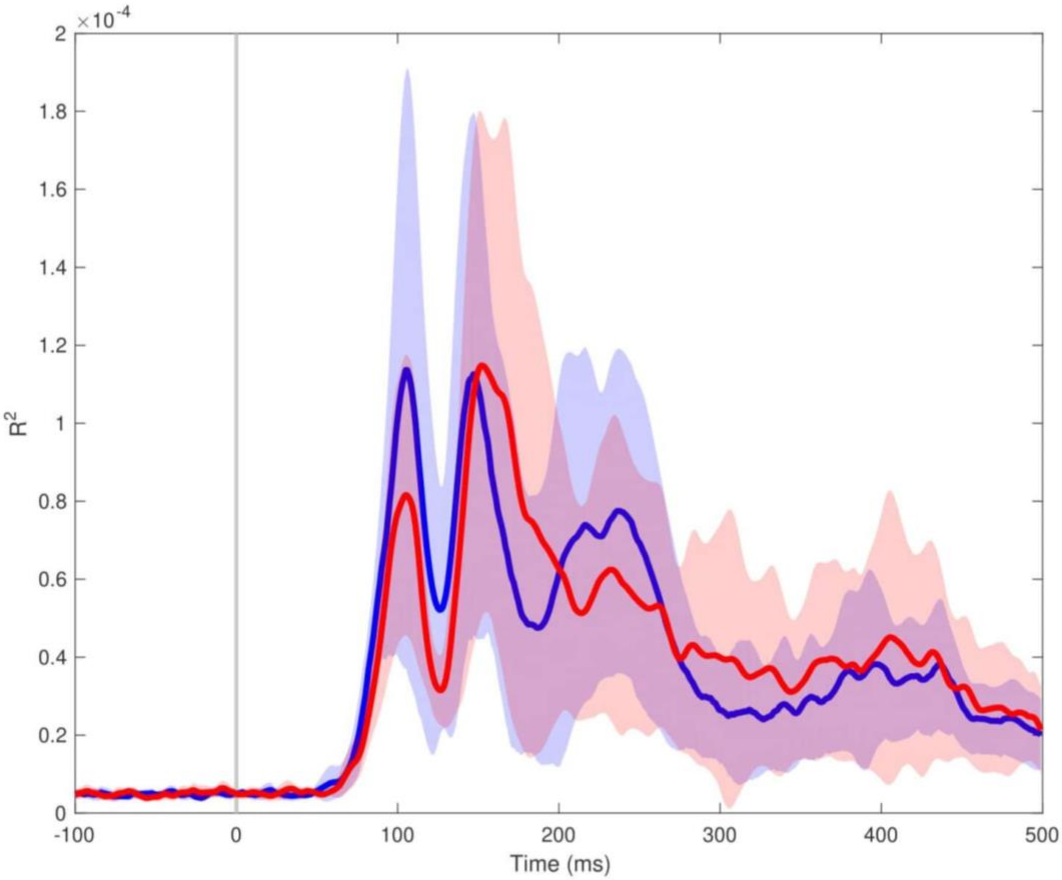

Figure 8. Regression analysis for the subordinate and basic task as a function of time. R^2values represent the extent to which all of the models combined resemble neural data in each task. Paired t-tests revealed that during the basic level task (blue), all models accounted more for early brain activity that during the subordinate task (red). The transparent curves represent error bars.

Discussion

The overall purpose of this experiment was to compare the temporal dynamics of neural responses for each of the three levels of object categorization. We found that the neural patterns resembled the basic level earlier than the subordinate or superordinate levels. The basic model had a larger neural response than both the subordinate and superordinate models, and the subordinate model had a greater response than the superordinate model. This supports the saliency of the basic level effect and suggests that categorization at this level is somewhat automatic. The superordinate level model had the weakest relationship with the neural data we observed, indicating that neural responses do not seem particularly well-tuned to this very abstract level of categorization. It accounts for less of the neural response. Neither the intensity of the neural response (as indexed by the regression coefficients) nor the time of maximum intensity of the neural activity varied as a function of the task performed. This suggests that despite performing different tasks, there was no significant difference in the strength of the relationship between the category models and the neural data.

From analyzing the R^2 results, we determined that when participants performed the basic level task, all three models can explain brain activity earlier than when they are doing the subordinate level task. The category models had equal efficacy in explaining the neural activity to both tasks. This means that apart from the early neural responses, both tasks engaged category-related neural processing equally. There is no difference in how much the models resembled brain activity according to task. Altogether, these results show that the neural responses are optimized for the basic level of categorization even when the participants performed the subordinate level task.

Many of our results were contrary to our a priori hypotheses. We expected to see greater categorization accuracy in the basic level task compared to the subordinate level task due to the basic level effect, but this was not the case. All of the participants had very high accuracy at both tasks (94.6% for basic and 93.6% for subordinate), and if any basic level effect were present, it was likely masked by the very high performance in both tasks.

We expected to see an early advantage for subordinate-level categorization due to the increased visual similarity that exists between members of these categories. To our surprise, we found that the basic level has an onset advantage. Categorizing pictures at a more abstract level takes longer than less abstract tasks (T. Carlson et al., 2013; Cichy et al., 2017). Between the two tasks in our experiment, the basic level was the most abstract level, and we therefore expected the onset time for this level to be later than those for the subordinate task, however this was not what we observed. The stimuli we used differed from those that Carlson and Cichy used. In Cichy et al.'s experiment, the background of the images were still present, providing more low-level features that can be taken advantage of such as the blue sky behind an airplane (Cichy et al., 2017). When the background of an object is present, it is less clear whether the neural response is specific to the object per se, or to the background. Furthermore, the current experiment featured 40 images per subordinate category, while Cichy and Carlson had only three images in the more specific categories (T. Carlson et al., 2013; Cichy et al., 2017). This difference may explain the discrepancy in our results. With a small number of images in each category, the neural activity might reflect the low-level features of those specific images rather than generalizing across the category. By having more images per category, we are better reflecting the natural variability of each subordinate category and thus focusing our analysis to generalize across the category rather than just a couple of specific examples.

Because our method forces generalization, unlike Carlson, we did not see any category-specific brain activity around 60-80 ms post-image onset that is typical of the processing of low-level features. There are at least two possible explanations for this discrepancy. Firstly, it is possible that the low-level features of the objects were less salient for this categorization task compared to those the participants viewed in Carlson's experiment because we sampled from a relatively large number of exemplars within each subordinate category, which reduces the probability that neural activity would match based solely on low-level visual features. Secondly, the earlier neural signals may have been less intense compared to those that occurred later. This may have occurred due to two reasons, the first being that only a subset of our electrodes showed the effect. Secondly, we may not have detected the early category-specific brain activity because the brain signals were not observable with EEG. The former explanation seems much more likely because early visually evoked potentials seem to originate from V1 and other early areas (Jeffreys & Axford, 1972). Therefore, we know that it is possible for the EEG to detect signals from V1, so it is less likely that we are not observing the signal, and more likely that we are averaging it out across electrodes.

With fMRI, Iordan at al. found some subordinate-level category grouping in early visual areas, including V1(Iordan et al., 2015). We used the same stimuli as this previous experiment, so it might seem unusual that we did not observe early neural responses that were associated with the subordinate level. However, we grouped over both central occipital electrodes as well as occipitotemporal ones in our analysis. Therefore, by only examining neural responses in central occipital electrodes, we might observe a similar early subordinate advantage.

The earlier onset on the basic level suggests that even the earliest category-related brain signals seem preferentially tuned to the basic level, which is the most behaviorally relevant level. Additionally, we are seeing that the brain optimizes for subsequent behavior at the earlier time points, suggesting that this may be a result of top-down activity. These ideas also make sense in terms of the Reverse Hierarchy theory (Ahissar & Hochstein, 2004). In this theory, the order of processing of our experience is the reverse order of processing performed by the visual system (Ahissar & Hochstein, 2004). In essence, V1 is responsible for processing very basic features such as edges and colors, while high-level areas such as the IT process categories. We are first aware of the results from the higher-level areas such as the IT rather than those from lower-level areas. In our experiment, we recorded neural activity over occipitotemporal areas, thus the activity what we recorded may have resulted from information being fed back to early visual areas from the higher visual areas. Secondary analyses could examine more broad activity across the scalp to determine the origin of category information.

At a theoretical level, the results of our study support the primacy of the basic level. Regardless of the task the participants performed, the regression coefficients were the highest for the basic level. The lack of task-related difference may be because the behavioral accuracy for both tasks was very high, suggesting that the tasks were easy. Given the simplicity of the task, we might not have been maximizing changes in the task. Should the task be made more difficult by lowering the contrast, adding noise to the images, or by reducing the time the images were on the screen, we might get more task-related engagement and therefore see stronger task effects.

Our results also suggest that the processes of object recognition and categorization cannot be divorced with this paradigm. This aligns with Bruner's idea that to recognize an object is to link it to a specific category (Bruner, 1957). Our results suggest that there is a basic level advantage even when the participants are performing a subordinate task. This aligns with some research suggesting that objects automatically activate their conceptual representations regardless of whether this process is required for the task at hand (Greene & Fei-Fei, 2014b; Mathis, 2002).

In our experiment, the superordinate level contributed very little to the observed neural patterns. This hierarchical level is the most proves to be difficult for children to use (Rosch et al., 1976; Singer-Freeman & Bauer, 1997). In addition to testing the subordinate and basic level tasks, it would be interesting to introduce a superordinate condition wherein participants categorized objects at the superordinate level. This may create a larger superordinate effect than the one we observed.

Furthermore, as previously discussed, the hierarchical categories are not the only categories available. A similar experiment to the one conducted here could use ad hoc categories instead of a hierarchical taxonomy. For instance, participants could categorize objects into categories such as "objects that can be used as emergency firewood" and "objects costing less than one million dollars". This type of categorization requires more conscious processing than the typical categorization task we asked participants to perform, and perhaps in this type of experiment there would be an effect of task. Therefore, it is possible that the onset times and the latencies of the maxima in an ad hoc categorization task will be different from the hierarchical categorization task. Finally, we decided to focus our analysis on the occipitotemporal electrodes because these record the majority of visual processing. Future analysis could investigate the patterns across the scalp at every time point. Decoding would allow us to determine the extent to which the brain has information of the three different hierarchical levels.

Altogether, our experimental results indicated that the earliest neural responses seem optimized for basic-level categorization, and this remains true even when participants performed a different task. Therefore, these results emphasize the primacy of the basic level and provide further evidence for the automatic nature of object categorization.

References

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. https://doi.org/10.1016/j.tics.2004.08.011

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience*, *4*(3), 324–330. https://doi.org/10.1038/85201

Bankson, B. B., Hebart, M. N., Groen, I. I. A., & Baker, C. I. (2018). The temporal evolution of conceptual object representations revealed through models of behavior, semantics and deep neural networks. *NeuroImage*, *178*(Journal Article), 172–182. https://doi.org/10.1016/j.neuroimage.2018.05.037

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*(3), 211–227.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147. https://doi.org/10.1037/0033-295X.94.2.115

Brown, R. (1958). How shall a thing be called? *Psychological Review*, *65*(1), 14–21. https://doi.org/10.1037/h0041727

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*(2), 123–152. https://doi.org/10.1037/h0043805

Campbell, J. K., O'Rourke, M., & Slater, M. H. (2011). *Carving Nature at Its Joints: Natural Kinds in Metaphysics and Science*. MIT Press.

Carlson, T. A., Hogendoorn, H., Kanai, R., Mesik, J., & Turret, J. (2011). High temporal resolution decoding of object position and category. *Journal of Vision*, *11*(10). https://doi.org/10.1167/11.10.9

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision*, *13*(10). https://doi.org/10.1167/13.10.1

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review*, *103*(3), 566–581. https://doi.org/10.1037/0033-295X.103.3.566

Cichy, R. M., Khosla, A., Pantazis, D., & Oliva, A. (2017). Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. *NeuroImage*, *153*(Journal Article), 346–358. https://doi.org/10.1016/j.neuroimage.2016.03.063

Cichy, R. M., Khosla, A., Pantazis, D., Torralba, A., & Oliva, A. (2016). Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence. *Scientific Reports*, *6*(1), 27755. https://doi.org/10.1038/srep27755

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience*, *17*(3), 455–462. https://doi.org/10.1038/nn.3635

Cichy, R. M., Ramirez, F. M., & Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *NeuroImage*, *121*, 193–204. https://doi.org/10.1016/j.neuroimage.2015.07.011

Connolly, A. C., Guntupalli, J. S., Gors, J. D., Hanke, M., Halchenko, Y. O., Wu, Y., Abdi, H., & Haxby, J. V. (2012). The representation of biological classes in the human brain. *The*

Journal of Neuroscience : The Official Journal of the Society for Neuroscience. https://doi.org/10.1523/JNEUROSCI.5547-11.2012

Corter, J. E., & Gluck, M. A. (1992). Explaining Basic Categories: Feature Predictability and Information. *Psychological Bulletin*, *111*(2), 291–303.

de la Rosa, S., Choudhery, R. N., Curio, C., Ullman, S., Assif, L., & Bülthoff, H. H. (2014). Visual categorization of social interactions. *Visual Cognition*, *22*(9–10), 1233–1271. https://doi.org/10.1080/13506285.2014.991368

Deng, J., Socher, R., Fei-Fei, L., Dong, W., Li, K., & Li, L.-J. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, *00*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Folkens, B. (2017, May 26). *Chihuahua or Muffin?* Medium. https://blog.cloudsight.ai/chihuahua-or-muffin-1bdf02ec1680

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

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition*, *52*(2), 125–157. https://doi.org/10.1016/0010-0277(94)90065-5

Gosselin, F., & Schyns, P. G. (2001). Why do we SLIP to the basic level? Computational constraints and their implementation. *Psychological Review*, *108*(4), 735–758.

Greene, M. R., & Fei-Fei, L. (2014a). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, *14*(1), 14–14. https://doi.org/10.1167/14.1.14

Greene, M. R., & Fei-Fei, L. (2014b). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, *14*(1), 14–14. https://doi.org/10.1167/14.1.14

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10), 1409–1422. https://doi.org/10.1016/S0042-6989(01)00073-6

Iordan, M. C., Greene, M. R., Beck, D. M., & Fei-Fei, L. (2015). Basic Level Category Structure Emerges Gradually across Human Ventral Visual Cortex. *Journal of Cognitive Neuroscience*, *27*(7), 1427–1446. https://doi.org/10.1162/jocn_a_00790

Ishihara, S. (1936). *The series of plates designed as tests for colour-blindness* (7th (complete) 32 plates.). Kanehara & Co.

Isik, L., Meyers, E. M., Leibo, J. Z., & Poggio, T. (2014). The dynamics of invariant object recognition in the human visual system. *Journal of Neurophysiology*, *111*(1), 91–102. https://doi.org/10.1152/jn.00394.2013

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Experimental Brain Research*, *16*(1), 1–21. https://doi.org/10.1007/BF00233371

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General*, *126*(3), 248–277. https://doi.org/10.1037/0096-3445.126.3.248

Johnson, K. E., Scott, P., & Mervis, C. B. (1997). Development of Children's Understanding of Basic-Subordinate Inclusion Relations. *Developmental Psychology*, *33*(5), 745–763. https://doi.org/10.1037/0012-1649.33.5.745

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology*, *16*(2), 243–275.

Julian, J. B., Ryan, J., & Epstein, R. A. (2017). Coding of Object Size and Object Category in Human Visual Cortex. *Cerebral Cortex*, *27*(6), 3095–3109. https://doi.org/10.1093/cercor/bhw150

Kanwisher, N., Chun, M. M., McDermott, J., & Ledden, P. J. (1996). Functional imaging of human visual recognition. *Cognitive Brain Research*, *5*(1), 55–67. https://doi.org/10.1016/S0926-6410(96)00041-9

Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipito-temporal cortex. *Neuron*, *74*(6), 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596–9601. https://doi.org/10.1073/pnas.092277599

Mack, M. L., & Palmeri, T. J. (2011). The Timing of Visual Object Categorization. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00165

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., Ledden, P. J., Brady, T. J., Rosen, B. R., & Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(18), 8135–8139.

Malcolm, G. L., Nuthmann, A., & Schyns, P. G. (2014). Beyond Gist: Strategic and Incremental Information Accumulation for Scene Categorization. *Psychological Science*, *25*(5), 1087–1097. https://doi.org/10.1177/0956797614522816

Mathis, K. M. (2002). Semantic interference from objects both in and out of a scene context. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *28*(1), 171–182. https://doi.org/10.1037/0278-7393.28.1.171

Mervis, C. B. (1983). Acquisition of a lexicon. *Contemporary Educational Psychology*, *8*(3), 210–236. https://doi.org/10.1016/0361-476X(83)90015-2

Mervis, C. B., & Crisafi, M. A. (1982). Order of Acquisition of Subordinate-, Basic-, and Superordinate-Level Categories. *Child Development*, *53*(1), 258–266. https://doi.org/10.1111/j.1467-8624.1982.tb01315.x

Mervis, C. B., Golinkoff, R. M., & Bertrand, J. (1994). Two-Year-Olds Readily Learn Multiple Labels for the Same Basic-Level Category. *Child Development*, *65*(4), 1163–1177. https://doi.org/10.1111/j.1467-8624.1994.tb00810.x

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An On-line Lexical Database*. *International Journal of Lexicography*, *3*(4), 235–244. https://doi.org/10.1093/ijl/3.4.235

Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin*, *91*(1), 174–177. https://doi.org/10.1037/0033-2909.91.1.174

Murphy, G. (2004). *The Big Book of Concepts*. MIT Press.

Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *11*(1), 70–84. https://doi.org/10.1037//0278-7393.11.1.70

Murphy, G. L., & Smith, E. E. (1982). Basic-Level Superiority in Picture Categorization. *Journal of Verbal Learning and Verbal Behavior; New York*, *21*(1), 1–20.

Nelson, K. (1974). Concept, word, and sentence: Interrelations in acquisition and development. *Psychological Review*, *81*(4), 267–285. https://doi.org/10.1037/h0036592

Pellegrino, J. W., Rosinski, R. R., Chiesi, H. L., & Siegel, A. (1977). Picture-word differences in decision latency: An analysis of single and dual memory models. *Memory & Cognition*, *5*(4), 383–396. https://doi.org/10.3758/BF03197377

Piaget, J., & Cook, M. (1952). The origins of intelligence in children. *New York: International Unitversities Press*, *8*, 442.

Pomerantz, J. R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization: Simplicity and likelihood principles. In *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance* (pp. 1–46). John Wiley & Sons.

Porphyry, & Strange, S. K. (1992). *On Aristotle's categories*. Cornell University Press.

Pothos, E. M., & Chater, N. (1998). *Rational categories*.

Pothos, Emmanuel M., & Chater, N. (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science*, *26*(3), 303–343. https://doi.org/10.1207/s15516709cog2603_6

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*(3), 192–233. https://doi.org/10.1037/0096-3445.104.3.192

Rosch, E., & Lloyd, B. B. (1978). *Cognition and Categorization*. L. Erlbaum Associates.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Ryle, G. (1940). Categories. *Oxford University Press on Behalf of The Aristotelian Society*, *5*(2), 83–84. https://doi.org/10.2307/2266206

Sandberg, C., Sebastian, R., & Kiran, S. (2012). Typicality mediates performance during category verification in both ad-hoc and well-defined categories. *Journal of Communication Disorders*, *45*(2), 69–83. https://doi.org/10.1016/j.jcomdis.2011.12.004

Schyns, P. G. (1998). Diagnostic recognition: Task constraints, object information, and their interactions. *Cognition*, *67*(1), 147–179. https://doi.org/10.1016/S0010-0277(98)00016-X

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*(3), 243–265. https://doi.org/10.1016/S0010-0277(98)00069-9

Singer-Freeman, K. E., & Bauer, P. J. (1997). Sorting out language and level: Examining the relation between productive vocabulary and category differentiation. *First Language*, *17*(51), 241–270. https://doi.org/10.1177/014272379701705111

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology*, *23*(3), 457–482. https://doi.org/10.1016/0010-0285(91)90016-H

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. https://doi.org/10.1038/381520a0

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327–352. https://doi.org/10.1037/0033-295X.84.4.327

Vallée-Tourangeau, F., Anthony, S. H., & Austin, N. G. (1998). Strategies for generating multiple instances of common and ad hoc categories. *Memory (Hove, England)*, *6*(5), 555–592. https://doi.org/10.1080/741943085

Wisniewski, E. J., & Murphy, G. L. (1989). Superordinate and Basic Category Names in Discourse: A Textual Analysis. *Discourse Processes*, *12*(2), 245–261. https://doi.org/10.1080/01638538909544728

Wittgenstein, L. (2001). *Philosophical Investigations: The German Text, with a Revised English Translation*. Blackwell.

Yeung, J. F. Y., Chan, A. P. C., & Chan, D. W. M. (2012). Defining relational contracting from the Wittgenstein family-resemblance philosophy. *International Journal of Project Management*, *30*(2), 225–239. https://doi.org/10.1016/j.ijproman.2011.06.002

Zipf, G. K. (1935). *The psycho-biology of language*. Houghton, Mifflin.