Advanced Review

Visual category learning

Jennifer J. Richler and Thomas J. Palmeri*

Visual categories group together different objects as the same kinds of thing. We review a selection of research on how visual categories are learned. We begin with a guide to visual category learning experiments, describing a space of common manipulations of objects, categories, and methods used in the category learning literature. We open with a guide to these details in part because throughout our review we highlight how methodological details can sometimes loom large in theoretical discussions of visual category learning, how variations in methodological details can significantly affect our understanding of visual category learning, and how manipulations of methodological details can affect how visual categories are learned. We review a number of core theories of visual category learning, specifically those theories instantiated as computational models, highlighting just some of the experimental results that help distinguish between competing models. We examine behavioral and neural evidence for single versus multiple representational systems for visual category learning. We briefly discuss how visual category learning influences visual perception, describing empirical and brain imaging results that show how learning to categorize objects can influence how those objects are represented and perceived. We close with work that can potentially impact translation, describing recent experiments that explicitly manipulate key methodological details of category learning procedures with the goal of optimizing visual category learning.

How to cite this article:
WIREs Cogn Sci 2014, 5:75–94. doi: 10.1002/wcs.1268

INTRODUCTION

The ability to recognize objects as *kinds of things* is a valuable survival skill. It lets us apply what we have learned about one thing and generalize that knowledge to other things of the same kind. For example, after learning the hard way that a particular mushroom is probably poisonous, it is highly adaptive to generalize that knowledge to other similar mushrooms than to have to learn the hard way every time a new mushroom is encountered. Once a thing is categorized as a kind of thing, we can use our previously acquired knowledge about other members of its category to determine the best course of action. Should we eat it or not?

*Correspondence to: thomas.j.palmeri@vanderbilt.edu

Department of Psychology, Vanderbilt University, Nashville, TN, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.

Humans categorize to dizzying degrees. Some categories are commonplace, like when we categorize chairs from tables, trees from shrubs, or cats from dogs. Some are remarkable, like when an expert mycologist rapidly categorizes an edible mushroom from its similar but potentially deadly cousins. The same thing can be categorized at different levels of abstraction. That object you are sitting on might be a piece of furniture, a chair, and a leather wing chair, of a particular make you have long forgotten and may have never known. That thing you spot along your hike is a fungus, a mushroom, a type of Morel, and specifically a *Morchella deliciosa*.

Indeed, any object can be categorized an infinite number of ways. A mushroom be categorized broadly as edible or poisonous, scientifically by its family, genus, or species, but also as *things found in a kitchen*, *things smaller than a breadbox*, or *things an assassin might use*. But here our focus is on categories that may be defined in some way by perceptual similarity, perceptual categories defined in some way by their perceptual features. While most of what we describe likely applies broadly to most perceptual categories, regardless of modality, our focus is on visual categories. In part this is simply because the bulk of the current literature focuses on visual categories. But in addition, it is not uncommon in the vision literature to focus on categories defined by visual features since mechanisms that apply to vision may not all apply to modalities like audition, olfaction, or somatosensation.

Visual categories span levels of abstraction. They bridge vision and cognition. And they give meaning to visual perception by connecting specific visual experiences with general knowledge. In this article, we review research on how visual categories are learned. The scientific literature on visual category learning is vast, spanning at least four decades of effort, with contributions from psychology, computer science, neuroscience, anthropology, and philosophy. We discuss just some of the computational, behavioral, and neural findings that have informed our understanding of visual category learning, touching on some classic research as well as some newer findings and directions. Our selection of research to review is informed by a particular theoretical viewpoint, and we indulge in highlighting some of our own research on visual category learning. We encourage readers to consider other reviews that have been framed by other theoretical views and have highlighted other research.[1–4]

Our review is organized into five main sections. *A Brief Guide to Visual Category Learning Experiments* describes the basic elements of category learning experiments, including the choice of objects, the psychological space of objects, how that space is divided into categories, the nature of those categories, how those categories are learned, and manipulations that can be used to influence learning. Naturally, this section introduces methodological details commonly used in behavioral and neural studies of visual category learning as a roadmap to readers of this literature. Furthermore, as we will see, these methodological details can have important empirical and theoretical implications. *Rules, Prototypes, and Exemplars* reviews the assumptions of some core theories and models, highlighting a handful of results that have helped distinguish between competing models. *Single Versus Multiple Systems for Visual Category Learning* considers evidence for independent representational systems for different kinds of category learning, focusing on critical methodological factors that significantly impact experimental findings and their theoretical implications. *How Visual Category Learning Influences Visual Representations* highlights how learning to categorize objects can influence how those objects are visually represented and perceived. *Optimizing Visual Category Learning* introduces recent work that reframes the question from *how are visual categories learned?* to *how are visual categories learned best?* We briefly close with some *Conclusions*.

A BRIEF GUIDE TO VISUAL CATEGORY LEARNING EXPERIMENTS

People know that some objects are *planes*, others are *trains*, and others are *automobiles*. In a visual category learning experiment, subjects learn that some objects are *mogs*, others are *bliks*, and others are *rabs*. The goal is to understand how real world object categories are learned and how object categories are represented by observing how experimental subjects learn novel object categories in the laboratory. Designing a visual category learning experiment involves selecting a set of objects to learn to categorize, characterizing the space in which those objects vary in terms of their physical or psychological similarity, defining how different objects are grouped together into categories, deciding how subjects will be taught those categories, and determining how to measure their knowledge of learned categories. Many of these methodological decisions are based on theoretical considerations, some are based on practical considerations, and others are based on mere convention.

Understanding the building blocks of visual category learning experiments serves as an introduction and overview of empirical work common to this area of cognition. Perhaps more importantly, because these are learning experiments, the methodological details of what is learned, how it is learned, and how learning is measured can have important theoretical implications as well. Sometimes seemingly inconsequential methodological details can prove critical to fully understanding how visual categories are learned. Conversely, certain methodological details sometimes loom large—perhaps too large—in theoretical discussions even when other equally critical details are downplayed or ignored. Finally, if someone has a goal to optimize category learning, then it is important to understand how various methodological details impact the speed, quality, and type of learning. These points will be revisited throughout our review.

We begin with the objects that subjects learn to categorize. Mirroring the multitude of objects and object categories in the real world, a multitude of objects and object categories have been used in visual category learning experiments, a small selection of which are illustrated in Figure 1. These range from simple stimuli based on elementary visual features

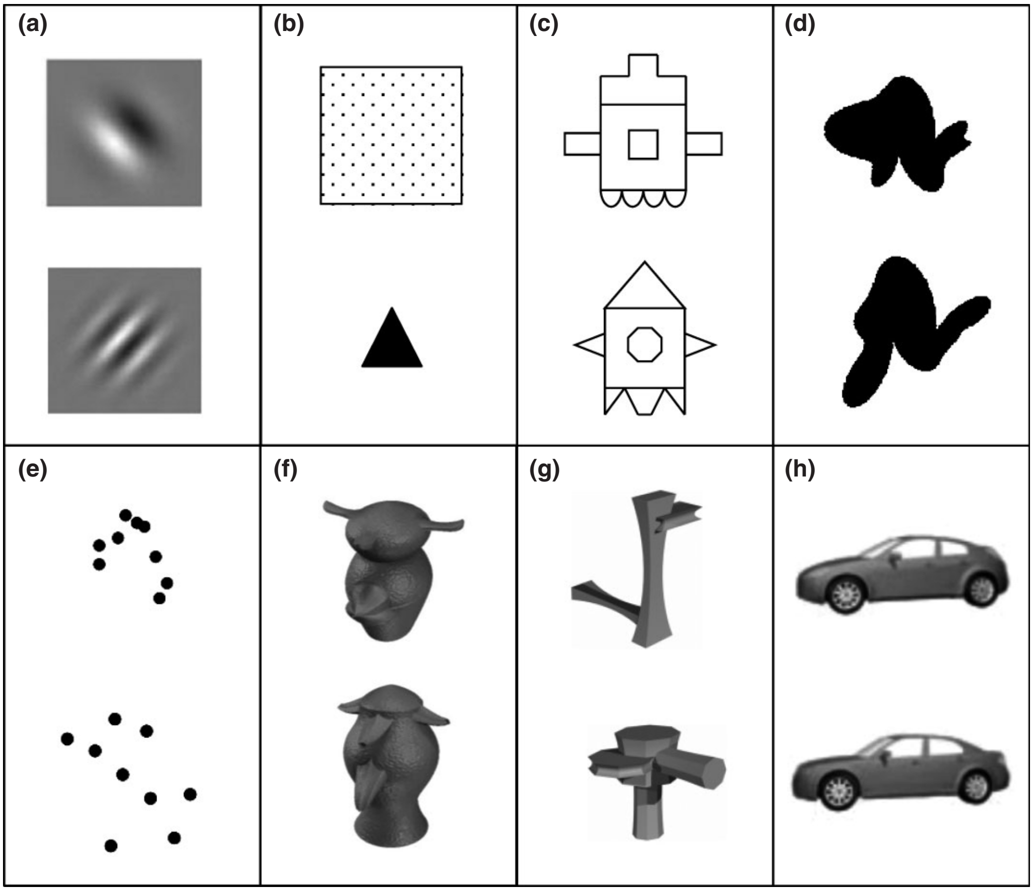

FIGURE 1 | Examples of objects used in visual category learning experiments: (a) Gabor patches varying in orientation and spatial frequency[5,6]; (b) simple objects varying in shape, size, and shading[7,8]; (c) simple line drawings of ships varying in the shape of the wings, porthole, tail, and nosecone[9,10]; (d) novel contour shapes defined by Fourier descriptors[11]; (e) random dot patterns[12–16]; (f) greebles that vary in body shape and the shapes of three appendages[17]; (g) ziggerins that vary in shape and style[18]; (h) novel cars created my morphing.[19,20]

(e.g., Gabor patches,[5,6] colors,[21] simple shapes[7,8]) to line drawings of objects with multiple parts (e.g., ships,[9] lamps,[22,23] creatures[24]), to novel objects (e.g., fourier-defined contours,[11] dot patterns,[12–14] greebles,[17] ziggerins[18]), and finally to known objects (e.g., animals,[25] cars[19,20]). All of these objects are composed of multiple dimensions or features, some of which are explicitly varied across different objects and categories, and others of which remain constant across all objects in an experiment.

While it is true that very abstract or ad hoc object categories can be defined by unseen semantic properties that group very dissimilar objects together into coherent categories,[2,26,27] visual category learning experiments most often use object categories defined by psychological similarity.[28,29] Different experiments manipulate similarity between objects in different ways, the collection of objects and their similarities defines a psychological space, and how that space is divided up defines object categories.

Objects can differ along discrete dimensions, such as large or small, black or white, circle or square.[8] Objects in experiments are often denoted using discrete-valued dimensions and abstract notation, such that a large, white, square would be object *1 2 2*, and a large, black, circle, would be object *1 1 1*. With binary dimensions, individual objects occupy the vertices of a multidimensional cube (e.g., Figure 2), but multiple nominal values along each dimension are also possible.[10,31]

Objects can also vary continuously (Figure 3). Dimensions of continuous variation can be based on psychophysics, such as angle and spatial frequency,[5,6] multidimensional psychological scaling (MDS), such as hue and saturation of colors,[21] or continuous physical variation, such as the length or location of object attributes.[34,35] Continuous dimensions can also be based on parameters having only an indirect relationship to object shape, such as Fourier descriptors,[11] or can be defined by various

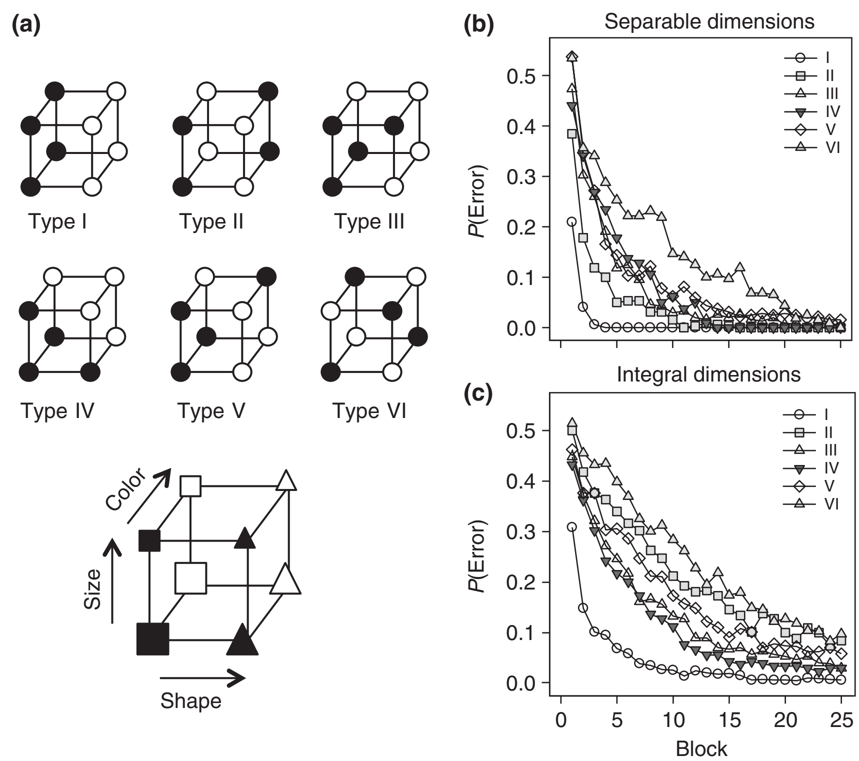

FIGURE 2 | (a) Six types of category structures tested by Shepard et al.[8]; see also Refs 7, 21. These types reflect the six possible ways of dividing up a set of stimuli defined along three binary-valued dimensions into categories of equal size (see also Ref 30). The three dimensions of the cube represent the three psychological dimensions, with each corner of the cube representing a possible object. For each type, the color coding (white or black) denotes members of one category versus the other category. Example of an assignment of logical dimensions to physical dimensions shown below. (b) Category learning data as *P*(Error) as a function of learning block for the six types when object dimensions are separable (shape, size, and color). (c) Category learning data as *P*(Error) as a function of learning block for the six types when object dimensions are integral (hue, saturation, and brightness of color patches). Note the different qualitative ordering of category learning difficulty depending on whether separable or integral dimension objects are used.

morphing techniques applied to real or novel objects.[19,20,36–38]

Objects can also vary continuously with respect to their similarity to a prototype that defines a particular category. In this case, while the objects are indeed multidimensional, the manipulations are ones of distortion with respect to a prototype object, not explicit variation along specific known psychological dimensions. A classic example in visual category learning is the random dot pattern (Figure 4) and its variants,[13–16,40] where the physical level of distortion of a prototype dot pattern creates new dot patterns with varying psychological dissimilarity to the prototype.

In principle, any given space of objects can be divided up into a large[30]—and for continuous spaces infinite—number of possible visual categories. Categories can be defined explicitly by a category boundary that separates members of one category from members of other categories (e.g., Figure 3(c)). Categories can also be defined by the members that make up that category (e.g., Figures 3(a) and (b), 4). In that case, we can distinguish between deterministic categories, where an object is a member of one and only one category, and probabilistic categories, where an object can sometimes be a member of one category and other times be a member of another category.

While it is common to test categorization performance using natural, real-world categories of objects,[41] when testing category learning it far more common to use novel categories and psychological spaces of objects that may appear artificial or arbitrary. But these categories and spaces are often explicitly designed to test predictions of particular theories. For example, one theory might predict that one kind of category is easier to learn than another, or that there will be more errors or slower responses learning some category members than others, while a competing theory might predict the opposite.

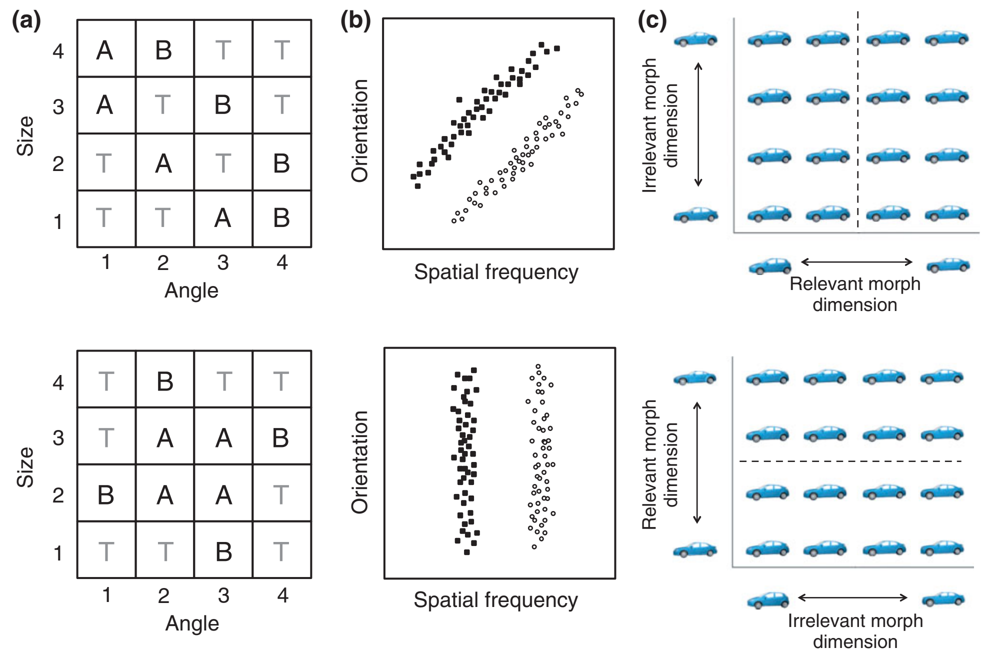

FIGURE 3 | Examples of psychological spaces of objects and category structures that are learned. (a) Objects are semicircles that vary in size (four levels) and the angle of a radial line drawn from the center to the edge (four levels).[32] Subjects learn objects as members of two categories (A and B) and are tested after category learning on the learned objects (As and Bs) and new transfer objects (Ts). The top category structure is linearly separable; the bottom category structure is not. (b) Objects are Gabor patches that vary in orientation and spatial frequency. Each point is an object in psychological space. Subjects learn objects as members of category A (black squares) or category B (open circles). Top shows an information integration (II) category structure that requires both dimensions. Bottom shows a rule-based (RB) category structure that requires only one dimension (in this case, spatial frequency). (c) Objects are morphs of cars. The two dimensions are themselves morphs between two parent cars (four different parent morphs in total). Each car in the morph space is a morph between a morph along the horizontal dimension and a morph along the vertical dimension. Two possible category structures are shown, one with a vertical category boundary defining horizontal morph dimension as relevant, and another with a horizontal category boundary defining the vertical morph dimension as relevant.[19,20]

Next, we consider different kinds of visual categories that might be learned. First, consider the case of objects with binary-valued dimensions, such as those used in the classic Shepard et al.[8] category learning experiment (see also Refs 7, 21). This experiment tested how well people could learn six different types of categorization problems (Figure 2), representing the six logically possible ways that eight objects defined along three binary-valued dimensions can be divided into two categories of four exemplars each (see also Ref 30 for a generalization of this approach). Categories can be well-defined verbally, such as a simple rule along one dimension (type I), or an exclusive-or rule along two dimensions (type II). Categories can also be more ill-defined, in the sense of being less amendable to description in terms of easily verbalizable logical rules. Type IV is ill-defined and linearly separable, in that although a plane perfectly separates members of one category from members of the other category, this plane cannot be described using am easily verbalizable rule along the perceptual dimension; in this way, it can also be characterized as a family resemblance structure, in that category exemplars are all similar to a particular prototype. The other types (III, V, and VI) are ill-defined and nonlinearly separable—no single plane can separate members of the two categories. The extreme case is type VI, where for every given object, the most similar objects are members of the opposite category. In addition to these classic category structures, many other structures with discrete dimensions have been created and used to test various competing hypotheses about how visual categories are learned and represented.[16,42,43]

In a continuous object space, analogous manipulations of category structures are possible (Figure 3). Categories can be defined by a single psychological dimension or multiple psychological dimensions. Categories may be amenable to a verbal description by logical rules or not, linearly separable or not, deterministic or probabilistic, defined by a fixed number of exemplars or by a continuous probability distribution of exemplars. As we noted earlier, while the experimental choice of how categories are defined

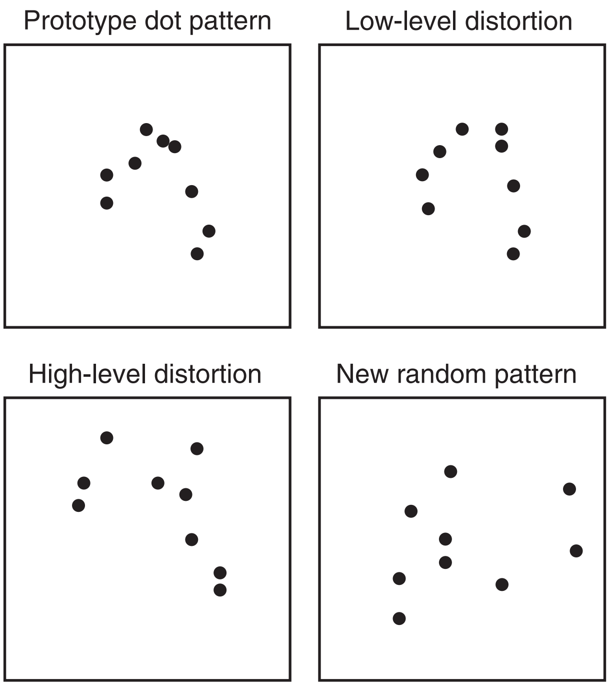

FIGURE 4 | Examples of a prototype dot pattern, low-level distortion of the prototype, high-level distortion of the prototype, and new unrelated random dot pattern used in dot pattern category learning studies.[13,39]

can sometimes appear rather arbitrary, it often depends on theory, as we will see in later sections of this review.

Now that we have considered the objects, psychological space of objects, and how that space of objects is divided into categories, how are those categories learned? In principle, an experimenter could teach a subject a collection of visual object categories in a wide variety of ways. If the categories can be described by verbal rules, even imperfect verbal rules, explicit instruction in those rules is one approach.[40,43–45] Alternatively, subjects could be shown all of the examples of a particular category or set of categories at the same time, allowed to study what is similar and different across the various category exemplars, and interpret what makes an object a member of a particular category.[24,46,47]

The most common experimental procedure is to show category exemplars one at a time (Figure 5). Typically, a randomly selected exemplar from a randomly selected category is displayed to the subject on each trial. Alternatively, the selection of categories and exemplars from a category could be biased if a goal is to test the effects of frequency on category learning,[53] or carefully controlled if a goal is to test the effects of staging of categories and exemplars on the speed, quality, and type of category learning.[54–56]

The particular details of each individual experimental trial of category learning can vary, sometimes as an intentional manipulation within an experiment, other times as a mere difference in convention across experimenters and laboratories. For example, each exemplar and its category label could be presented simultaneously,[57] or, each exemplar could be shown without a label and the subject could be asked to categorize it; subjects may[13,42] or may not[58] receive corrective feedback, that itself may vary.[5,39,49] In some cases learning may be entirely implicit, in the sense that subjects are neither provided explicit feedback nor told that the goal is to learn categories of objects.[59]

Finally, category learning can be measured in different ways. For example, experiments can focus on the number of trials required to reach some learning criterion, or measure decreases in errors[7,8,21,60–62] or response times[40,61] with learning; performance can be measured on all category learning trials,[7] or on transfer trials (with either trained or new exemplars) without feedback at the end of learning[42] or interspersed over the course of training.[9]

Clearly, discussing the full factorial combination of all of these experimental manipulations of objects, spaces, categories, and procedures over several decades of research would be impossible within any single review. In the following sections, we highlight just a small subset of empirical and theoretical results. In particular, we highlight studies that emphasize the importance of considering some of the methodological and procedural details of category learning experiments and how those details impact empirical results and their theoretical implications.

RULES, PROTOTYPES, AND EXEMPLARS

Categories are abstractions, with collections of different objects treated as the same kind of thing. Categorization is a form of abstraction. But does that imply that the mental representations and processes involved in categorization are inherently abstract and that category learning involves creating abstractions?

Abstraction defined much of the early research and debate about category learning.[2] Early accounts equated category learning with logical rule learning.[63,64] While rule-based accounts provided good descriptions of how people learned artificial categories defined by explicit logical rules, natural categories were argued to have a graded structure that suggested instead notions like 'family resemblance' and 'similarity.'[65,66] For example, it is easier to categorize a robin as a bird than an ostrich as a

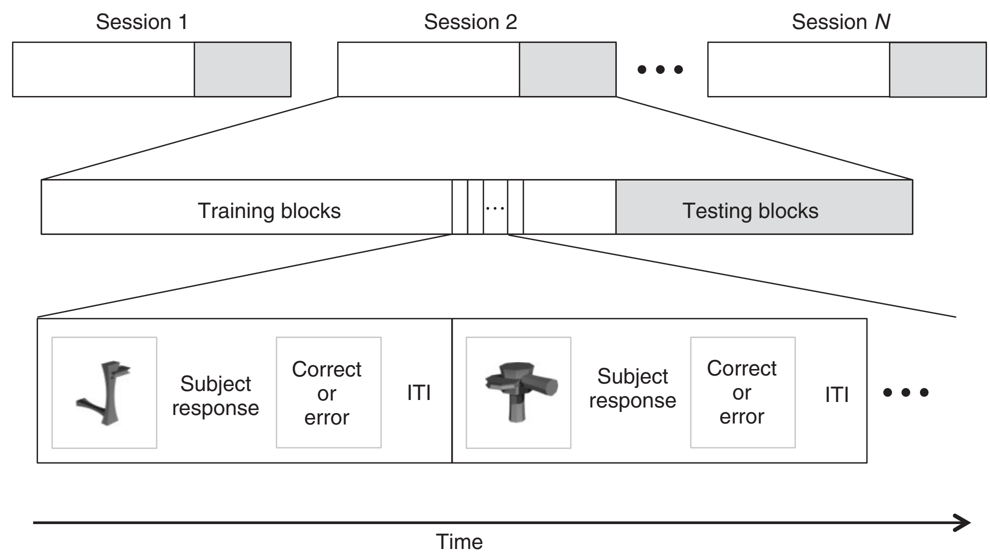

FIGURE 5 | Schematic illustration of a possible visual category learning experiment. Some category learning experiments are conducted in a single session (e.g., Refs 7, 42, 43) and some over multiple sessions (e.g., Refs 18, 40, 48). Most category learning experiments include explicit training blocks, where typically there is a set of objects that are used as training items and corrective category feedback is provided on every trial. These may be followed by test blocks that include training items as well as transfer items and no corrective feedback is provided (e.g., Refs 42, 43). Sometimes transfer blocks are also interspersed at various key time points over the course of category learning (e.g., Ref 9). A common structure of a single training trial is to present an object, either for unlimited time or for an experimentally determined amount of time, record the subject category response, and provide feedback, sometimes with the object still visible and sometimes not. As discussed in the text, the timing of the feedback and whether intervening tasks are used within a trial can be manipulated (e.g., Refs 5, 49). An intertrial interval (ITI) separates one trial from the next. The trial-by-trial contingencies can influence performance, as discussed in the text (e.g., Refs 50–52).

bird because a robin is more similar to the prototypical bird.[64] Experiments followed that examined how subjects learned novel categories defined by prototypes rather than rules, classic experiments using dot pattern stimuli being well known examples.[13,67] When subjects were trained on category exemplars that were distortions of a prototype, during transfer tests after learning, subjects categorized the prototype as well as, and often better than, the trained category exemplars, even though the prototype was never seen during learning, suggesting that a prototype had been abstracted from the experienced exemplars during category learning.

Computational models that formally describe theories using mathematical expressions and computer simulations led to new insights regarding the potential role of abstraction in category learning. Models of object recognition often hypothesize the stages of visual processing through the retina, lateral geniculate nucleus, and visual cortex[4,68,69]; these models often recognize and categorize images (see also Refs 70, 71). But many models of visual category learning simplify the details of visual processing stages by assuming that objects are represented as points in a multidimensional psychological space,[72] bypassing the details of early visual processing stages, to focus on testing competing hypotheses about how visual categories are learned and represented, and how categorization decisions are made. This multidimensional space can be defined by psychophysics, the physical properties of the objects, or techniques like multidimensional scaling.[28,73] Psychological similarity can then be defined as a function of the distance between objects in that psychological space,[28] as we outline below in a more formal manner.

Exemplar or instance-based models assume that a category is mentally represented in terms of the particular exemplars that have been experienced during learning. Exemplars stored in visual memory are activated according to their similarity to the object to be categorized. The degree to which an object is similar to objects from just a single category determines how well that object will be categorized. In this way, categorization behaviors that seem to suggest abstraction, such as the prototype effects observed in dot pattern category learning tasks, are well-predicted by exemplar models that assume no abstraction.[12,14,50,73,74] They also predict effects of specific exemplar similarity,[75] advantages learning some nonlinearly separable versus linearly separable categories,[76] and a host of other categorization phenomena inconsistent with prototype abstraction.[75]

These models have also been extended to jointly predict how response times and response probabilities vary according to factors like exemplar similarity and the amount of category learning.[61,77]

Exemplar models do not simply assume a passive registration of category exemplars in visual memory that are associated with learned categories. Instead, a key assumption of many exemplar models is that dimensions diagnostic of category membership will be weighted more heavily than less diagnostic dimensions when determining similarity.[33,60,78] For example, the generalized context model[33] assumes that the psychological similarity s_{ij} between object i and stored exemplar j is given by

$$S_{ij} = \exp\left(-c \cdot d_{ij}\right), \quad (1)$$

where the distance d_{ij} is given by

$$d_{ij} = w_m \left[\sum_m \left| i_m - j_m \right|^r \right]^{1/r}, \quad (2)$$

where m indexes the dimensions of the multidimensional psychological space, i_m and j_m are the values of object i and stored exemplar j along dimension m, and r reflects the metric used to calculate distance. Key are the w_m parameters that weight each dimension m according to its diagnosticity for the learned categories. So even if an object and stored exemplars differ along highly salient dimensions, if those dimensions are nondiagnostic of category membership—if their weights w_m are close to zero—the differences among those dimensions will not contribute to psychological similarity s_{ij} calculated for the purposes of deciding category membership.

Consider the six category learning problems from Shepard et al.[8] illustrated in Figure 2. Prototype abstraction would predict better and faster learning for Types I and IV than the other category problems because these are both linearly separable categories, amendable to the formation of a single prototype for each category. An exemplar model without any selective attention to diagnostic dimensions would predict an order of I < IV, III, V < II < VI based on overall similarity to exemplars.[8,79] In fact, the empirically observed ordering of category learning difficulty is I < II < IV, III, V < VI.[7,8] This is precisely the ordering that is predicted by the generalized context model when optimal dimensional weights based on category diagnosticity are assumed[79] (see also Ref 80), and when dimensional weights are learned by error correction[7,60] (see also Ref 81).

This ordering of category learning problem difficulty does depend on choosing objects and psychological spaces where diagnostic dimensions can be selectively weighted, like size and shape; for example, it is easy to discriminate triangles from squares regardless of their size, and large from small objects regardless of their shape. However, it is difficult to selectively attend to other psychological dimensions because they are not perceived independently; for example, the perceived richness of a red patch (its saturation) influences how much that patch is perceived as the color *red* (its hue).[82] When dimensions are defined by psychological dimensions like these, dimensions that are difficult to selectively weight, the ordering predicted by an exemplar model without any selective attention is observed, I < IV, III, V < II < VI.[21] We revisit the theoretical importance of selective attention to diagnostic dimensions again later in this review.

SINGLE VERSUS MULTIPLE SYSTEMS FOR VISUAL CATEGORY LEARNING

More recent research on category learning has considered more complicated models that jointly consider abstract and specific category representations. For example, one class of models considers category learning as cluster formation.[83,84] These models allow for a continuum of category representations, including large clusters akin to category prototypes, small clusters akin to specific exemplars, or narrow clusters akin to rules, depending on the nature of the categorization problem and the values of parameters that control cluster formation. In some cases, when these models are fit to observed category learning data, the parameters that control cluster size converge on category learning computations that are mathematically similar to exemplar models.[85] More generally, we can consider a representational continuum from models that abstract category representations[86] to those that assume specific category representations.[75,87] When we consider the breadth of models along this continuum, category learning behavior often falls within the range predicted by models either assuming specific category representations like exemplars, or those assuming relatively small levels of abstraction that make predictions largely indistinguishable from exemplar models.[60,87–91]

Other category learning models have considered mixtures of abstract representations, like rules or prototypes, with more specific representations, like exemplars. The rule-plus-exception (RULEX) model of category learning[43,62,90] hypothesized that people try to explicitly learn categories by forming simple rules, often along a single dimension. For ill-defined categories, these rules are imperfect, so RULEX hypothesized that people remember exceptions to those rules (hence rule-plus-exception), and may remember some aspects of the other category exemplars in visual memory.[43] Despite its focus on abstraction, RULEX predicts many of the same fundamental category learning phenomena as exemplar models,[90] including, for example, the ordering of category learning problems from Shepard et al.,[8] described earlier.

Key to distinguishing predictions of RULEX from those of exemplar models was reconsidering how category learning performance was measured. It is common to characterize the probability with which individual test objects are assigned to particular categories by averaging responses across subjects in an experiment. Of course, the potential for averaging artifacts, where the average performance across subjects may not characterize the behavior of any individual subject, are well known.[92] One of the key assumptions of RULEX is that the formation of rules and exceptions is idiosyncractic, governed by a stochastic learning process (an assumption true for RULEX, but not necessarily a property of other models in its class). One subject might learn a rule along one psychological dimension and remember its exceptions, while another subject might learn a rule along a completely different psychological dimension and remember its exceptions. Consequently, generalization to test objects during transfer could be quite different for different subjects, and these differences could be completely obscured by averaging. Indeed, with modest amounts of category learning, patterns of generalization observed at the individual subject level were consistent with the formation of simple rules.[90] Over the course of category learning, however, patterns of generalization shifted from those consistent with the application of rules to those consistent with similarity to exemplars[9] (see also Ref 43). These qualitative changes as a function of category learning are only observed when behavior is measured as individual subject generalization patterns, not average categorization responses. Details of how behavioral data are measured and analyzed matter.

The initial theoretical suggestion of a hybrid model like RULEX was followed by a number of other hybrid accounts of category learning. For example, Palmeri[40] taught subjects an explicit rule to categorize dot pattern stimuli by their numerosity, but over multiple sessions with those patterns a shift to categorization based on exemplar similarity instead of the rule was observed. Changes in response times with learning and generalization to new dot patterns was well-predicted by a model that assumed a race[93] between the explicit counting rule and the exemplar-based random walk model.[61] Other models similarly propose combinations of rule-based and similarity-based representations.[34,94]

There continues to be ongoing debate about the nature of different forms of category representations, how they are learned, and where they are represented and processed in the brain. For example, some have characterized rule-based representations versus more similarity-based or implicit representations as candidates for independent memory systems,[1,95] others have characterized interactions where exemplar-based representations gate application of rule-based representations,[34] while others have suggested a single system account.[50,96] To be clear, by multiple memory systems researchers typically mean multiple *representational* systems, say one for representing learned rules, one for representing remembered exemplars, and so forth. Category learning clearly involves many different functional components in addition to learning, storing, and retrieving representations, like attending to features and dimensions, processing feedback, or engaging executive control. A multiple memory systems view suggests that separate and independent systems are engaged for learning different kinds of category representations. According to a single system view, learned category representations may be shared, but the different functional systems that perform specific computations, like processing feedback or attending to features and the like, may be differentially recruited depending on the categorization problem to be learned. For example, while Johansen and Palmeri[9] observed shifts from rule-based to exemplar-based patterns of generalization with category learning, they accounted for behavior with a single-system exemplar model that allowed qualitatively different patterns of dimension weights to emerge for different simulated subjects. They suggested that explicit top-down executive control[97] might be responsible for different patterns of dimension weights established early in category learning. This executive could indeed be engaged in hypothesis testing, but it manifests those hypotheses in terms of parameter settings in the exemplar model, not as an independent rule-based category representation that is functionally independent of exemplar-based category representations.

One source of evidence for independent memory systems comes from a series of studies reporting behavioral dissociations using rule-based versus information–integration categories (Figure 3(b)). In these experiments, categories are defined by multivariate normal distributions with category exemplars that are random samples from each distribution.[98] Not surprisingly, unidimensional rule-based categories vary along a single dimension. More generally, rule-based categories can be verbalized, with independent decisions about particular values along a subset of dimensions (see also Ref 45). By contrast, information integration categories cannot be easily verbalized and categorization responses require combining values along two or more dimensions. The COVIS model[94] proposes a COmpetition between a Verbal and Implicit System, and predicts that different procedural manipulations will differentially affect rule-based versus information-integration category learning.

For example, according to COVIS, if information-integration category learning relies on a procedural learning system. Given its assumption that the procedural learning system is tightly linked with the motor system, swapping the response keys associated with different categories after initial learning should significantly interfere with performance on information-integration categories. By contrast, performance on rule-based categories, which do not depend on procedural learning according to COVIS, should be relatively unaffected. That is precisely what Ashby et al.[99] found. However, as noted by Nosofsky et al.,[100] and as acknowledged by Ashby et al.,[99] in addition to the obvious structural difference between information-integration and rule-based categories, they also differ considerably in difficulty. When these differences in difficulty were carefully considered, Nosofsky et al. observed analogous interference effects for rule-based categorization as Ashby et al. observed for information-integration categorization; Nosofsky et al. also observed interference effects when more difficult explicit rules were learned. When the methodological confounds were controlled, no dissociation was observed.

Analogously, according to COVIS, rule-based category learning requires working memory and attention to process and consider verbal rules, while information–integration category learning does not. Maddox et al.[6] had subjects learn a unidimensional rule-based category or an information-integration category. During learning, after making a categorization response to each stimulus, subjects completed a memory-scanning task either immediately or 2500 milliseconds after receiving corrective feedback. Maddox et al. observed interference from the memory-scanning task only when it was introduced immediately after feedback, and only for rule-based category learning. This dissociation suggests a different processing locus for rule-based and information–integration category learning. However, Stanton and Nosofsky[101] noted that while Maddox et al. aimed to equate the difficulty of rule-based and information-based category learning, they did so by making many of the stimuli in the rule-based category learning condition almost perceptually indistinguishable from one another; while difficulty was equated in terms of overall accuracy, difficulty of information–integration categories stemmed from the complexity of the categorization problem while difficulty of rule-based categories was perceptual. When Stanton and Nosofsky equated difficulty of categorization without making stimuli perceptually indistinguishable, they failed to observe any differential interference based on whether the memory scanning task was immediate or delayed (see also Ref 102 for another example of how equating difficulty eliminates dissociations predicted by multiple memory systems models). Moreover, when Stanton and Nosofsky made stimuli in an information-integration category task perceptually difficult to discriminate, like the rule-based category items in Maddox et al., they observed interference for immediate compared to delayed memory scanning. These behavioral dissociations appear far less stable than would be predicted by an independent memory systems model like COVIS, and they can be accounted for by other aspects of the experimental design that are not obviously related to COVIS.

Now consider a provocative finding from Filoteo et al.[103] They reported that subjects who learned a complex three-dimensional information–integration category learning task performed better when they also engaged in a secondary memory scanning task than subject who did not. They explained this seemingly counterintuitive result—phrased as a category learning advantage that comes from ‘removing the frontal lobes’—using COVIS. They suggested that occupying working memory with the memory scanning task renders the rule-based category learning system largely unavailable, such that category learning must rely solely on the procedural learning system that is optimally suited for learning difficult information-integration categories. However, Newell et al.[104] highlighted a problem with their experimental procedures, namely that subjects given the memory scanning task were also inexplicably given an additional 2500ms with which to process the corrective category feedback. Is this a counterintuitive advantage that comes from a concurrent memory scanning task, or a less surprising advantage that comes from simply having more time available to process feedback? Newell et al. found evidence consistent with the latter: subjects given more time to process feedback performed better, irrespective of whether they engaged in a concurrent memory scanning task or not. Procedural details of the category learning tasks are important.

Other work has dissociated exemplar learning from prototype abstraction as independent memory systems, the former considered a form of explicit declarative memory and the latter considered a form of implicit procedural learning.[105] One well-known demonstration of a behavioral dissociation tested amnesic individuals and controls on category learning and recognition memory with a variant of the dot pattern task.[39,106] After studying a small set of random dot patterns, amnesic individuals were significantly impaired compared to controls at discriminating studied dot patterns from new dot patterns during a recognition memory test. However, after studying a set of distortions of a prototype dot pattern, there was no difference between amnesic individuals and controls in their ability to categorize members versus nonmembers at test. By itself, a significantly larger impairment at recognition memory compared to categorization is not necessarily inconsistent with predictions of exemplar models.[107,108] However, observing chance recognition memory with above chance categorization[106] is more theoretically challenging.[50]

However, it is important to consider procedural details of the dot pattern tasks used in these particular experiments. For the recognition memory task, individuals studied five entirely random dot patterns presented eight times each. Then they were tested on those five studied patterns and five new random patterns. Without memory for studied dot patterns, there is simply no way to discriminate old from new patterns. For the categorization task, individuals studied forty high-level distortions of a single category prototype. At test, they were shown four repetitions of the previously unseen prototype, 20 low-level distortions, 20 high-level distortions, and 40 completely new random patterns, and were asked to discriminate studied category members from nonmembers. While no feedback was provided in the categorization test, it is important to note that during the test subjects were shown many dot patterns that were very similar to one another and other dot patterns that were completely dissimilar to one another (see Figure 4). Is memory even needed to discriminate members from nonmembers at test given the highly structured nature of the test itself?

Palmeri and Flanery[50,109] conducted experiments where subjects were told that they had been subliminally shown a series of dot patterns during an initial unrelated task at the start of the experimental session. In fact, they had never seen any dot patterns whatsoever. Subjects were then tested on either the recognition memory or categorization dot pattern task. Not surprisingly, subjects were completely at chance at recognizing 'old' dot patterns from new dot patterns; without memory (since no dot patterns had ever been seen before) there was no way to make that discrimination. However, even without ever having seen any previous distortions of the prototype, subjects were significantly above chance at categorization. In fact, these subjects who studied nothing performed just as well as subjects who had actually studied dot patterns earlier. Subjects had learned during the test, and Palmeri and Flanery surmised that this learning could be supported by working memory, which is spared in many cases of amnesia. The potential for learning during test has been since shown to play a key role in several other experiments purporting to show dissociations between categorization and recognition memory[50,110,111] and prototype abstraction more generally.[51,112] Seemingly inconsequential procedural details, such as the structure of a categorization test, can have profound consequences for how we interpret behavioral and neuropsychological dissociations.

Debates about single versus multiple memory systems in category learning are part of larger theoretical debates about brain organization. Some have suggested that there are independent neural systems for learning rule-based, information-integration, prototype-defined, and probabilistic categories.[1,95] Single-system accounts instead suggest that different neural systems are responsible for different computations necessary for categorization, with some neural systems recruited more or less for certain category learning problems than others. Unfortunately, the 'single' in 'single system' is often unfairly portrayed as some kind of cartoon, something akin to Lashley's equipotentiality,[113] when in fact single system models consistent of multiple components, each of which is likely mapped onto different neural mechanisms. Exemplar models like ALCOVE[60] or EBRW[61] assume perceptual processing, perceptual memories, selective weighting based on diagnosticity, learned associations between exemplars and categories, and categorization decision mechanisms, all of which can be subject to top-down executive control (see Refs 9, 50). Emerging fMRI evidence seems consistent with this single system, multiple mechanism view of category learning.[114–117] The vigorous debates about single versus multiple systems in category learning mirror the vigorous debates about modularity in visual object recognition,[118] where there the debate is not about tasks but whether different kinds of objects recruit different neural mechanisms in an all-or-none (multiple systems) or graded (single system) manner.[119]

HOW VISUAL CATEGORY LEARNING INFLUENCES VISUAL REPRESENTATIONS

How does visual category learning influence visual representations of objects belonging to those learned categories? We know from the large body of work in categorical perception (recently reviewed by Goldstone and Hendrickson[120]) that pairs of stimuli that vary by a particular physical increment can appear quite similar or quite different depending on whether the stimuli in that pair belong to the same category or different categories. Two different shades of red appear more similar than a shade of red versus a shade of orange, even if the physical difference is the same. We know that categorical perception is not merely a fixed biological constraint on perception since the category boundaries separating colors depend on how the acquired language of the observer parses those colors into different category names. Moreover, categorical perception effects have been created in category learning experiments using nonspeech auditory stimuli,[121] novel color categories,[122] and novel face categories.[123] But beyond elementary psychophysical dimensions or potentially special perceptual categories like speech sounds or faces, there is continued debate about how visual category learning influences representations of objects, behaviorally and neurally.

Folstein et al.[19] recently demonstrated an effect of category learning on perceptual discrimination using novel cars and novel car categories. Adapting a technique used previously with faces,[37] they first created a two-dimensional morph space of cars from a set of morph parents. Adapting a technique used previously with colors,[122] they next had subjects learn to categorize these novel cars into one of two novel car categories; the categories were simply defined by a linear category boundary that separated the morph space into two regions. In a subsequent discrimination test, subjects showed significantly better perceptual discrimination for cars that differed along the relevant dimension (parallel to the category boundary) than those that differed along the irrelevant dimension (orthogonal to the category boundary), demonstrating an effect of category learning on visual discrimination performance.

What is the neural locus of these perceptual changes[118,124]? According to some theories, object representations in visual cortex are not systematically influenced by the way objects have been learned to be categorized.[69] Indeed, Jiang et al.[38] found neural modulation according to learned category in frontal areas but not in visual areas of the brain (see also Ref 36). Like Folstein et al.,[19] Jiang et al.[38] had subjects learn to categorize objects in a morph space. Critically, unlike Folstein et al., subjects in Jiang et al. were not better at discriminating objects depending on whether the difference between objects was along a relevant versus irrelevant dimension. Why the difference?

While both studies used objects in morph spaces, Folstein et al.[19] demonstrated that the experimental details of how those morphspaces are created matter a lot. Without explicating those details here, when morphspaces were created by factorially combining morphlines,[37] enhanced perceptual discrimination along category-relevant object dimensions was observed. By contrast, when morphspaces were created by blending different morphparents together, as in Jiang et al., no enhanced perceptual discrimination according to learned category was observed. In a companion piece, Folstein et al.[20] further showed, using their factorial morphspace of objects, neural evidence via fMRI for changes in perceptual representations of objects as a result of category learning within visual cortex. It is clear that seemingly minor experimental details, such as the way a morph space is created, can have profound effects on behavioral and neural data, which in turn can have profound consequences for our theoretical understanding of neural plasticity as a consequence of category learning.

A related source of debate concerns how visual category learning influences the organization of visual object representations across inferotemporal cortex. According to modular accounts, regions of the visual cortex become specialized for particular kinds of visual entities, including faces, places, and body parts.[125] According to a competing view, regions of cortex are recruited depending on the representations and processes necessary to learn particular kinds of object categories, not according to the categories themselves.[119,126] By this view, faces recruit the fusiform face areas (FFA) not because they are faces, but because the FFA supports representations and processes necessary for fine-grained identification. Perceptual expertise demands fine-grained categorization, so bird experts recruit FFA to identify birds, car experts to identify cars, and greebles experts (a laboratory-induced form of expertise) to identify greebles.[48,127] Moreover, bird, car, and greebles experts display a range of behaviors qualitatively similar to face recognition (see Ref 128).

To more directly test the hypothesis that it is the type of category learning experience that creates behaviors and patterns of brain activity similar to face recognition, Wong and colleagues[18,129] had two groups of subjects learn to categorize the same set of novel objects (called ziggerins) in two very different ways. One aimed to capture aspects of face recognition. This group completed individuation training, where they learned unique names for many individual ziggerins. Another aimed to capture key aspects of letter recognition. This group completed basic-level categorization training, where they learned to categorize ziggerins into broad categories defined by overall object shape and to scan for particular exemplars of ziggerins in an array of ziggerin 'text'. The two groups were trained for the same amount of time, with the same number of objects and categories, but they differed in their learning goals and the tasks used to achieve those goals. Only subjects who learned to individuate ziggerins, like people learn to individuate faces, showed behavioral markers of face recognition[18] (see also Ref 130). In addition, while subjects who learned to individuate ziggerins showed an increase in face-selective regions of visual cortex when viewing new untrained ziggerins, those who learned to categorize and scan ziggerins showed more distributed brain activity, with specific increases in medial areas of occipital cortex,[129] not dissimilar to activity observed for letter recognition. Given the same set of objects, the procedural details of visual category learning can have significant effects on behavior and brain activity.

A central message of much of the work described thus far is that procedural details of category learning experiments can matter, whether it be the nature of the stimuli, how categories are defined, how people are trained, or how categorization performance is measured or manipulated. This should not be interpreted as suggesting that incisive questions about category learning cannot be answered with any degree of certainty. Rather, it is to combat an unsettling rhetorical strategy of claiming that particular theoretical or empirical questions have already been settled and that the field has moved on. In fact, many of these 'settled' matters, such as the existence of independent memory systems for category learning, are often based on a limited set of experimental results, finely tailored and tuned to observe an effect with limited generalizability, an effect that can often be explained by far simpler principles. Questions concerning the nature of abstraction in category learning, or whether there are independent representational systems supporting different forms of category learning, or whether the visual system plays any role in category learning are so fundamental and foundational that they demand the kind of detailed empirical investigation illustrated by the work described here.

In the final section, we briefly review recent work that represents a converse of the work described so far, work that aims to manipulate procedural details in order to optimize visual category learning.

OPTIMIZING VISUAL CATEGORY LEARNING

The vast majority of research on visual category learning has focused on how categories are learned and represented. In addition to theoretical questions regarding mechanism, we can also ask more practical questions about how to optimize category learning. Given a particular visual category learning problem, what can be done procedurally to learn to categorize objects more quickly, retain that category knowledge for longer, and generalize that knowledge to new objects most appropriately? The answers to these questions need not be the same, of course. One procedural variation might lead to rapid learning but less robust learning, while a different procedural variation might cause slower learning but more enduring learning.

Some recent work has examined how manipulating the presentation of category learning trials may optimize learning. Giguère and Love[55] considered how people learn categories with overlapping distributions of category members. Because the distributions overlap, categorization is probabilistic, in the sense that the same object is sometimes associated with one category and other times associated with another category. As a result, these category learning tasks can be quite difficult, since corrective feedback can seem entirely conflicting at times. Giguère and Love found that training on idealized distributions of category items, essentially removing the probabilistic aspects entirely, led to better categorization of trained and untrained exemplars than training using the true probabilistic structure (but see Ref 131). In that sense, idealized training can be considered optimal. However, the authors did suggest that this idealized training could lead to unwanted consequences, such as overconfidence on borderline cases. Overconfidence might have little consequence for a behavioral experiment, but could be problematic in real-world situations. For a physician deciding whether a skin condition is a serious one that requires a painful treatment or is a benign one that will clear up on its own, well-calibrated confidence in categorization is critical.

It has also been shown that interleaving items from different categories during observational learning leads to better categorization of both studied and new exemplars at test compared with massed category presentations.[132–136] These results highlight the known benefits of contrast in category learning, where direct contrast between members of different categories highlights category differences[137] (see also Refs 133, 138). Indeed, disrupting the ability to contrast sequentially presented items by introducing a secondary task between study items actually leads to a disadvantage in learning for spaced items.[132]

The benefit of drawing attention to distinctions between categories is also maximized in ‘fading’ procedures, where perceptual distinctions are initially exaggerated during learning. Pashler and Mozer[56] demonstrated that such procedures led to a learning advantage (better generalization to new items) when the relevant—faded—dimension was embedded among task-irrelevant dimensions that also varied. The critical factor was whether subjects were required to determine which feature was relevant for category learning. In other words, exaggerating the relevant dimension leads to a learning advantage because it helps subjects discover what dimension to attend to (see Ref 139). Consequently, fading did not help acquisition when the discrimination relied on a single perceptual dimension that could be well-specified to the learner either because stimuli themselves were one-dimensional or because subjects were explicitly told which dimension was relevant.

However, benefits of contrast in category learning seem dependent on other aspects of the learning context. For example, contrast may be less useful for category learning under conditions where determining similarity within a category is more critical than determining how categories differ, such as when all the items being categorized are highly dissimilar and items from the same category share very few features that are not easy to identify. Empirical work supports this intuition,[140,141] showing that, for example, interleaving exemplars from different categories is advantageous when categories are very perceptually similar, whereas massing is more effective for high-discriminability categories where all items are more different from each other. Thus, category structure and the kind of learning task can modulate whether interleaving, which promotes learning from contrast, or massing, which promotes learning from similarity, is a preferable manipulation to optimize category learning.

It is important to realize that what is considered optimal can depend on how categorization performance is being measured. For example, recall that many category learning experiments teach categories to subjects in a trial-by-trial manner, presenting a randomly sampled object, asking subjects to label that object as a member of a particular category, and then providing corrective feedback. Categorization performance is often measured as accuracy on those category learning trials. Stewart et al.[52] reported that subjects can strategically optimize their performance by capitalizing on this trial-by-trial structure, using their memory for the previously presented object and the corrective feedback provided to inform categorization on the current trial. Using this strategy, subjects were better at categorizing difficult objects near the category boundary when those objects were immediately preceded by a trial with a distant object from the opposite category, compared to a trial with a distant object from the same category. A significant amount of categorization performance in the lab is based on short-term strategies available because of the structure of the category learning experiment, and may not be based on long-term representations of categories in visual memory; this is analogous to the arguments made by Palmeri and Flanery[50,109] regarding categorization tests described earlier.

Palmeri and Flanery[51] explored the theoretical consequences of this trial-by-trial strategy on category learning experiments with only two categories, which is neither necessary nor universal, but extremely common in the literature. Imagine categorization decisions on the current trial of a category learning experiment are based simply on how similar the object on the current trial is to the object on the previous trial. If they are sufficiently similar, subjects categorize the object on the current trial into the same category as the object on the previous trial. Otherwise, subjects categorize it into the other category. Palmeri and Flanery simulated the use of this simple strategy on a series of category learning problems examined by Smith and colleagues.[142,143] For every category learning problem where the simple strategy simulated by Palmeri and Flanery produced above-chance categorization performance, Smith and colleagues had reported evidence for prototype abstraction. Moreover, data simulated using this simple strategy was better fit by a prototype model than an exemplar model. Clearly, this simple strategy relies only on short-term memory for the previous trial, not long-term memory for enduring category representations or prototypes. Thus, some measures of categorization performance may reflect online strategies, not long-term category representations, which is a cautionary note for any work considering how to optimize category learning. Excellent performance during a category learning task where online strategies are possible does little good when knowledge needs to be used later in the real world to categorize an isolated object that is not presented within an experimental trial structure.

CONCLUSIONS

Visual category learning connects specific perceptual experience with abstract conceptual knowledge. Understanding this fundamental aspect of visual cognition means weaving together an understanding of visual perception, visual memory, visual knowledge, and decision making, integrating behavioral, theoretical, and cognitive neuroscience research.[144] Our review has only been able to sample some of this research, highlighting some of the key findings, theories, and controversies.

We have intentionally organized much of our discussion in this review around methodological aspects of category learning experiments. There is growing recognition in psychology and neuroscience that inappropriate statistical tests, questionable statistical practices, and low statistical power can impede scientific progress.[145–148] Similarly, methodological choices, limitations, and confounds can profoundly impact empirical results and their theoretical implications in behavioral and brain imaging research on visual category learning, as reviewed here, as well as other aspects of visual categorization and recognition.[70,149–152] Our message here is not that it is impossible to draw conclusions from category learning experiments because seemingly inconsequential methodological details may matter. Rather, we emphasize the importance of paying careful attention to all procedural details, even those that may on first blush appear tangential or inconsequential. Experimental work that explicitly tests the impact of these procedural details is invaluable. Moreover, it appears that highly modular theories of category learning, those that assume independent memory systems for learning rule-based versus prototype-based versus exemplar-based categories, appear to be empirically and theoretically rather brittle notions. As we have shown throughout our review, predictions of multiple systems theories are often falsified with relatively minor changes to experimental procedures. And finally, new findings suggest that systematic modification of experimental procedures can profoundly affect how quickly and how robustly people learn new visual categories, promising new avenues for optimizing visual category learning.

ACKNOWLEDGMENT

This work was supported by NSF grants SMA-1041755 (Temporal Dynamics of Learning Center) and BCS-1257098.

REFERENCES

1. Ashby FG, Maddox WT. Human category learning. *Annu Rev Psychol* 2005, 56:149–178.
2. Murphy GL. *The Big Book of Concepts*. Cambridge, MA: MIT Press; 2002.
3. Pothos EM, Wills AJ. *Formal Approaches in Categorization*. Cambridge: Cambridge University Press; 2011.
4. Serre T. Models of visual categorization. *WIREs Cogn Sci* 2013. Submitted for publication.
5. Maddox WT, Ashby FG, Bohil CJ. Delayed feedback effects on rule-based and information-integration category learning. *J Exp Psychol Learn Mem Cogn* 2003, 29:650–662.
6. Maddox WT, Ashby FG, Ing AD, Pickering AD. Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Mem Cogn* 2004, 32:582–591.
7. Nosofsky RM, Gluck M, Palmeri TJ, McKinley SC, Glauthier P. Comparing models of rule-based classification learning: a replication and extension of Shepard, Hovland, and Jenkins (1961). *Mem Cogn* 1994, 22:352–369.
8. Shepard RN, Hovland CI, Jenkins HM. Learning and memorization of classifications. *Psychol Monograph Gen Appl* 1961, 75:1–42.
9. Johansen MK, Palmeri TJ. Are there representational shifts during category learning? *Cogn Psychol* 2002, 45:482–553.
10. Palmeri TJ. Learning categories at different hierarchical levels: a comparison of category learning models. *Psychon Bull Rev* 1999, 6:495–503.
11. Op de Beeck H, Wagemans J, Vogels R. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci* 2001, 4:1244–1252.
12. Palmeri TJ, Nosofsky RM. Central tendencies, extreme points, and prototype enhancement effects in ill-defined perceptual categorization. *Q J Exp Psychol* 2001, 54:197–235.

13. Posner MI, Keele SW. On the genesis of abstract ideas. *J Exp Psychol* 1968, 77:353–363.
14. Shin HJ, Nosofsky RM. Similarity-scaling studies of dot-pattern classification and recognition. *J Exp Psychol Gen* 1992, 121:278–304.
15. Homa D. On the nature of categories. *Psychol Learn Motiv* 1984, 18:49–94.
16. Smith JD, Minda JP. Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning. *J Exp Psychol Learn Mem Cogn* 2002, 28:800–811.
17. Gauthier I, Williams P, Tarr MJ, Tanaka J. Training 'greeble' experts: a framework for studying expert object recognition processes. *Vis Res* 1998, 38:2401–2428.
18. Wong AC-N, Palmeri TJ, Gauthier I. Conditions for face-like expertise with objects: Becoming a Ziggerin expert—but which type? *Psychol Sci* 2009, 20:1108–1117.
19. Folstein J, Gauthier I, Palmeri TJ. How category learning affects object perception: not all morph spaces stretch alike. *J Exp Psychol Learn Mem Cogn* 2012, 38:807–820.
20. Folstein JR, Palmeri TJ, Gauthier I. Category learning increases discriminability of relevant object dimensions in visual cortex. *Cereb Cortex* 2013, 23:814–823.
21. Nosofsky RM, Palmeri TJ. Learning to classify integral-dimension stimuli. *Psych Bull Rev* 1996, 3:222–226.
22. Lamberts K. Categorization under time pressure. *J Exp Psychol Gen* 1995, 124:161–180.
23. Lamberts K, Brockdorff N. Fast categorization of stimuli with multi-valued dimensions. *Mem Cogn* 1997, 25:296–304.
24. Regehr G, Brooks LR. Category organization in free classification: the organizing effect of an array of stimuli. *J Exp Psych* 1995, 21:347.
25. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 2001, 291:312–316.
26. Barsalou LW. Ad hoc categories. *Mem Cogn* 1983, 11:211–227.
27. Malt BC. Category coherence in cross-cultural perspective. *Cogn Psychol* 1995, 29:85–148.
28. Shepard RN. Toward a universal law of generalization for psychological science. *Science* 1987, 237:1317–1323.
29. Tversky A. Features of similarity. *Psychol Rev* 1977, 84:327–352.
30. Feldman J. Minimization of Boolean complexity in human concept learning. *Nature* 2000, 407:630–633.
31. Folstein JR, Gauthier I, Palmeri TJ. Mere exposure alters category learning of novel objects. *Front Cogn Sci* 2010, 1:40.
32. Shepard RN. Attention and the metric structure of the stimulus space. *J Math Psychol* 1964, 1:54–87.
33. Lamberts K. The time course of categorization. *J Exp Psychol Learn Mem Cogn* 1998, 24:695–711.
34. Erickson MA, Kruschke JK. Rules and exemplars in category learning. *J Exp Psychol Gen* 1998, 127:107–139.
35. Kruschke JK. Dimensional relevance shifts in category learning. *Connect Sci* 1996, 8:225–248.
36. Freedman DJ, Riesenhuber M, Poggio T, Miller EK. A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *J Neurosci* 2003, 23:5235–5246.
37. Goldstone RL, Steyvers M. The sensitization and differentiation of dimensions during category learning. *J Exp Psychol Gen* 2001, 130:116–139.
38. Jiang X, Bradley E, Rini RA, Zeffiro T, Vanmeter J, Riesenhuber M. Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 2007, 53:891–903.
39. Ell SW, Ing AD, Maddox WT. Critrial noise effects on rule-based category learning: the impact of delayed feedback. *Attent Percept Psychophys* 2009, 71:1263–1275.
40. Palmeri TJ. Exemplar similarity and the development of automaticity. *J Exp Psychol Learn Mem Cogn* 1997, 23:324–354.
41. Mack ML, Palmeri TJ. The timing of visual object recognition and categorization. *Front Percept Sci* 2011, 2:165.
42. Medin DL, Schaffer MM. A context theory of classification learning. *Psychol Rev* 1978, 85:207–238.
43. Palmeri TJ, Nosofsky RM. Recognition memory for exceptions to the category rule. *J Exp Psychol* 1995, 21:548–568.
44. Noelle DC, Cottrell GW. Individual differences in exemplar-based interference during instructed category learning. In: *Proceedings of the 22nd Annual Cognitive Science Society Conference*. Mahwah, NJ: Lawrence Erlbaum; 2000, 358–363.
45. Nosofsky RM, Clark SE, Shin HJ. Rules and exemplars in categorization, identification, and recognition. *J Exp Psychol Learn Mem Cogn* 1989, 15:282–304.
46. Ahn WK, Medin DL. A two-stage model of category construction. *Cogn Sci* 1992, 16:81–121.
47. Medin DL, Wattenmaker WD, Hampson SE. Family resemblance, concept cohesiveness, and category constructure. *Cogn Psychol* 1987, 19:242–279.
48. Gauthier I, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. Activation of the middle fusiform ‘‘face area’’ increases with expertise in recognizing novel objects. *Nat Neurosci* 1999, 2:568–573.

49. Maddox WT, Love BC, Glass BD, Filoteo JV. When more is less: feedback effects in perceptual category learning. *Cognition* 2008, 108:578–589.
50. Palmeri TJ, Flanery MA. Memory systems and perceptual categorization. *Psychol Learn Motivat* 2002, 41:141–189.
51. Palmeri TJ, Flanery MA. Prototype abstraction in category learning? *Proceedings of the Twenty-Third Annual Meeting of the Cognitive Science Society*, Edinburgh, Scotland, 2001.
52. Stewart N, Brown GDA, Chater N. Sequence effects in categorization of simple perceptual stimuli. *J Exp Psychol* 2002, 28:3–11.
53. Nosofsky RM. Similarity, frequency, and category representations. *J Exp Psychol Learn Mem Cogn* 1988, 14:54–65.
54. Church BA, Mercado E III, Wisniewski MG, Liu EH. Temporal dynamics in auditory perceptual learning: impact of sequencing and incidental learning. *J Exp Psychol Learn Mem Cogn* 2013, 39:270–276.
55. Giguère G, Love BC. Limits in decision making arise from limits in memory retrieval. *Proc Natl Acad Sci U S A* 2013, 110:7613–7618.
56. Pashler H, Mozer MC. When does fading enhance perceptual category learning? *J Exp Psycho* 2013, 39:1162–1173.
57. Poldrack RA, Clark J, Pare-Blagoev EJ, Shohamy D, Moyano JC, Myers C, Gluck MA. Interactive memory systems in the human brain. *Nature* 2001, 414:546–550.
58. Clapper JP, Bower GH. Learning and applying category knowledge in unsupervised domains. *Psychol Learn Motiv* 1991, 27:65–108.
59. Knowlton BJ, Squire LR. The learning of categories: parallel brain systems for item memory and category knowledge. *Science* 1993, 262:1747–1749.
60. Kruschke JK. ALCOVE: an exemplar-based connectionist model of category learning. *Psychol Rev* 1992, 99:22–44.
61. Nosofsky RM, Palmeri TJ. An exemplar-based random walk model of speeded classification. *Psychol Rev* 1997, 104:266–300.
62. Nosofsky RM, Palmeri TJ. A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psych Bull Rev* 1998, 5:345–369.
63. Bourne LE. Knowing and using concepts. *Psychol Rev* 1970, 77:546.
64. Trabasso T, Bower GH. *Attention in learning: Theory and Research*. New York: John Wiley & Sons; 1968.
65. Rosch E, Mervis CB. Family resemblances: studies in the internal structure of categories. *Cogn Psychol* 1975, 7:573–605.
66. Wittgenstein L. *Philosophical Investigations*. New York: Macmillan; 1953.
67. Homa D, Cross J, Cornell D, Goldman D, Schwartz S. Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. *J Exp Psychol* 1973, 101:116–122.
68. Dailey MN, Cottrell GW. Organization of face and object recognition in modular neural network models. *Neural Netw* 1999, 12:1053–1073.
69. Riesenhuber M, Poggio T. Hierarchical models of object recognition in cortex. *Nat Neurosci* 1999, 2:1019–1025.
70. Mack ML, Palmeri TJ. Modeling categorization of scenes containing consistent versus inconsistent objects. *J Vis* 2010, 10:1–11.
71. Ross DA, Deroche M, Palmeri TJ. Not just the norm: Exemplar-based models also predict face aftereffects. *Psych Bull Rev*. In press.
72. Ashby FG. *Multidimensional Models of Perception and Cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 1992.
73. Nosofsky RM. Similarity scaling and cognitive process models. *Annu Rev Psychol* 1992, 43:25–53.
74. Hintzman DL. Schema abstraction in a multiple-trace memory model. *Psychol Rev* 1986, 93:411–428.
75. Nosofsky RM. Exemplar representation without generalization? Comment on Smith and Minda's (2000). Thirty categorization results in search of a model. *J Exp Psychol* 2000, 26:1735–1743.
76. Medin DL, Schwanenflugel PJ. Linear separability in classification learning. *J Exp Psychol Hum Learn Mem* 1981, 7:355–368.
77. Nosofsky RM, Palmeri TJ. Exemplar-based random walk model. In: Busemeyer JR, Townsend J, Wang ZJ, Eidels A, eds. *Mathematical and Computational Models of Cognition*. Oxford University Press. In press.
78. Nosofsky RM. Choice, similarity, and the context theory of classification. *J Exp Psychol Learn Mem Cogn* 1984, 10:104–114.
79. Nosofsky RM. Attention, similarity, and the identification-categorization relationship. *J Exp Psychol Gen* 1986, 115:39–57.
80. Nosofsky RM. Optimal performance and exemplar models of classification. In: Oaksford M, Chater N, eds. *Rational Models of Cognition*. London: Oxford University Press; 1998, 218–247.
81. Pitt MA, Kim W, Navarro DJ, Myung JI. Global model analysis by parameter space partitioning. *Psychol Rev* 2006, 113:57–83.
82. Shepard RN, Chang JJ. Stimulus generalization in the learning of classifications. *J Exp Psychol* 1963, 65:94–102.
83. Anderson JR. The adaptive nature of human categorization. *Psychol Rev* 1991, 98:409–429.

84. Love BC, Medin DL, Gureckis TM. SUSTAIN: a network model of category learning. *Psychol Rev* 2004, 111:309–332.
85. Nosofsky RM. Relation between the rational model and the context model of categorization. *Psychol Sci* 1991, 2:416–421.
86. Ashby FG, Maddox WT. Relations between prototype, exemplar, and decision bound models of categorization. *J Math Psychol* 1993, 37:372–400.
87. Ashby FG, Alfonso-Reese L. Categorization as probability density estimation. *J Math Psychol* 1995, 39:216–233.
88. McKinley SC, Nosofsky RM. Investigations of exemplar and decision bound models in large, ill-defined category structures. *J Exp Psychol* 1995, 21:128–148.
89. Nosofsky RM, Kruschke JK. Investigations of an exemplar-based connectionist model of category learning. *Psychol Learn Motivat* 1992, 28:207–250.
90. Nosofsky RM, Palmeri TJ, McKinley SC. Rule-plus-exception model of classification learning. *Psychol Rev* 1994, 101:53–79.
91. Rosseel Y. Mixture models of categorization. *J Math Psychol* 2002, 46:178–210.
92. Estes WK. The problem of inference from curves based on group data. *Psychol Bull* 1956, 53:134–140.
93. Logan GD. Toward an instance theory of automatization. *Psychol Rev* 1988, 95:492–527.
94. Ashby FG, Alfonso-Reese LA, Turken AU, Waldron EM. A neuropsychological theory of multiple systems in category learning. *Psychol Rev* 1998, 105:442–481.
95. Ashby FG, Maddox WT. Human category learning 2.0. *Ann N Y Acad Sci* 2011, 1224:147–161.
96. Nosofsky RM, Johansen MK. Exemplar-based accounts of multiple-system phenomena in perceptual categorization. *Psychon Bull Rev* 2000, 7:375–402.
97. Logan GD, Gordon RD. Executive control of visual attention in dual-task situations. *Psychol Rev* 2001, 108:393–434.
98. Ashby FG, Gott RE. Decision rules in the perception and categorization of multidimensional stimuli. *J Exp Psychol* 1988, 14:33–53.
99. Ashby FG, Ell SW, Waldron EM. Procedural learning in perceptual categorization. *Mem Cogn* 2003, 31:1114–1125.
100. Nosofsky RM, Stanton RD, Zaki SR. Procedural interference in perceptual classification: implicit learning or cognitive complexity? *Mem Cogn* 2005, 33:1256–1271.
101. Stanton RD, Nosofsky RM. Feedback interference and dissociations of classification: evidence against the multiple-learning-systems hypothesis. *Mem Cogn* 2007, 35:1747–1758.
102. Lewandowsky S, Yang LX, Newell BR, Kalish ML. Working memory does not dissociate between different perceptual categorization tasks. *J Exp Psychol* 2012, 38:881–904.
103. Filoteo JV, Lauritzen S, Maddox WT. Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychol Sci* 2010, 21:415–423.
104. Newell BR, Moore CP, Wills AJ, Milton F. Reinstating the Frontal Lobes? Having More Time to Think Improves Implicit Perceptual Categorization. A Comment on Filoteo, Lauritzen, and Maddox (2010). *Psychol Sci* 2013, 24:386–389.
105. Squire LR, Zola SM. Structure and function of declarative and nondeclarative memory systems. *Proc Natl Acad Sci* 1996, 93:13515–13522.
106. Squire LR, Knowlton B. Learning about categories in the absence of memory. *Proc Natl Acad Sci* 1995, 92:12470–12474.
107. Nosofsky RM, Zaki S. Dissociations between categorization and recognition in amnesic and normal individuals: an exemplar-based interpretation. *Psychol Sci* 1998, 9:247–255.
108. Zaki SR. Is categorization really intact in amnesia? A meta-analysis. *Psychon Bull Rev* 2005, 11:1048–1054.
109. Palmeri TJ, Flanery MA. Learning about categories in the absence of training: profound amnesia and the relationship between perceptual categorization and recognition memory. *Psychol Sci* 1999, 10:526–530.
110. Nosofsky RM, Denton SE, Zaki SR, Murphy-Knudsen AF, Unverzagt FW. Studies of implicit prototype extraction in patients with mild cognitive impairment and early Alzheimer's disease. *J Exp Psychol Learn Mem Cogn* 2012, 38:860–880.
111. Zaki SR, Nosofsky RM. A single-system interpretation of dissociations between recognition and categorization in a task involving object-like stimuli. *Cogn Affect Behav Neurosci* 2001, 1:344–359.
112. Zaki SR, Nosofsky RM. A high-distortion enhancement effect in the prototype-learning paradigm: dramatic effects of category learning during test. *Mem Cogn* 2007, 35:2088–2096.
113. Wideman N. *Constructing Scientific Psychology: Karl Lashley's Mind-Brain Debates*. Cambridge: Cambridge University Press; 1999.
114. Nosofsky RM, Little DR, James TW. Activation in the neural network responsible for categorization and recognition reflects parameter changes. *Proc Natl Acad Sci* 2012, 109:333–338.
115. Gureckis TM, James TW, Nosofsky RM. Re-evaluating dissociations between implicit and explicit category learning: an event-related fMRI study. *J Cogn Neurosci* 2011, 23:1697–1709.
116. Davis T, Love BC, Preston AR. Striatal and hippocampal entropy and recognition signals in category learning: simultaneous processes revealed by model-based fMRI. *J Exp Psychol Learn Mem Cogn* 2012, 38:821–839.

117. Davis T, Love BC, Preston AR. Learning the exception to the rule: model-based fMRI reveals specialized representations for surprising category members. *Cereb Cortex* 2012, 22:260–273.

118. Palmeri TJ, Gauthier I. Visual object understanding. *Nat Rev Neurosci* 2004, 5:291–303.

119. Behrmann M, Plaut DC. Distributed circuits, not circumscribed centers, mediate visual cognition. *Trends Cogn Sci* 2013, 17:210–219.

120. Goldstone RL, Hendrickson AT. Categorical perception. *WIREs Cogn Sci* 2010, 1:69–78.

121. Pisoni DB. Identification and discrimination of the relative onset time of two-component tones: implications for voicing perception in stops. *J Acoust Soc Am* 1977, 61:1352–1361.

122. Goldstone RL. Influences of categorization on perceptual discrimination. *J Exp Psychol Gen* 1994, 123:178–200.

123. Gureckis TM, Goldstone RL. The effect of internal structure of categories on perception. In: *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*. Washington, DC: Cognitive Science Society; 2008, 1876–1881.

124. Gauthier I, Palmeri TJ. Visual neurons: categorization-based selectivity. *Curr Biol* 2002, 12:R282–R284.

125. Kanwisher N. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc Natl Acad Sci* 2010, 107:11163–11170.

126. Bukach CM, Gauthier I, Tarr MJ. Beyond faces and modularity: the power of an expertise framework. *Trends Cogn Sci* 2006, 10:159–166.

127. Gauthier I, Skudlarski P, Gore JC, Anderson AW. Expertise for cars and birds recruits brain areas involved in face recognition. *Nat Neurosci* 2000, 3:191–197.

128. Gauthier I, Tarr M, Bub D. *Perceptual Expertise: Bridging Brain and Behavior*. New York: Oxford University Press; 2009.

129. Wong AC-N, Palmeri TJ, Rogers BP, Gore JC, Gauthier I. Beyond shape: experience can determine patterns of category selectivity in the visual system. *PLoS One* 2009, 4:e8405.

130. Gauthier I, Tarr MJ. Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *J Exp Psychol Hum Percept Perform* 2002, 28:431–446.

131. Little DR, Lewandowsky S. Better learning with more error: probabilistic feedback increases sensitivity to correlated cues in categorization. *J Exp Psychol Learn Mem Cogn* 2009, 35:1041–1061.

132. Birnbaum MS, Kornell N, Bjork EL, Bjork RA. Why interleaving enhances inductive learning: the roles of discrimination and retrieval. *Mem Cogn* 2012:1–11.

133. Kang SH, Pashler H. Learning painting styles: spacing is advantageous when it promotes discriminative contrast. *Appl Cogn Psychol* 2012, 26:97–103.

134. Kornell N, Bjork RA. Learning concepts and categories: is spacing the "enemy of induction"? *Psychol Sci* 2008, 19:585–592.

135. Kornell N, Castel AD, Eich TS, Bjork RA. Spacing as the friend of both memory and induction in young and older adults. *Psychol Aging* 2010, 25:498.

136. Wahlheim CN, Dunlosky J, Jacoby LL. Spacing enhances the learning of natural concepts: an investigation of mechanisms, metacognition, and aging. *Mem Cogn* 2011, 39:750–763.

137. Goldstone RL. Isolated and interrelated concepts. *Mem Cogn* 1996, 24:608–628.

138. Davis T, Love BC. Memory for category information is idealized through contrast with competing options. *Psychol Sci* 2010, 21:234–242.

139. Kruschke JK. Learning involves attention. In: Houghton G, ed. *Connectionist Models in Cognitive Psychology*; East Sussex, UK: Psychology Press; 2005, 113–140.

140. Carvalho PF, Goldstone RL. Category structure modulates interleaving and blocking advantage in inductive category acquisition. *Proceedings of the Thirty-Fourth Annual Conference of the Cognitive Science Society*, 2012, 186–191.

141. Zulkiply N, Burt JS. The exemplar interleaving effect in inductive learning: moderation by the difficulty of category discriminations. *Mem Cogn* 2013, 41:16–27.

142. Smith JD, Minda JP. Prototypes in the mist: the early epochs of category learning. *J Exp Psychol Learn Mem Cogn* 1998, 24:1411–1436.

143. Smith JD, Murray MJ, Minda JP. Straight talk about linear separability. *J Exp Psychol Learn Mem Cogn* 1997, 23:659–680.

144. Lewandowsky S, Palmeri TJ, Waldmann MR. Introduction to special section on theory and data in categorization: integrating computational, behavioral, and cognitive neuroscience approaches. *J Exp Psychol Learn Mem Cogn* 2012, 38:803–806.

145. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013, 14:365–376.

146. Francis G. Publication bias and the failure of replication in experimental psychology. *Psychon Bull Rev* 2012, 19:975–991.

147. Nieuwenhuis S, Forstmann BU, Wagenmakers E-J. Erroneous analyses of interactions in neuroscience: a problem of significance. *Nat Neurosci* 2011, 14:1105–1107.

148. Vul E, Harris C, Winkielman P, Pashler H. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect Psychol Sci* 2009, 4:274–290.

149. Mack ML, Palmeri TJ. The dynamics of categorization: unraveling rapid categorization. Submitted for publication.
150. Richler JJ, Gauthier I, Palmeri TJ. Automaticity of basic-level categorization accounts for labeling effects in visual recognition memory. *J Exp Psychol Learn Mem Cogn* 2011, 37:1579–1587.
151. Richler JJ, Palmeri TJ, Gauthier I. Meanings, mechanisms, and measures of holistic processing. *Front Percep Sci* 2012, 3:553.
152. Richler JJ, Palmeri TJ, Gauthier I. How does using object names influence visual recognition memory? *J Mem Lang* 2013, 68:10–25.

FURTHER READING

Luck S, Hollingworth A. *Visual Memory*. New York: Oxford University Press; 2008.

Margolis E, Laurence S. *Concepts: Core Readings*. Cambridge, MA: MIT Press; 1999.

