Vision Research 49 (2009) 1961–1968

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Time course of visual object categorization: Fastest does not necessarily mean first

Michael L. Mack [a], Alan C.-N. Wong [b], Isabel Gauthier [a], James W. Tanaka [c], Thomas J. Palmeri [a,*]

[a] Psychology Department, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, United States
[b] Department of Psychology, Sino Building, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong, China
[c] Psychology Department, University of Victoria, P.O. Box 3050, STN CSC, Victoria, BC, Canada V8W 3P5

ARTICLE INFO

Article history:
Received 10 February 2009

Keywords:
Object recognition
Object categorization
Basic level
Time course
Temporal dynamics

ABSTRACT

Perceptual categorization at the basic level is generally faster than categorization at more superordinate or subordinate level [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]. But, what does it mean to be fastest? One possibility is that levels of abstraction that are categorized fastest are processed first. In this vein, the basic level is often considered the "entry level" into our knowledge about categories in the world [Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16*(2), 243–275]. We tested this "fastest means first" hypothesis by contrasting the time course of basic- and subordinate-level categorization of objects in a signal-to-respond experiment. This method probes subjects to respond at systematically varying points in time after the onset of the object. The time course function relating performance to time is characterized by its onset, growth rate, and asymptote. While basic and subordinate categorization differed significantly in growth rate and asymptote, they did not differ significantly in onset. If a basic-level stage preceded a subordinate-level stage, we should have observed a difference in onset. We conclude that fastest does not necessarily mean first in perceptual categorization.

1. Introduction

The human visual system allows us to rapidly and accurately recognize objects in the world (Thorpe, Fize, & Marlot, 1996). At a glance, we can detect that an object is there, categorize it as a bird, or identify it as blue jay. An important and long-standing question about object processing is when these different levels of abstraction become available to the perceiver. Some of these perceptual decisions are made more quickly than others (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). But does fastest mean first? Do certain perceptual decisions start earlier than others during visual object recognition (Grill Spector & Kanwisher, 2005; Mack, Gauthier, Sadr, & Palmeri, 2008; Palmeri, Wong, & Gauthier, 2004)?

Rosch et al. (1976) found that participants were faster at verifying that objects matched labels at the so-called *basic level* (e.g., dog) than more superordinate (e.g., animal) or subordinate (e.g., beagle) levels of abstraction. The fastest level of categorization was later termed the *entry level* by Jolicoeur, Gluck, and Kosslyn (1984) to acknowledge that the level at which perceptual information makes first contact with a stored visual representation. For many category members, "basic-level [categorization] occurs first and is followed, some time later, by subordinate-level identification" (Jolicoeur et al., 1984, p. 270). Rosch et al. (1976) argued that the advantage for the basic level arises because the basic level is the level at which objects show the largest gain in structural similarity independent of the perceiver; representations of basic-level categories follow the natural correlations and divisions of features found in objects and, as a consequence, are available first during recognition. Various factors can influence which level is fastest. For example, atypical category members that are structurally dissimilar to their subordinate counterparts can be categorized faster at subordinate levels than the basic level (Jolicoeur et al., 1984; Murphy & Brownell, 1985). Furthermore, for objects of perceptual expertise (Tanaka & Taylor, 1991), subordinate-level identification occurs as quickly and as accurately as basic-level categorization.[1] This has been characterized as an *entry-level shift* with expertise: For novice categories, the basic level is the entry level, but for expert categories, more subordinate levels become the entry level (but see Johnson & Mervis, 1997).

But what does it mean for a particular level of abstraction to be the "entry level"? One straightforward possibility is illustrated in Fig. 1. In this simple box-and-arrow model, objects from novice

[1] For clarity and ease of prose, we often use the terms basic-level *categorization* and subordinate-level *identification* in this paper. In this sense of the terms, both "categorization" and "identification" are categorizations, albeit at different levels of abstraction.

* Corresponding author. Fax: +1 615 343 8449.
E-mail address: thomas.j.palmeri@vanderbilt.edu (T.J. Palmeri).

0042-6989/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2009.05.005

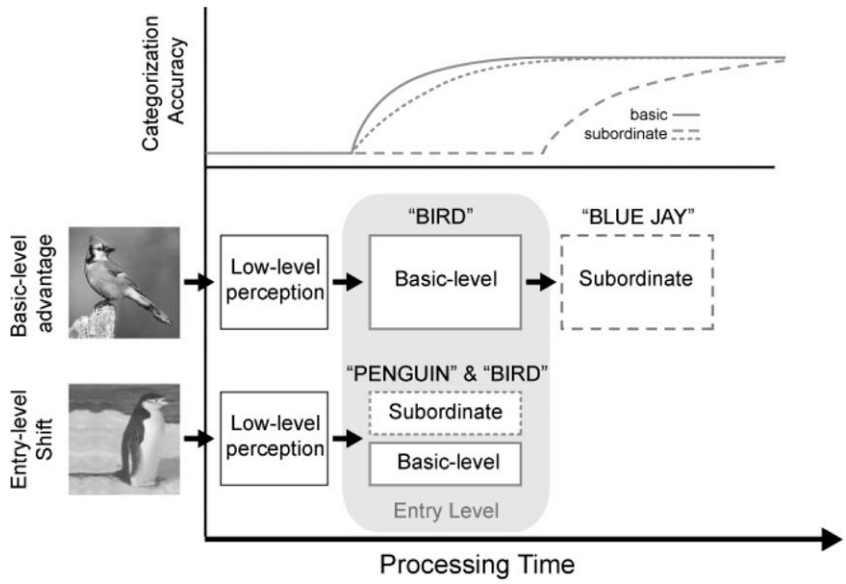

Fig. 1. One possible descriptive model of basic-level advantage (top) and entry-level shift (bottom). Speed-accuracy tradeoff functions show the (exaggerated) hypothetical time course of processing associated with novice and experiment categories according to such a model.

categories, after low-level visual processing, are categorized first at the basic level (the entry level) before being categorized at more subordinate (or superordinate) levels. Basic-level categorization is faster than subordinate-level identification because basic-level categorization occurs *before* subordinate-level identification begins – fastest means first. But for objects from expert categories or atypical objects, there is an entry level shift: Objects are categorized at subordinate levels of abstraction without first being categorized at the basic level. As illustrated in Fig. 1, by this account learning about expert categories and atypical objects creates special-purpose machinery for rapidly recognizing subordinate categories that bypasses an initial basic-level stage of processing.

Characterizing stages of visual processing with levels of categorization has obtained some currency in visual cognition and visual neuroscience. For example, Grill Spector and Kanwisher (2005) contrasted the time course of object detection, basic-level categorization, and subordinate-level identification and observed the very same rapid time course for detection and categorization compared to identification. These results suggest an early stage of image segmentation that both detects that an object is there and tells you what basic-level category the object belongs to. Subordinate-level identification takes place in a subsequent stage of processing.

However, many extant computational models of object recognition and categorization propose no such preliminary basic-level stage. Instead, basic-level categorization and subordinate-level identification are perceptual decisions at the end of the line of processing, not sequential stages of visual processing. For example, exemplar-based models of categorization (Kruschke, 1992; Nosofsky, 1992) can account for both categorization and identification performance. In broad strokes, these models assume some initial stage of perceptual processing that provides the perceptual representation of an object. This perceptual representation activates stored exemplars in memory according to their similarity to the presented object, with perceptual dimensions more diagnostic of category or identity contributing more to similarity than non-diagnostic dimensions. Stored exemplars are associated with basic-level categories or subordinate-level identities through weights learned by Hebbian or error-driven learning rules (depending on the particular model). A stochastic random walk decision process at this final decision stage accounts for both errors and variability in response time (Nosofsky & Palmeri, 1997; Palmeri, 1997). In the parlance of neural network models, decisions about both category and identity are made within the final output layer, not in earlier layers (Nosofsky & Kruschke, 1992).

A similar hierarchy of information processing is seen in other models. One neural network model of object recognition (Joyce & Cottrell, 2004) assumes that an object goes through stages of Gabor filtering, principal component analysis (PCA), and a neural network mapping PCA representations onto category labels. Decisions about basic-level category or subordinate-level identity are driven by trained weights leading to units at the same final output layer of the neural network. Similarly, another neural network model of object recognition (Riesenhuber & Poggio, 2000) assumes a hierarchy of information processing that begins with low-level features, moves on to view-based representations, object representations, and ultimately to labels for category and identity. Like the other models, perceptual decisions at different levels of abstractions are instantiated at the same output layer of the network. Critically, none of the current models postulate an explicit basic-level stage of processing that precedes the subordinate and superordinate stages. A basic-level advantage arises in many models because of the greater level of structural similarity among basic-level category members and the greater dissimilarity to other categories. If there is truly a basic-level stage of processing – as suggested by some interpretations of entry-level phenomena – then this would challenge many current computational models of perceptual categorization and object recognition.

We attempted to unravel the time course of basic-level categorization and subordinate-level identification. Is basic-level categorization a stage prior to subordinate-level identification? If so, then most models are wrong.

Specifically, we asked whether the onset of processing for basic-level categorization occurs *before* the onset of processing for subordinate-level identification. Our first experiment verified that basic-level categorization is significantly *faster* than subordinate-level categorization for typical members of novice categories; by contrast, atypical members should be categorized as fast at subordinate and basic levels. We then asked if fastest might also mean *first*. The second experiment used a signal-to-respond (STR) technique to examine the time course of basic-level categorization and subordinate-level identification. STR probes perceptual decisions at various time points after the stimulus appears. Of particu-

lar interest are potential temporal markers for understanding the source of the basic-level advantage. If subordinate-level identification of novice categories is performed by a stage of processing that begins only after basic-level categorization finishes (as illustrated in top part of Fig. 1), then we should find a delay in the initial onset of subordinate-level decisions relative to basic-level decisions over time.

In addition to testing participants on novice categories of objects, we also tested participants on faces. Faces provide an interesting contrast category for two reasons. First, there have been explicit suggestions that an initial stage of processing categorizes a stimulus as a face prior to a stage that identifies the unique individual. This suggestion has been supported by time course measures using EEG (e.g., Anaki, Zio-Golumbic, & Bentin, 2007) and MEG (Liu, Harris, & Kanwisher, 2002). So perhaps faces are like common objects from novice categories. They are categorized at the basic level as a person. Then in a subsequent stage of processing they are identified uniquely. If that were true, then like novice categories, during the STR task we might expect a significant difference in the onset of processing for basic-level categorization and subordinate-level identification of faces.

Second, others have suggested that normally-functioning adults can be considered face experts (Carey, 1992; Carey & Diamond, 1994; Diamond & Carey, 1986; Gauthier & Tarr, 1997; Tanaka, 2001), though whether face expertise is qualitatively different from other forms of perceptual expertise is hotly debated (Bukach, Gauthier, & Tarr, 2006; McKone, Kanwisher, & Duchaine, 2007). It is true, however, that faces show qualitatively the same entry-level shift as other categories of expertise. Specifically, pictures of highly familiar faces are identified as quickly as unique individuals as they are categorized as people (Tanaka, 2001), much in the same way that for bird experts, pictures of birds are identified as quickly at a subordinate level as they are categorized as birds (Tanaka & Taylor, 1991). By this alternative account, during the STR task we might expect no difference in the onset of processing for basic-level categorization and subordinate-level identification for faces.

2. Experiment 1

2.1. Methods

2.1.1. Participants

Fifteen Vanderbilt University undergraduates participated in two 1 h sessions for course credit or monetary compensation.

2.1.2. Stimuli

Images of objects from three categories (faces, dogs, and birds) were used. Each category consisted of about 320 images from eight different subordinate-level categories: faces – Arnold Schwarzenegger, Jennifer Aniston, Britney Spears, Nicole Kidman, George W. Bush, Mel Gibson, Hillary Clinton, Bill Clinton; dogs – sharpei, beagle, chihuahua, chow chow, golden retriever, german shepherd, weimaraner, poodle; birds – robin, dove, crow, hawk, duck, penguin, ostrich, owl. Two of the dogs (chihuahua and poodle) and two of the birds (penguin and ostrich) are atypical members of the category as defined in previous work (Jolicoeur et al., 1984; Rosch et al., 1976), the rest of the dogs and birds are typical. Images were presented in grayscale and subtended approximately $5.2^\circ \times 5.2^\circ$ of visual angle.

2.1.3. Procedure

Participants were seated approximately 60 cm from the computer display and performed a speeded category verification task. Each trial began with a basic- or subordinate-level category label displayed for 1000ms, followed immediately by the test image. The test image remained on the screen until the participant responded. Participants responded by hitting a "yes" key if the label matched the object shown in the test image, and a "no" key if it did not. Half of the category verifications were made at the basic level (face, dog, or bird), and half were made at the subordinate level (Jennifer Aniston, beagle, robin, etc.). On true trials, the category label and the object in the test image matched. On false basic-level trials, another basic-level category was shown (e.g., a label BIRD for the image of a german shepard). On false subordinate-level trials, another category label from the same basic-level category was displayed (e.g., a label BEAGLE for the image of a german shepherd); for faces, the label on false trials was a person of the same gender as the one depicted in the image. Participants were instructed to respond as quickly and accurately as possible. Before the experimental trials began participants completed a short practice session; the practice stimuli were drawn from other basic-level categories. Each session consisted of 960 trials and lasted approximately 1 h.

2.2. Results and discussion

Verification response times and accuracy on true trials from each of the object categories are shown in Fig. 2. A basic-level advantage was found for birds and dogs, replicating Tanaka and Taylor (1991), but not faces, replicating Tanaka (2001). Both the response time and accuracy data were subjected to a within-subjects domain (dog, bird, and face) $\times$ level (basic, subordinate) ANOVA. Overall, responses were faster ($F(2, 28) = 23.4$, MSE = 2571, $p < 0.001$) and more accurate ($F(2, 28) = 4.48$, MSE = 0.00024, $p < 0.05$) for faces than dogs or birds, and responses were faster ($F(1, 14) = 23.3$, MSE = 2330, $p < 0.001$) and more accurate ($F(1, 14) = 17.29$, MSE = 0.00066, $p < 0.01$) for basic-level than subordinate-level verifications. Critically, significant domain $\times$ level interactions were observed for both response time ($F(2, 28) = 7.69$, MSE = 1120, $p < 0.01$) and accuracy ($F(2, 28) = 8.91$, MSE = 0.00029, $p < 0.01$). Planned comparisons were conducted on the difference between the basic- and subordinate-level verifications for each domain. For both birds and dogs, responses were faster and more accurate for basic than subordinate verifications [birds – RT $t(14) = 4.01$, $p < 0.01$, accuracy $t(14) = 3.93$, $p < 0.01$; dogs – RT $t(14) = 5.24$, $p < 0.001$, accuracy $t(14) = 5.04$, $p < 0.001$]. For faces, no significant difference was found for either response time [$t(14) = 1.81$, $p = 0.203$] or accuracy [$t(14) < 1.0$, $p = 0.87$].

The above analyses included all birds and dogs, regardless of their typicality. Not surprisingly, when we excluded the atypical objects from the analyses conducted above, all of statistical contrasts were at least as strong (Table 1). We then analyzed the atypical objects separately (Table 1). For the atypical birds, subordinate-level categorization was actually significantly faster than basic-level categorization [$t(14) = 2.98$, $p = 0.01$], with no significant difference in accuracy [$t(14) = 1.02$, $p = 0.327$].[2] For the atypical dogs, basic-level categorization was still significantly faster than subordinate-level categorization [$t(14) = 4.19$, $p = 0.001$] with no significant difference in accuracy [$t(14) = 0.82$, $p = 0.424$].

In addition to examining average accuracy and response times, we also examined the full response time distributions for basic-level categorization and subordinate-level identification. We were specifically interested in the fastest tail of the RT distributions. If

[2] We also analyzed data for each individual bird and dog separately. In all but two cases, typical objects were categorized faster at the basic than subordinate levels and atypical objects were categorized at least as fast at the subordinate level as the basic level. The only exceptions were owl, which showed response times like atypical objects, and poodle, which showed response times like typical objects.

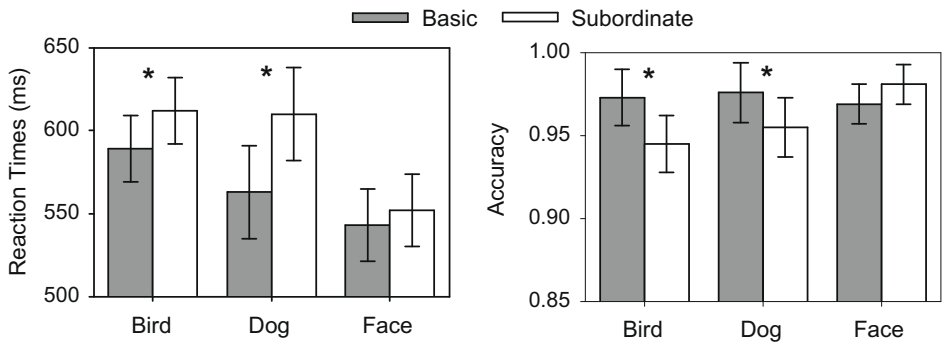

Fig. 2. RT and accuracy for speeded verification in Experiment 1. Gray and white bars represent basic- and subordinate-level performance, respectively. Asterisks (*) represent significant differences ($p < 0.05$) between basic- and subordinate-level performance and error bars represent 95% confidence intervals.

Table 1
Average accuracy and RT in Experiment 1 for typical and atypical objects in each domain. 95% confidence intervals are shown in parentheses as well as significant differences between basic-level categorization and subordinate-level identification at $p < 0.05$ (*).

Condition		Accuracy	RT
Typical			
Bird	Basic	.982 (.951, 1.00)*	572 (539, 605)*
	Subordinate	.941 (.909. ,972)	647 (614, 680)
Dog	Basic	.982 (.965, .999)*	563 (532, 595)*
	Subordinate	.952 (.935, .968)	613 (581, 645)
Person	Basic	.969 (.957, .984)	544 (521, 566)
	Subordinate	.981 (.966, .993	553 (530, 575)
Atypical			
Bird	Basic	.970 (.958, .983)	607 (581, 632)*
	Subordinate	.980 (.967, .992)	568 (543, 594)
Dog	Basic	.972 (.944, .999)	564 (541,586)*
	Subordinate	.961 (.933, .989)	608 (565, 630)

there is clear separation between the RT distributions at these fastest RTs, such that the fastest basic-level categorizations are faster than the fastest subordinate-level identifications, this could provide some converging evidence for a basic-level stage preceding subordinate-level processing.

To do this comparison, we created Vincentized (Ratcliff, 1979; Vincent, 1912) RT distributions for basic-level categorization and subordinate-level identification, as shown in Fig. 3. Vincentizing is a technique for creating an average RT distribution that preserves the shape of the individual-participant RT distributions; it is well known that if individual-participant RTs are simply piled together into a single group-defined RT distribution, then the group RT distribution can have a very different shape from any of the individual RT distributions. Vincentizing first creates a cumulative RT distribution for each individual participant. At each quantile of the distribution, the RT at that quantile for each individual RT distribution is averaged together. For Fig. 3, we chose a fine-grained Vincentizing at each 5% (1% for zoomed figure insets). The shaded region is a confidence interval on the Vincentized RT distribution generated with a bootstrapping procedure.[3] For typical objects, even though the distributions for basic- and subordinate-level decisions are separated over the bulk of the RT distributions, the separation for the fastest RTs is less clear. For atypical objects, we see a great deal of overlap in the RT distributions over their full extent. We see a similar degree of overlap in the distributions for faces.

To summarize, for objects from novice categories, verifications at the basic level were faster and more accurate than those at a subordinate level. But for faces, there was an "entry-level shift", with comparable speed and accuracy at the subordinate level as the basic level. Comparing RT distributions did not show clear evidence for a difference in onset of correct responses for basic-level categorization versus subordinate-level identification. We further explored the time course of categorization and identification in Experiment 2.

3. Experiment 2

For novice categories, objects are categorized faster at the basic level than subordinate levels. But are these objects categorized at the basic level *before* subordinate identification begins? Does fastest mean first?

To answer this question, we contrasted the time course of categorizing expert and novice objects at basic and subordinate levels using a *signal-to-respond* (STR) technique, or also called a *response-signal* technique (Corbett & Wickelgren, 1978; Dosher, 1981; Hintzman, Caulton, & Curran, 1994; Reed, 1973). This task can be used to unravel the time course of visual object processing by systematically varying the amount of time a participant is given to process a test object and measuring how categorization performance changes as a function of processing time.

Another common technique for probing the time course of object processing involves systematically manipulating the exposure duration of images (rather than manipulating the time to make an object decision). This is a useful technique, especially for understanding what kinds of perceptual decisions are possible at a glance. One potential limitation is that, especially for very rapid exposure durations, it is difficult to disentangle the time-course of processing per se from the quality of the perceptual representation being processing. If early visual areas need to temporally integrate sensory information, then very rapidly presented images will have a degraded sensory representation. By holding exposure duration constant, we can focus our lens on the time-course of decisions. For this reason, STR techniques are commonly used when asking questions about potential stages of processing that lead to those decisions (e.g., Hintzman & Curran, 1997).

We introduced a STR version of the category verification task, using the same collection of objects and categories. At systematically varying lags after the appearance of each image, we presented a response signal (a tone) that instructed the participant to immediately make a response. Specifically, the participant was instructed to respond within a short time window after hearing the tone.

[3] The confidence interval was generated by creating 5000 Vincentized RT distributions and setting the upper and lower bounds of the confidence interval to the 2.5% and 97.5% extent of a distribution of distributions. Specifically, we used a bootstrap procedure whereby on each of the 5000 simulated runs, we created a RT distribution for each participant by sampling their observed RTs with replacement, and then created a sample Vincentized RT distribution using the same approach we used for the actual sample of observed data.

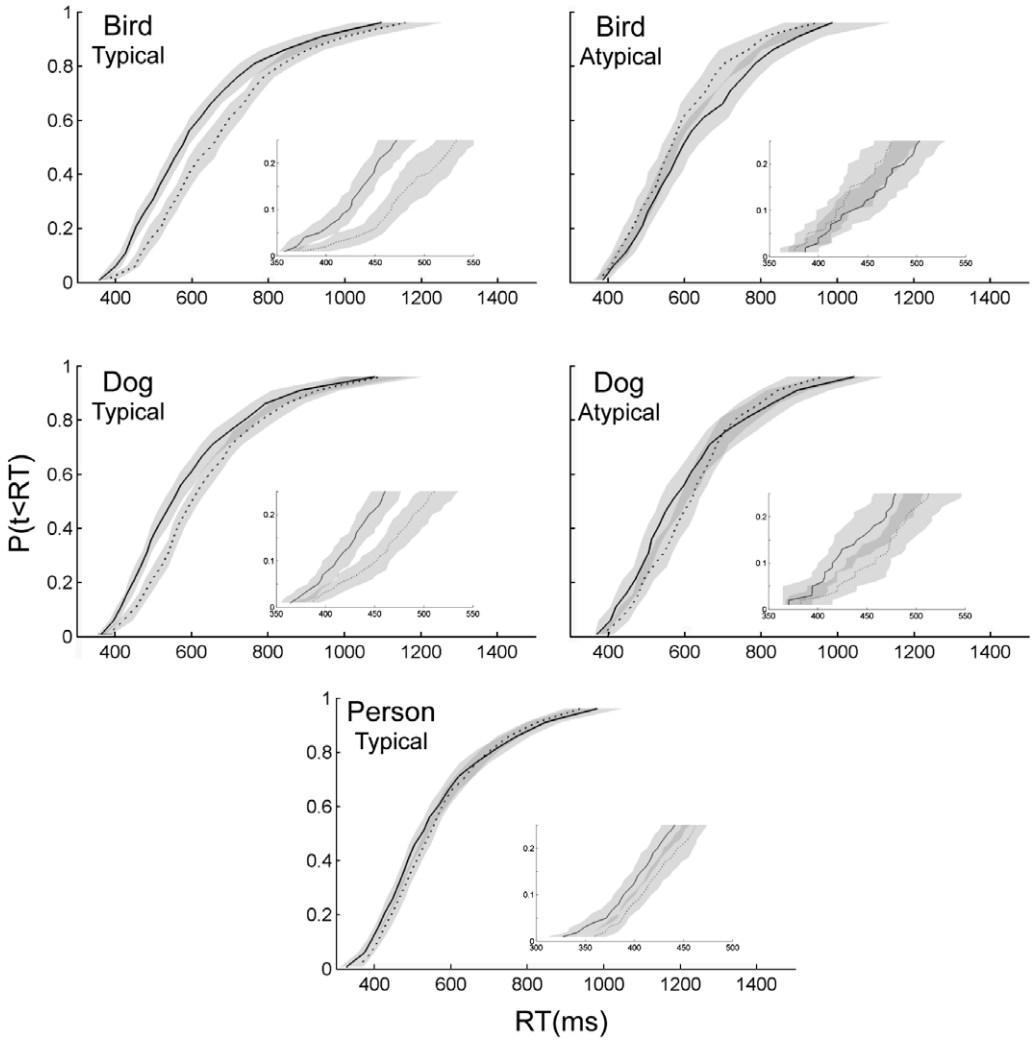

Fig. 3. Cumulative distributions of correct true trials in Experiment 1; basic-level categorization (solid line) and subordinate-level identification (dotted line), shaded regions represent 95% confidence intervals. Insets zoom in on the fastest tail of distributions.

Varying the lag from image appearance to response signal allows us to examine how category verification performance changes over time. Fig. 1 illustrates the *speed-accuracy tradeoff functions* (SATF) that are typically observed in STR paradigms. These curves can be characterized by their *onset*, the time at which categorization performance begins to grow above chance, *growth rate*, how steeply categorization performance increases with increasing time, and *asymptote*, the maximum level of categorization performance possible.

Of particular interest to us are differences in onset for basic-level categorization and subordinate-level identification. As illustrated in Fig. 1, if for novice categories the onset of subordinate-level identification requires completion of a basic-level categorization stage, then there should be some window of time where above-chance performance is possible for basic-level categorization but not subordinate-level identification – this would be reflected by a significant onset difference in the SATF. Of course, the illustration in Fig. 1 exaggerates the onset difference we might expect. But given that there was a 50ms difference between basic- and subordinate-level decisions for novice categories in Experiment 1, there is a potential opportunity to uncover significant onset differences using a STR paradigm. Indeed, it is common to find significant differences in the onset of speed-accuracy tradeoff functions for a variety of simple decisions that can be made as rapidly as the categorization decisions under consideration in this article, from lexicality and memory decisions (e.g., Hintzman & Curran, 1997; Hintzman et al., 1994) to categorization decisions (e.g., Lamberts, 2000; Lamberts & Freeman, 1999) to visual perceptual decisions (e.g., Carrasco, McElree, Denisova, & Giordano, 2003).

Now as for faces, the previous research using MEG and EEG suggests that we might expect to find a significant onset difference for basic-level categorization versus subordinate-level identification (e.g., Anaki et al., 2007; Liu et al., 2002). On the other hand, Experiment 1 showed no significant difference in category verification at the basic and subordinate levels, so we might also expect to find no onset difference whatsoever.

3.1. Methods

3.1.1. Participants

Five of the participants from the first experiment took part in this experiment and were paid $12 per session.

3.1.2. Stimuli

The same stimuli were used as Experiment 1.

3.1.3. Procedure

Participants completed a category verification task like Experiment 1, but with the inclusion of a signal-to-respond manipula-

tion. On each trial, a category label was displayed for 1000ms, and then a premask was displayed for a variable duration, followed by the presentation of the stimulus image for 200ms, followed by a postmask. An auditory signal to respond was presented to the participants after a variable lag (12, 24, 35, 47, 94, 188, 376, 753, or 1506 ms) from image appearance. Masking was used to limit the amount of perceptual processing in order to make the task more difficult than unmasked viewing; note that the same limits from masking were imposed at all signal-to-respond levels. As in Experiment 1, participants verified the match or mismatch between the category label (basic or subordinate) and object in the stimulus image, but they could only respond after hearing the auditory signal. A warning message was presented if the participants responded before the signal or if the response time after the signal was smaller than 180ms or greater than 350ms. Participants responded by pressing keys marked as "yes" and "no" on a keyboard. Participants completed 16 sessions with each session consisting of 864 trials and lasting approximately 1 h. This resulted in 256 trials for each lag in every domain (dog, bird, and person) and category label (basic or subordinate).

3.2. Results

Fig. 4 displays the average observed speed-accuracy tradeoff functions in terms of discriminability (d') as a function of processing time for basic- and subordinate-level categorization of typical and atypical (insets) objects in each of the three object domains. In order to quantitatively compare the temporal dynamics of speed-accuracy tradeoff functions, d' values from individual participants were fitted with an exponential function widely used to analyze STR data (Wickelgren & Corbett, 1977)

$$d' = \lambda(1 - e^{-\beta(t-\delta)}),$$

where t is the lag until the response signal plus the response time after the signal (i.e., t = signal lag + RT), λ is the asymptote, β is the growth rate, and δ is the onset. The asymptote represents an expected maximum accuracy for the task given unlimited time; the growth rate represents the rate at which relevant information is extracted; the onset represents when performance begins to grow above chance during the time course of processing. By fitting this function to each participant's data, we can statistically compare the resulting parameters values (λ, β, and δ) for basic- and subordinate-level categorization in each domain. If, for example, we find statistically shorter onsets for the basic than subordinate levels, then this indicates a delay in initial processing of subordinate-level categories, perhaps because subordinate-level categorization follows basic-level categorization.

After fitting the exponential function to each individual participant's speed-accuracy tradeoff data, we conducted planned comparisons testing for differences in asymptote, growth rate, and onset parameters between basic- and subordinate-level decisions. Average values of the asymptote, growth, and onset parameters are shown in Table 2. For typical objects from novice categories (birds and dogs), no significant difference was observed for the onsets [$t(4) < 1.0$]. For birds, the growth rate [$t(4) = 3.64$, $p = 0.003$] was significantly higher for basic-level categorizations and the asymptote [$t(4) = 2.33$, $p = 0.079$] was marginally higher for basic-level categorizations. For dogs, a marginally significant difference was observed in asymptote [$t(4) = 2.18$, $p = 0.094$] with a higher asymptote for basic-level categorizations. For faces, planned comparisons revealed a marginally significant difference in the growth rate [$t(4) = 2.59$, $p = 0.061$]. Interestingly, a small but significant difference in onset, $t(4) = 4.53$, $p = 0.027$, was observed, with the basic-level condition having the shorter onset. For atypical objects, no significant differences were observed in any of the SATF parameters [$t(4) < 1$].

In addition to simply fitting the exponential functions to the individual speed-accuracy tradeoff functions, we also tested hypotheses by fitting special cases of the function. Specifically, we tested whether the onset for basic and subordinate decisions was the same by constraining the onset to be identical for basic and subordinate decisions but allowing the growth rate and asymptote to vary. We contrasted the fit of the "full model", with three parameters for basic and three parameters for subordinate, with a "restricted model", with a common onset parameter for

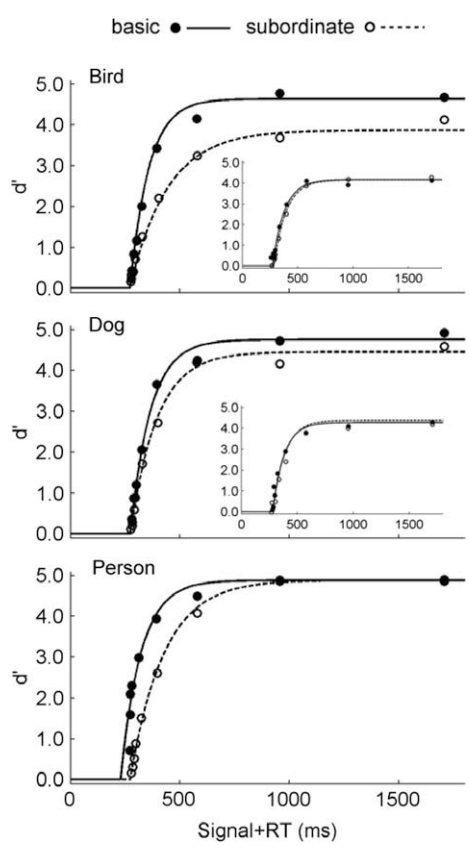

Fig. 4. Speed-accuracy tradeoff functions in Experiment 2; behavioral time course data (close and open circles) and exponential curve fits (solid and dotted lines) for typical objects in each domain (atypical objects are shown in the insets). Performance (d') is plotted along the y-axis and response time plus lag is plotted along the x-axis.

Table 2
Average best fitting parameters from the basic-level categorization and subordinate-level identification SATF for typical and atypical objects in each domain. Significant differences between parameters for basic-level categorization and subordinate-level identification are labeled at $p < 0.05$ (**) and $p < 0.10$ (*).

Condition		Parameters		
		Asymptote	Growth rate	Onset
Typical				
Bird	Basic	4.67*	10.51**	0.272
	Subordinate	3.95	6.07	.0267
Dog	Basic	4.76*	10.91	0.273
	Subordinate	4.46	8.23	0.273
Person	Basic	4.88	11.25*	0.228**
	Subordinate	4.89	6.46	0.272
Atypical				
Bird	Basic	4.10	11.42	0.279
	Subordinate	4.29	9.03	0.282
Dog	Basic	4.181	9.21	0.269
	Subordinate	4.17	8.90	0.283

both basic and subordinate but separate growth and asymptote parameters for basic and subordinate. If the restricted model with a common onset fits significantly worse than the full model, we can reject the hypothesis that the onset is the same for basic- and subordinate-level categorizations. Following Dosher (1981), the quality of the fitted models was assessed using an R^2 statistic that represents the proportion of variance accounted for by the model and was adjusted by the number of free parameters in the model. Model comparisons are based on the direct comparison of the R^2 values with higher R^2 values indicating better accounts of the observed data. For typical objects from novice categories (birds and dogs), the restricted models with equal onsets fitted as well as the full model where onsets could be different (for birds, average R^2 was .948 for the full model and .957 for the restricted model; for dogs, average R^2 was .912 for the full model and .932 for the restricted model). For faces, the restricted model with equal onsets (average R^2 = .858) fitted as well as the full model with unequal onsets (average R^2 = .861). Fits of the SATF for atypical objects showed similar results as for typical objects, with the restricted model fitting as well as the full model for both atypical birds (full: R^2 = .928, restricted: R^2 = .934) and atypical dogs (full: R^2 = .873, restricted: R^2 = .897).

3.3. Discussion

For the novice categories, even though basic-level categorization was faster and more accurate than subordinate-level identification in Experiment 1, there was no significant delay in the onset of subordinate-level identification compared to basic-level categorizations in the STR task. Basic-level categorizations may be made faster than subordinate-level identification, but basic-level categorization does not appear to be a stage of processing that precedes subordinate-level identification.

One potential concern could be that we failed to muster sufficient statistical power to detect significant differences in the onset when the exponential functions were fitted to the observed data.[4] In fact, we were able to observe a small but statistically significant onset difference of only 27 ms for faces. This onset effect is one half the size of the basic-level advantage in response time we observed in the first experiment. This suggests that we had sufficient statistical power to detect a putative onset difference with dogs and birds.

Turning now to the significant onset difference with faces, recall that the stage model outlined in Fig. 1 predicts an onset difference for novice objects, with little or no onset differences for expert categories. That is clearly not what we found. So why might basic-level categorizations of faces show an earlier onset than subordinate-level identification of faces? Perhaps a face is first categorized as a "face" before it is identified uniquely. Indeed, this stage-like processing of faces has been suggested by some (e.g., Anaki et al., 2007; Liu et al., 2002). It is surprising, however, that such stage-like processing would be found for faces, which do not show any basic-level advantaged in speeded categorization, and not for objects from novice categories, which do show a basic-level advantage. Perhaps this is a special property of face processing that is not true for other categories of object.

Alternatively, it may be that basic-level categorization – *is there a "face" in the image?* – could be driven by low-level image properties available very early in visual processing. It is known that people can rapidly categorize based on perceptually salient features and that such salient features are often available before less salient, but potentially more diagnostic features (e.g., Lamberts, 2000). To gain some further insight into this puzzling result, we averaged together all of the face images, bird images, and dog images without controlling for viewpoint differences across the images. For face images, this average was roughly an oval face contour. The average of the bird images and dog images did not look like a bird or a dog or any other clearly identifiable basic shape. Further research is necessary to completely understand this phenomenon (e.g., are onset differences found with face stimuli that include more variable viewpoints?). What is clear, however, is that non-face (non-expert) objects did not show any stage-like processing effects in the speed-accuracy tradeoff function.

4. General discussion

For novices, objects are categorized faster at the basic level than at more subordinate levels (Rosch et al., 1976). Jolicoeur et al. (1984) noted that such data are consistent with a model where objects must first be categorized at the basic level before they can be categorized at coarser or finer levels, speculating that "every object has one particular level at which contact is made first with semantic memory" (p. 272). According to this view, basic-level categorization is fast because it is the "entry level" into semantic knowledge. In other words, basic-level categorization is fast because it is completed before other stages of categorization can begin. Interestingly, the difference between basic-level categorization and subordinate-level identification typically disappears with expertise (Gauthier & Tarr, 1997; Tanaka & Taylor, 1991). This has been characterized as an "entry-level shift", whereby expert objects and atypical objects can be identified at a subordinate level without first being categorized at the basic level.

This stage-like view of visual object processing is implicit in some writings and has been explicitly suggested recently by Grill Spector and Kanwisher (2005). They contrasted the time course of object detection, basic-level categorization, and subordinate-level identification by systematically varying the exposure duration of images in the experiment. Performance on subordinate-level identification was significantly worse than object detection and basic-level categorization at all exposure durations, but performance on object detection and basic-level categorization was identical. Their conclusion was stated in the paper's subtitle, "As soon as you know it is there, you know what it is." They suggested that image segmentation – detecting that an object is there – and basic-level categorization – knowing what it is – could be intimately linked as a stage of visual processing prior to subordinate-level identification. The tight temporal coupling between object detection and basic-level categorization has been decoupled in recent work (Bowers & Jones, 2008; Mack et al., 2008). The present work examined the temporal relationship between basic-level categorization and subordinate-level identification.

Our experimental results argue against any stage-like process of basic-level categorization preceding subordinate-level identification. While categorization of typical objects from novice categories is faster at the basic level than the subordinate level in a speeded category verification task, no qualitative difference was observed in the time course of decisions in a signal-to-respond paradigm. Specifically, there was no observed delay in the onset of the speed-accuracy tradeoff function for identification relative to categorization.

As we noted in the introduction, many computational models of object recognition and object categorization assume no stage-like processing architecture (Joyce & Cottrell, 2004; Lamberts, 2000; Nosofsky & Kruschke, 1992; Nosofsky & Palmeri, 1997; Palmeri,

[4] We also conducted another experiment that used essentially the same stimuli and procedures and observed qualitatively the same results for non-face objects. The only difference between that study and the one reported here is that we did not include as many response signal delays, especially at the longer times; this gave us more trials per condition. However, while long response signals are unimportant for our primary question about the onset of the SATF, data from those long response signals is important for getting adequate fits of the asymptote parameter of the SATF, which is why we reported the present study instead of this one.

1997; Riesenhuber & Poggio, 2000). But if that is the case, why are subordinate-level identifications slower than basic-level categorizations in novice domains? And what happens when this difference goes away, as in expert domains or with atypical objects? In many models, basic-level categorization and subordinate-level identification are both perceptual decisions at the same stage of processing (Palmeri & Cottrell, 2008; Palmeri & Tarr, 2008). Some models explicitly propose that these perceptual decisions are made in prefrontal cortex (Riesenhuber & Poggio, 2000) or other brain areas (Ashby, Ennis, & Spiering, 2007) but that they are not made in visual cortex (Jiang, Bradley, Rini, Zeffiro, VanMeeter, & Riesenhuber, 2007). The speed of these perceptual decisions can be influenced by a variety of factors, such as the speed of perceptual processing (Lamberts, 2000), the ease or difficulty of discriminations (D'Lauro, Tanaka, & Curran, 2008), or the quality of the visual representations used to drive those perceptual decisions (Palmeri et al., 2004).

Indeed, in other work, we are exploring how one model of object categorization, the exemplar-based random walk (EBRW) model (Nosofsky & Palmeri, 1997; Palmeri, 1997), accounts naturally for both the basic-level advantage for novices and the entry-level shift with expertise, without assuming stage of processing for the basic level or any qualitative change in representations over learning (Mack, Wong, Gauthier, Tanaka, & Palmeri, 2007; see also Palmeri et al., 2004). Models like EBRW do not propose any qualitative reconfiguration with learning. Instead they assume gradual quantitative changes – a sharpening of representations over time (see also Jiang et al., 2006; Jiang et al., 2007). Yet these quantitative changes can give rise to qualitatively different patterns of results across novice and expert categorization (see also Joyce & Cottrell, 2004). Faster categorization is predicted without assuming different stages of processes. Fastest does not necessarily mean first.

Acknowledgments

This work was supported by the Temporal Dynamics of Learning Center NSF SBE-0542013, NSF Grant HSD-DHBS05, and a Grant from the James S. McDonnell Foundation.

References

Anaki, D., Zio-Golumbic, E., & Bentin, S. (2007). Electrophysiological neural mechanisms for detection, configural analysis and recognition of faces. *Neuroimage, 37*, 1407–1416.

Ashby, F. G., Ennis, J. M., & Spiering, B. J. (2007). A neurobiological theory of automaticity in perceptual categorization. *Psychological Review, 114*, 632–656.

Bowers, J. S., & Jones, K. W. (2008). Detecting objects is easier than categorizing them. *Quarterly Journal of Experimental Psychology, 61*(4), 552–557.

Bukach, C. M., Gauthier, I., & Tarr, M. J. (2006). Beyond faces and modularity: The power of an expertise framework. *Trends in Cognitive Sciences, 10*, 159–166.

Carrasco, M., McElree, B., Denisova, K., & Giordano, A. M. (2003). Speed of visual processing increases with eccentricity. *Nature Neuroscience, 6*, 699–700.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions Royal Society London, B, 335*, 95–102.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115*(3), 107–117.

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than children? *Visual Cognition, 1*, 253–274.

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

D'Lauro, C., Tanaka, J. W., & Curran, T. (2008). The preferred level of face categorization depends on discriminability. *Psychonomic Bulletin and Review, 15*, 623–629.

Dosher, B. A. (1981). The effects of delay and interference: A speed-accuracy study. *Cognitive Psychology, 13*(4), 551–582.

Gauthier, I., & Tarr, M. J. (1997). Becoming a 'Greeble' expert: Exploring mechanisms for face recognition. *Vision Research, 37*, 1673–1682.

Grill Spector, K., & Kanwisher, N. (2005). Visual recognition: As soon as you know it is there, you know what it is. *Psychological Science, 16*, 152–160.

Hintzman, D. L., Caulton, D. A., & Curran, T. (1994). Retrieval constraints and the mirror effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(2), 275–289.

Hintzman, D. L., & Curran, T. (1997). Comparing retrieval dynamics in recognition memory and lexical decision. *Journal of Experimental Psychology: General, 126*, 228–247.

Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron, 50*, 159–172.

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

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.

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

Joyce, C., & Cottrell, G. W. (2004). Solving the visual expertise mystery. In H. Bowman & C. Labiouse (Eds.), *Connectionist models of cognition and perception II: Proceedings of the eighth neural computation and psychology workshop*. World Scientific.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 22–44.

Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review, 107*, 227–260.

Lamberts, K., & Freeman, R. P. J. (1999). Building object representations from parts: Tests of a stochastic sampling model. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 904–926.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*(9), 910–916.

Mack, M. L., Wong, A. C. -N., Gauthier, I., Tanaka, J., & Palmeri, T. J. (2007). Unraveling the time course of perceptual categorization: Does fastest mean first? In Proceedings of the twenty-ninth annual meeting of the Cognitive Science Society (pp. 1253–1258).

Mack, M. L., Gauthier, I., Sadr, J., & Palmeri, T. J. (2008). Object detection and basic-level categorization: Sometimes you know it is there before you know what it is. *Psychonomic Bulletin and Review, 15*(1), 28–35.

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*, 8–15.

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.

Nosofsky, R. M. (1992). Exemplar-based approach to relating categorization, identification, and recognition. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition*. Hillsdale, NJ: Erlbaum.

Nosofsky, R. M., & Kruschke, J. K. (1992). Investigations of an exemplar-based connectionist model of category learning. *The Psychology of Learning and Motivation, 28*, 207–250.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*(2), 266–300.

Palmeri, T. J. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 324–354.

Palmeri, T. J., & Cottrell, G. (2008). Modeling perceptual expertise. In D. Bub, M. Tarr, & I. Gauthier (Eds.), *Perceptual expertise: Bridging brain and behavior*. Oxford University Press.

Palmeri, T. J., & Tarr (2008). Object recognition and long-term visual memory for objects. In S. Luck & A. Hollingsworth (Eds.), *Visual memory*. Oxford University Press.

Palmeri, T. J., Wong, A. C.-N., & Gauthier, I. (2004). Computational approaches to the development of perceptual expertise. *Trends in Cognitive Sciences, 8*(8), 378–386.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*, 446–461.

Reed, A. V. (1973). Speed-accuracy tradeoff in recognition memory. *Science, 181*, 574–576.

Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience, 3*(Suppl.), 1199–1204.

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.

Tanaka, J. W. (2001). The entry point of face recognition: Evidence for face expertise. *Journal of Experimental Psychology: General, 130*(3), 534–543.

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.

Thorpe, S., Fize, D., & Marlot, M. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

Vincent, S. B. (1912). The function of vibrissae in the behavior of the white rat. *Behavioral Monographs, 1* (Whole No. 5).

Wickelgren, W. A., & Corbett, A. T. (1977). Associative interference and retrieval dynamics in yes-no recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 3*(2), 189–202.