


The world we experience is formed by our perceptual processing. However, it is not viciously circular to argue that our perceptual systems are reciprocally formed by our experiences. In fact, it is because our experiences are necessarily based on our perceptual systems that these perceptual systems must be shaped so that our experiences are appropriate and useful for dealing with our world.

The notion that our perceptions are experience driven has been construed as everything from mundane to magical. At the mundane (at least, well-understood) pole, much is known of the mechanisms underlying simple sensitization and habituation. Through sensitization, repetitive presentation of a stimulus leads to increased sensory response to the stimulus. Through habituation, repetitive presentation of a stimulus leads to decreased responsiveness. The neurochemical circuitry underlying the sensitization of the aplysia's gill withdrawal response is well understood (Frank & Greenberg, 1994). Closer to the magical end of the continuum, scholars have argued for profound perceptual consequences of experience. Kuhn (1962) described how scientists, when exposed to a particular theoretical paradigm, see physical phenomena in new ways: “Though the world does not change with a change of paradigm, the scientist afterward works in a different world” (p. 121), and "during [scientific] revolutions, scientists see new and different things when looking with familiar instruments in places they have looked before" (p. 111).
Similarly, the new look movement in psychology (Bruner, Postman, & Rodriguez, 1951) and cultural anthropologists (Whorf, 1941/1956) have claimed that learned culture and experiences constrain our perceptions, partially blinding us to alternative perceptual descriptions that are inconsistent with our expectations.

**TUNED PERCEPTUAL SYSTEMS**

Before describing my laboratory’s experiments on and models of experience-induced perceptual learning, it is useful to remember why an adaptive perceptual system is beneficial. Flexibility is beneficial when the world is variable. If everyone were confronted with the same environment, and this environment remained unchanged millennium after millennium, then perceptual systems could become hardwired for this particular environment. These perceptual systems would be efficient because they are specifically tuned to the unchanging environment. At a first pass, this does describe human perceptual systems. Humans from different cultures possess highly similar apparatus for vision (Kay & McDaniel, 1978, but see Schafer, 1983 for some apparent exceptions). Many general environmental factors, such as color characteristics of sunlight, the position of the horizon, and the change in appearance that an approaching object makes, have all been mostly stable over the time that the human visual system has developed.

However, if we look more closely, there is an important sense in which different people face different environments. Namely, to a large extent, a person’s environment consists of animals, people, and things made by people. Animals and people have been designed by evolution to show variability, and artifacts vary widely across cultures. Evolutionary pressures may have been able to build a perceptual system that is generally adept at processing faces (Bruce, 1998; Farah, 1992), but they could not have hardwired a neural system that was adept at processing a particular face, such as John Kennedy’s, for the simple reason that there is too much generational variability among faces. Individual faces show variability from generation to generation, and variability is apparent over only slightly longer intervals for artifacts, words, ecological environments, and animal appearances. Thus, we can be virtually positive that tools show too much variability over time for there to be a hardwired detector for hammers. Words and languages vary too much for there to be a hardwired detector for the written letter A. Biological organisms are too geographically diverse for people to have formed a hardwired cow detector, for example. When environmental variability is high, the best strategy for an organism is to develop a general perceptual system that can adapt to its local conditions.

There is an even deeper sense in which people face different environments. People find themselves in different worlds because they choose to specialize. English-speaking people become specialized at hearing and seeing English words. People familiar with a particular race become specialized at recognizing faces from that race (Shapiro & Penrod, 1986). Experts at wine tasting, chick sexing, X-ray diagnosing, identical twin identifying, and baseball pitch judging all have unique perceptual capabilities because of the tasks they perform. Experts typically have highly specialized skills, many of which are perceptual (Sowden, Davies, & Roling, 2000). Moreover, the examples of word and face recognition suggest that every person has domains in which they show expertise. Even if all people confronted the same world initially, they would create distinctive communities with unique languages, artifacts, and objects of importance. In large part, individuals decide for themselves what objects they will be exposed to, and one’s identity is defined by the particular specialized niche one assumes. One’s niche will depend on many factors, including proclivity, community needs, local support, random accidents, and positive feedback loops. Thus, it is again advisable to build a perceptual system with the flexibility needed to support any one of a large number of niches.

This argument is that evolutionary forces drive a perceptual system to not only adapt but also to have mechanisms that allow for adaptation within an organism’s lifetime when the environmental stability required for hardwiring is not present. An information-theoretic perspective suggests additional computational advantages for perceptual systems that become tuned to organisms’ environments within their lifetime. In particular, perceptual systems can create highly efficient, compressed encodings of stimuli if the encodings are tuned to statistical structures across a set of stimuli. One particularly powerful method of doing this is to devise a set of elements that, if combined together, would generate the set of objects in one’s world. Part of the power from this approach stems from compositionality—creating representations by composing building blocks. An exceedingly large number of objects can be created from a surprisingly small set of building blocks if the building blocks can be combined in different arrangements. A set of 10 building blocks can be combined to form 10^5 objects that each contains 5 building blocks simply superimposed on top of each other. If there are eight different relations (i.e., left-of, right-of, above, below, inside, attached-to, behind, and in-front-of relations) that specify how these five building blocks within an object interrelate, then there would be 107,374,182,400,000 possible configurations of these five building blocks. The number of configurations that P parts, sampled from a set of S parts, create when potentially related to every other part in one of R ways is

$$R^{P(P-1)/2} S^P.$$  \hspace{1cm} (1)

A moral that should be drawn from this combinatorial explosion of potential configurations is that complex stimuli should often be represented in terms of parts and their relations (Biederman, 1987; Hummel, 2000; Markman, 1999), rather than in terms of nondecomposed configurations. This is particularly true when there are a large number of possible stimuli and the stimuli share often-repeated parts or relations. However, the moral that is typically drawn from these considerations is slightly different—that stimuli ought to be represented in terms of fixed and a priori
parts. A frequent strategy in cognitive science has been to represent entities in terms of hardwired primitive features. In linguistics, phonemes have been represented by the presence or absence of fewer than 12 features such as voiced, nasal, and strident (Jakobson, Fant, & Halle, 1963). Scenarios, such as ordering food in a restaurant, have been represented by Schank (1972) in terms of a set of 23 primitive concepts such as Physical-transfer, Propel, Grasp, and Ingest. In the field of object recognition, Biederman (1987) proposed a set of 36 geometric shapes such as wedge and cylinder to be used for representing objects such as telephones and flashlights. Wierzbicka (1992) proposed a set of 30 semantic primitives, including good, want, big, and time, to be composed together in structured phrases to generate all other words. To summarize, in this widespread and successful approach toward representation, the ability to represent new entities derives from combining a fixed set of a priori primitives, features, or parts in novel arrangements.

In what follows, I will argue that the building blocks an observer uses for constructing their world depends on the observer’s history, training, and acculturation. These factors, together with psychophysical constraints, mold one’s set of building blocks. The researchers who proposed fixed sets of hardwired primitives are exactly right in one sense—the combinatorics of objects, words, scenes, and scenarios strongly favor componential representations. However, this does not necessitate that the components be hardwired. By developing new components to subserve particular tasks and environments, a newly important discrimination can generate building blocks that are tailored for the discrimination. Adaptive building blocks are likely to be efficient because they can be optimized for idiosyncratic needs and environments.

A neural network that provides an excellent initial conception of how a system can acquire useful building blocks is the expectation maximization (EM) algorithm for factorial learning (Dempster, Laird, & Rubin, 1977; Ghahramani, 1995; Hinton, Dayan, Frey, & Neal, 1995; Tenenbaum, 1996). This approach considers the statistics of component parts across a set of inputs. When presented with a set of inputs, this approach finds an underlying set of independent components that, when combined in different arrangements, reproduces the set of inputs. For example, Ghahramani (1995) created a set of 160 patterns by combining a horizontal line in any one of four positions with a vertical line in any one of four positions, together with added noise. The complete set of patterns is shown in Fig. 7.1. From these patterns, an EM algorithm could generate the set of eight horizontal and vertical lines that suffice for generating the 160 patterns. It does so by finding the weightings of different hidden dimensions (such as a bar) that would be most likely to have produced the 160 patterns. The algorithm is able to discover both the hidden dimensions and their weightings by iterating between two steps: (1) computing the expected hidden dimensions given the current weights and (2) maximizing the likelihood of the weights given these expected dimensions. Finding decompositions into parts is a useful enterprise in terms of information demands. Imagine wanting to represent each of the 16 basic configurations of four

horizontal and four vertical lines. Rather than creating 16 different complex configurations, the componential representation requires only eight simpler parts to be stored. Although the storage requirements are only halved in this case, the savings becomes far greater as the number of composed dimensions and the number of values per dimension increases. The neural network modeling that I later report will differ from the EM algorithm in several ways. The network will develop building blocks that are modulated online by categorization feedback and perceptual constraints. Still, my later modeling borrows the basic insight from this application of the EM algorithm. It is efficient to represent objects in terms of building blocks, particularly if exactly those building blocks can be discovered that would provide compressed encodings of the objects.

Considering perceptual learning from an information theoretic perspective, this claim about the advantages of learned componential representations should be qualified. Whether or not the objects comprising an observer’s world are efficiently represented by composing together parts or by nondecomposed, configurational units will depend on the amount and kinds of similarity shared by objects. A well-spaced set of objects that uniformly covers an abstract stimulus space will typically be efficiently represented by a compositional set of parts. It is the factorial combination of horizontal and vertical parts in Fig. 7.1 that makes it useful to extract them out as
parts. If each horizontal line occurred with only one particular vertical line, then it would be more efficient to represent only the four pairings of two lines rather than eight individual lines (Edelman, 1982; Kohnen, 1995). As such, the distribution of objects can inform whether the objects ought to be represented in terms of parts or holistically.

The choice between compositional and holistic representations is also influenced by how the objects will be used. Parts may need to be extracted if they predict an important categorization or use of the object (Gibson, 1969). Conversely, if a single complex configuration dictates the use of an object, then there should be a bias to represent the configuration holistically rather than by its parts. By analogy, if the efficacy of a medicine in treating a skin disease depends only on the skin’s dryness, then dryness should be extracted as an isolated component for representing patients. However, if the medicine is only effective for a particular combination of dryness, body location, radius, and patient age, then there should be a bias to connect together these diverse attributes into a single functional feature. The general principle for creating efficient codes for stimuli is to find the level of representation that captures the systematicities present between and within objects.

**VARIETIES OF PERCEPTUAL LEARNING**

In arguing for flexible tuning of perceptual systems, one natural question to ask is, “Is the adaptation truly perceptual, or does it perhaps occur at postperceptual stages?” I have argued elsewhere for a deliberate blurring of the boundary between perception and cognition (Goldstone & Barsalou, 1998). There are deep similarities between perceptual unitization and chunking in memory and between perceptual differentiation and association building (Hall, 1991). Although downplaying a definite boundary where perception leaves off and true cognition begins, it is still useful to highlight learned changes in processing that involve adjusting initial representations of stimuli. Perceptual learning exerts a profound influence on behavior precisely because it occurs early during information processing and thus shifts the foundation for all subsequent processes.

The neurocognitive evidence fails to support the notion that cortical areas involved with early perceptual processes are fairly fixed, context insensitive, and stable, whereas more central cortical processes show most of the lability. Consistent with an early stage of adaptation, Weinberger (1993) describes evidence that cells in the primary auditory cortex become tuned to the frequency of often-repeated tones. Practice in discriminating among small visual motions in different directions significantly alters electrical brain potentials that occur within 100 ms of the stimulus onset (Fahle & Morgan, 1996). These electrical changes are centered over the primary visual cortex, suggesting plasticity in early visual processing. Karni and Sagi (1991) find evidence, based on the specificity of training to eye (interocular transfer does not occur) and retinal location, that is consistent with early.

primary visual cortex adaptation in simple discrimination tasks (see also Fahle & Morgan, 1996; Shiu & Pashler, 1992). Changes in the primary somatosensory areas in the cortex are observed when environmental or cortical changes occur (Garraghty & Kaas, 1992). When cortical areas are lesioned, neighboring areas newly respond to sensory information formerly controlled by the lesioned area; when external sensory organs are disabled, cortical areas formerly activated by the organ become sensitive to sensory stimulation formerly controlled by its neighboring areas (Kaas, 1991). When two fingers are surgically fused, creating highly correlated inputs, a large number of cortical areas develop that respond to both fingers (Allard, Clark, Jenkins, & Merzenich, 1991). In sum, there is an impressive amount of converging evidence that training leads to changes to very early stages of information processing.

This neurophysiological evidence makes a good case for relatively early changes to perception with laboratory training or general experience. This case having been briefly made, the evidence that follows focuses not on the locus of changes within the brain or the stage of information processing but on the mechanisms that underlie some of these changes. I describe these mechanisms at a functional rather than concrete level. Even though neurocognitive details are known in some cases, a functional level of description is appropriate for interpreting the results from cognitive experiments and for unifying the accounts of human and computational learning. The following mechanisms of perceptual change are not exhaustive (a more complete organization is provided by Goldstone, 1998). Instead, mechanisms of perceptual change are described that will have particular relevance to a subsequently presented model of the interaction between perceptual and conceptual change.

**Imprinting**

By imprinting, perceptual detectors or receptors are developed that are specialized for stimuli or parts of a stimulus. Internalized detectors develop for repeated stimuli, and these detectors increase the speed, accuracy, and general fluency with which the stimuli are processed. In computational learning terms, imprinting is an unsupervised learning mechanism in that there is no need for a parent, teacher, or programmer to tell the learner what a stimulus is called. Learning proceeds by adapting internal detectors to become better specialized for the imprinting stimuli. The advantages of developing a detector that increasingly resembles an input stimulus are that small deviations from the input stimulus can be easily detected, particularly, strong responses can be quickly produced when the stimulus is presented again, and shorter codes can represent often-presented stimuli. Detectors may imprint on either a whole stimulus or a subset of the whole stimulus.

Imprinting on an entire stimulus is tantamount to developing a functional detector that is specialized for a specific stimulus. One of the most common paradigms
for studying this has been to show subjects a briefly presented stimulus and ask them to identify it. Such stimuli are more accurately identified when the subject has been previously exposed to them (Schacter, 1987). The identification advantage for previously familiarized instances lasts at least 3 weeks, is found for both words and pictures, requires as few as one previous presentation of an item, and is often tied to the specific physical properties of the item during its initial exposure. Several models in psychology and computer science have exploited the power of storing individual exposures to stimuli in a relatively raw, unabstracted form. Exemplar, instance-based, view-based, case-based, nearest neighbor, configural cue, and vector quantization models all share the fundamental insight that novel patterns can be identified, recognized, or categorized by giving the novel patterns the same response that was learned for similar, previously presented patterns. By creating detectors for presented patterns not only is it possible to respond to repetitions of these patterns, but it is also possible to give novel patterns responses that are likely to be correct by sampling responses to old patterns, weighted by their similarity to the novel pattern. Consistent with these models, psychological evidence suggests that people show good transfer to a new stimulus in perceptual tasks just to the extent that the new stimulus superficially resembles previously learned stimuli (Kolers & Rodiger, 1984; Palmeri, 1997).

In addition to imprinting on entire stimuli, unsupervised information from the environment can also lead to imprinting on selected parts or features from a stimulus. Whereas imprinting to whole stimuli supports configural, holistic processing, imprinting to parts supports componential processing of a stimulus by piecing it together out of a perceptual vocabulary. If a stimulus part varies independently of other parts, or occurs frequently, people may develop a specialized detector for that part. Schyns and Rodet (1997) found unfamiliar parts (arbitrary curved shapes within an object) that are important in one task are more likely to be used to represent subsequent categories. Their subjects were more likely to represent a conjunction of two parts, X and Y, in terms of these two components (rather than as a whole unit or a unit broken down into different parts) when they received previous experience with X as a defining part for a different category. Neural networks have proved to be highly effective at creating building blocks for describing sets of stimuli. In addition to the previously described work by Ghahramani (1995), researchers have discovered that simple exposure to photographs of natural scenes suffices to allow neural networks to create a repertoire of oriented line segments to be used to describe the scenes (Miikulainen, Bednar, Choe, & Sirosh, 1997; Obermayer, Sejnowski, & Blasdel, 1995; Schmidhuber, Eldracher, & Foltin, 1996). These feature detectors bear a strong resemblance to neural detectors found in the primary visual cortex and are created by learning algorithms that develop units that respond to independent sources of regularity across photographs.

Human and computer systems that develop perceptual vocabulary elements based on the objects in their environment have an advantage over systems that employ only fixed features. One difficulty with fixed sets of features is that it is hard to choose exactly the right features that will suffice to accommodate all possible future entities. On the one hand, if a small set of primitive elements is chosen, then it is likely that two entities will eventually arise that must be distinguished but cannot with any combination of available primitives. On the other hand, if a set of primitives is sufficiently large to accommodate all entities that might occur, then it will likely include many elements that lie unused, waiting for their moment of need to possibly arise (Schyns, Goldstone, & Thibaut, 1998). If entities are represented by a smaller set of fixed features that are composed as building blocks in different arrangements to accommodate a wide range of entities, then very small building blocks will be required, resulting in large and inefficient representations of whole entities. However, by developing new components as needed, newly important discriminations can cause the construction of detectors that are tailored for the discrimination.

### Task-Dependent Perceptual Learning

In imprinting, the stimuli themselves shape the perceptual change and can do so irrespective of what must be done with the stimuli. Not only is the perceptual change unsupervised in the sense of not requiring a label or response to be associated with the stimuli, but it is task independent in the more general sense of proceeding without any regard for context or the observer's goals. However, in many situations, what is to be done with a stimulus influences what information is extracted from the stimulus (Gibson, 1969; Goldstone, Lippa, & Shiffrin, 2001; Pick, 1992).

People can dynamically shift their attention to different stimulus features depending on their perceived importance. One way that perception becomes adapted to tasks and environments is by increasing the attention paid to perceptual features that are important, or by decreasing attention to irrelevant dimensions and features, or both. Whether accomplished by moving one's eyes or more covertly, shifts of visual attention play an important role in allowing us to learn needed categories. As such, it comes as no surprise that most successful theories of categorization and learning incorporate selective attention. In Sutherland and Mackintosh's (1971) analyzer theory, learning a discrimination involves strengthening the tendency to attend to relevant analyzers. In Nosofsky's (1986) exemplar model of categorization, the categorization of an object depends on its similarity to previously stored category members in a multidimensional space (see also Medin & Shaffer, 1978). Critically, distances between objects are compressed and expanded along dimensions in this space depending on the categorization required. Distances between objects on relevant dimensions are expanded, and irrelevant dimensions are compressed. For example, Nosofsky (1986) found that if participants are given a categorization where the angle of a line embedded in a circular form is relevant while the size of the circular form is irrelevant, then distances between objects on the angle dimension are increased and distances on the size dimension are decreased.
The next two mechanisms of perceptual learning I describe produce dynamic changes to the sizes of perceptual units used in representing an object. One of the mechanisms creates larger units out of smaller units, and the other mechanism divides a larger unit into smaller units. However, before describing these mechanisms, it is important to be clear about what it means for a system to have a perceptual unit. What could it mean for a system to learn a new feature? When writing of features, I am concerned with psychologically internalized rather than external features of an object. This is the only possible strategy because the only awareness of an object that we have is necessarily filtered through our perceptual system. The only way an objective property of an object can influence us is if it influences a psychologically represented feature. Even if a physicist can measure the illuminance of an object or a chemist can measure the tannin content of a Bordeaux, these stimulus properties are not psychological features unless the perceiving organism can isolate them as well. A psychological feature, then, is a set of stimulus elements that are responded to together, as an integrated unit. That is, a feature is a package of stimulus elements that is separated from other sets of elements and reflects the subjective organization of the whole stimulus into components.

In my usage of the term feature, I purposefully encompass several terms typically treated as distinct: features, dimensions, and parts. Features are typically understood as qualitative stimulus values that are either present or absent. Dimensions are stimulus values that have quantitative levels. Parts are spatially localized regions of a stimulus. One reason to define features as psychologically isolated packages of stimulus elements with the intention of referring to all three types of stimulus organization is that a clear distinction between these organizations is often hard to justify. For example, "has wings" is typically described as a feature of birds, but wings are also spatially localized parts. Does a red patch on an object count as a part but suddenly become a feature if it covers the entire object? Does the difference between a circle and a square suddenly shift from being featural to dimensional as soon as shapes with intermediate levels of curvature are introduced? A unified account of all three structures is desirable.

This definition corresponds closely to standard operationalizations of features. Garner (1976) measures the extent to which two stimulus elements are psychologically fused by determining how much categorizations made on the basis of one of the elements are slowed by irrelevant variation on the other. Treisman and Gelade (1980) argue for a vocabulary of features that includes color, orientation, and closure elements. Their empirical task involves visually searching for a target defined by a simple feature or conjunction of features, among a set of distractor objects. Features are empirically identified by response times to targets that are not influenced by the number of distractors if the targets are distinguished by a feature. Tests for featurehood typically involve showing that some information is dependent on other information (in the same feature) but is independent of still other information (in different features). For feature learning to occur simply means that the organization of elements into sets has changed. For example, although the saturation and brightness of a color are not strong psychological features for most people, they may be for color experts (Burns & Shepp, 1988), if the color expert can demonstrate that they can attend to saturation without being very influenced by brightness. In the modeling presented later, features are implemented as acquired detectors in a hidden layer of a neural network model, and two stimulus elements belong to the same feature to the extent that a single unit responds selectively to those elements. Thus, there need not be anything more magical about the development of novel features than reorganizing functional detectors so that they are influenced by only some sources of information.

**Unitization.** One result of category learning is to create perceptual units that combine stimulus components that are useful for the categorization. Such a process is one variety of the more general phenomenon of unitization, by which single functional units are constructed and triggered when a complex configuration arises. Cattell (1886) invoked the notion of perceptual unitization to account for the advantage that he found for tachistoscopically presented words relative to nonwords. Gestalt psychologists proposed the perceptual principle that objects tend to be perceived in terms of components that have acquired familiarity (Koffka, 1935). Weisstein and Harris (1974) found that briefly flashed line segments are more accurately identified when they are part of a set of lines forming a unitary object rather than an incoherent pattern. They interpreted this effect as showing that arrangements of lines can form configural patterns that are perceived before the individual lines are perceived. More recently, Gauthier and Tarr (1997; see also Gauthier, Williams, Tarr, & Tanaka, 1998) found that prolonged experience with a novel object leads to a configural representation of it that combines all of its parts into a single, viewpoint-specific, functional unit. Their evidence for such a representation is that recognition of these familiarized objects improved considerably with practice and was much more efficient when the object was in its customary upright form rather than inverted.

Unitization has also been explored in the field of attention. Using a task where participants decided whether or not two visual objects were identical, LaBerge (1973) found that when stimuli were unexpected, participants were faster at responding to actual letters than to letterlike controls. Furthermore, this difference diminished as the unfamiliar letterlike stimuli became more familiar over practice. He argued that the shape components of often-presented stimuli become processed as a single functional unit with practice. More recently, Czervinski, Lightfoot, and Shiffrin (1992) referred to a process of perceptual unitization in which conjunctions of stimulus features are chunked together so that they become perceived as a single unit. Shiffrin and Lightfoot (1997) argued that separated line segments can become unitized following prolonged practice with the materials. Their evidence comes from the slopes relating the number of distractor elements to response time in a feature search task. When participants learned a conjunctive search task in
which three line segments were needed to distinguish the target from distracters, impressive and prolonged decreases in search slopes were observed over 20 hour-long sessions. These prolonged decreases were not observed for a simple search task requiring attention to only one component.

Unification is also important during the development of object perception. Newborn infants fail to integrate the separate regions of an object that is occluded (Slater et al., 1990). However, by 4.5 months of age, babies form the same interpretation of displays whether they are fully visible or occluded (Johnson, 1997; Needham, 1997, this volume). This developed ability to integrate different parts into a single object representation depends on the featural similarity of these parts (Needham, 1999).

I recently completed experiments designed to test whether category learning can lead to stimulus unification and to explore the boundary conditions on unification related to stimulus characteristics and amount of training (Goldstone, 2000). Whenever the claim for the construction of new units is made, two objections must be addressed. First, perhaps people already possess the unit before categorization training. My stimuli were designed to make this explanation unlikely. Each unit to be sensitized was constructed by connecting five randomly chosen curves. There were 10 curves that could be sampled without replacement, yielding 30,240 (10 × 9 × 8 × 7 × 6) possible different units, an implausibly large number under the constraint that all units preexisted. The second objection is that no units need be formed; instead, people analytically integrate evidence from the five separate curves to make their categorizations. However, this objection was shown to be untenable because subjects, at the end of 20 hours of training, were faster at categorizing the five-element units than would be expected by an analytic approach. In this analytic account, the response time required to respond to a unit was predicted by the response times required to respond to individual components within the unit. The response-time modeling involved Fourier deconvolutions for isolating component processes within response-time distributions and is too involved to describe here (for discussions, see Goldstone, 2000, or Goldstone, Steyvers, Spencer-Smith, & Kersten, 2000). For present purposes, I simply treat pronounced response time improvements to respond to a multicomponent unit as suggestive evidence for unification, setting aside the more elaborate analyses.

In the experiments, subjects saw objects appear on the screen and were required to categorize them as quickly as possible. The categorization was designed so that evidence for five components had to be received before certain categorization responses were made. The stimuli and their category memberships are shown in Fig. 7.2. Each of the letters refers to a particular segment of a doodle. Each doodle was composed of five segments, with a semicircle below the segments added to create a closed figure. To reliably place the doodle labeled “ABCDE” into Category 1, all five components, “A,” “B,” “C,” “D,” and “E,” must be processed. For example, if the right-most component were not attended, then ABCDE could not be distinguished from ABCDZ, which belongs in Category 2. Not only does no single component suffice for accurate categorization of ABCDE, but two-way, three-way, and four-way conjunctions of components also do not suffice.

If unification occurs during categorization, then the stimulus ABCDE should become efficiently processed with practice. It is the only object that belongs to Category 1, and if reliably classify this item the entire object must be processed. Results from a speeded categorization task, shown in Fig. 7.3, show exactly this pattern. When all five components must be attended to reliably categorize “ABCDE” as belonging in Category 1, then there is a large improvement in speed over the course of four blocks and 1.5 hours (see the ordered-five condition in Fig. 7.3). This large improvement was not found when only one of the components was required to make a Category 1 response. In the ordered-one task, only one object, such as ABCDZ, belongs in Category 2, so detecting component “E” is sufficient for making a Category 1 response. The large improvement in speed was also not found when each of the five components of ABCDE needed to be identified but were randomly ordered. That is, ABCDE, DEBAC, and any other permutation of the five components A–E, were treated as equivalent. In this condition, a single photograph-like image cannot serve to represent ABCDE because many different images produce this pattern. The lack of impressive improvement in this random-five condition suggests that components that are not combined in a consistent manner to create a coherent image are not effectively unitized. Thus, the first constraint on the unitization process, illustrated in the top panel of Fig. 7.4, is that unitization only proceeds when a photograph-like template can be constructed for
the unit. Unitization apparently consists in establishing templates for patterns that are diagnostic for categorization.

The nature of these templates was tested in a second experiment that compared speed improvements for detecting conjunctions of five components that were either separated or spatially contiguous. The same components were used for the contiguous and separated displays, and the distance between the endpoints was equated, but in the separated condition the components were stacked on top of each other so that they no longer touched (as shown in the second panel of Fig. 7.4). As shown in Fig. 7.5, there was a small advantage found for contiguous over separated displays, found when either all five or only one component(s) needed to be identified for a Category 1 response. However, both the separated-five and contiguous-five conditions improved dramatically over blocks of practice, with almost equal rates of improvement for each. As such, the unitization process is not constrained to form images for contiguous images. Separating the components of an image does not interfere with unitization in the manner that randomly permuting the components did.

However, the third panel in Fig. 7.4 shows a constraint on people's ability to unitize across discontinuous components. In a third experiment, a conjunction of two (rather than five) components was critical for categorization. These two components were either contiguous to each other or were separated by a third irrelevant component. More precisely, in the Together condition, ABCDE belonged to Category 1, and ABCDZ and ABCYE belonged to Category 2. In the Apart condition, ABCDE belonged to Category 1, and ABCDZ and ABYDE belonged to Category 2, and, hence, Components C and E were critical for making a Category 1 response. Category 1 response times improved much more quickly in the Together than in the Apart condition. In fact, the pattern of improvement in the Apart condition closely resembled the improvement shown when a conjunction of three components was required. Apparently, in the Apart condition, subjects form a unit by combining the two disconnected components and the irrelevant component between them. This result can be reconciled with the previous paragraph by hypothesizing that units are formed for components within a single spatially
Impressively complex units can be formed, but as unit complexity increases, so does the time required for unitization. The important implicit message from this result is that unitization is not like taking photographic film to be developed. Although the end result of each may be similar, an image-like template recording an environmental input, the process by which these results are obtained is dissimilar. A photograph of a crowd takes no longer to develop than a photograph of a single face, but for human unitization development time is proportional to stimulus complexity. Once constructed, units of different complexity may be identified almost equally quickly, but this belies the large differences in time required for constructing the units.

**Dimension Differentiation.** Selective attention, described earlier, is a critical component of adaptive learning, but it may not be the only process that dynamically alters the description of an object in a categorization task. A second candidate process is dimension differentiation, by which dimensions that are originally psychologically fused together become separated and isolated. Selective attention presumes that the different dimensions that make up a stimulus can be selectively attended. To increase attention to size but not color, one must be able to isolate size differences from color differences. In his classic research on stimulus integrality and separability, Garner argued that stimulus dimensions differ in how easily they can be isolated or extracted from each other (Garner, 1976, 1978). Dimensions are said to be separable if it is possible to attend to one of the dimensions without attending to the other. Size and brightness are classic examples of separable dimensions; making a categorization on the basis of size is not significantly slowed if there is irrelevant variation on brightness. Dimensions are integral if variation along an irrelevant dimension cannot be ignored when trying to attend to a relevant dimension. The classic examples of integral dimensions are saturation and brightness, where saturation is related to the amount of white mixed into a color, and brightness is related to the amount of light coming off of a color. For saturation and brightness, it is difficult to attend to only one of the dimensions (Burns & Shepp, 1988; Melara, Marks, & Potts, 1993).

From this work distinguishing integral from separate dimensions, one might conclude that selective attention can proceed with separable but not integral dimensions. However, one interesting possibility is that category learning can, to some extent, change the status of dimensions, transforming dimensions that were originally integral into more separable dimensions. Experience may change the underlying representation of a pair of dimensions such that they come to be treated as relatively independent and noninterfering sources of variation that compose an object. Seeing that stimuli in a set vary along two orthogonal dimensions may allow the dimensions to be teased apart and isolated, particularly if the two dimensions are differentially diagnostic for categorization. There is developmental evidence that dimensions that are easily isolated by adults, such as the brightness and size of a square, are treated as fused together for 4-year old children (Kemler & Smith,
1978). It is relatively difficult for children to decide whether two objects are identical on a particular dimension but relatively easy for them to decide whether they are similar across many dimensions (Smith, 1989). Children show considerable difficulty in tasks that require selective attention to one dimension while ignoring another, even if the dimensions are separable for adults (Smith & Kemler, 1978). For example, children seem to be distracted by shape differences when they are instructed to make comparisons based on color. Adjectives that refer to single dimensions are learned by children relatively slowly compared with nouns (Smith, Gasser, & Sandhofer, 1997).

The developmental trend toward increasingly differentiated dimensions is echoed by adult training studies. Under certain circumstances, color experts (art students and vision scientists) are better able to selectively attend to dimensions (e.g., hue, chroma, and value) that comprise color than are nonexperts (Burns & Shepp, 1988). Goldstone (1994a) showed that people who learn a categorization in which saturation is relevant and brightness is irrelevant (or vice versa) can learn to perform the categorization accurately, and as a result of category learning, they develop selectively heightened sensitivity at making saturation, relative to brightness, discriminations. That is, categorization training that makes one dimension diagnostic and another dimension nondiagnostic can serve to split apart these dimensions, even if they are traditionally considered to be integral dimensions. These training studies show that to know how integral two dimensions are, one has to know something about the observer’s history.

Goldstone and Steyvers (2001) recently explored whether genuinely arbitrary dimensions can become isolated from each other. They explored dimension differentiation by a category learning and transfer paradigm. Their subjects first learned to categorize a set of 16 faces into two groups by receiving feedback from a computer and then were transferred to a second categorization. The stimuli varied along arbitrary dimensions that were created by morphing between randomly paired faces. As shown in Fig. 7.6, Dimension A was formed by gradually blending from Face 1 to Face 2, and Dimension B was formed by gradually blending from Face 3 to Face 4. Using a technique described by Steyvers (1999), a set of faces can be created from these two dimensions such that each face is defined half by its value on Dimension A and half by its value on Dimension B. Dimensions are thus formed by creating negative contingencies between two faces—the more of Face A that is present in a particular morphed face, the less of Face B there will be. The 4 × 4 matrix of faces in Fig. 7.6 shows how a set of faces can be constructed that varies independently along the two arbitrary dimensions. However, using the same dimensions, one can create a set of stimuli with even less intrinsic dimensionality by assigning 16 faces to coordinates that fall on a circle, rather than a grid, in the abstract space defined by Dimensions A and B. A variable $D$ was created and assigned 16 different values, from 0 to 360 in 22.5° steps. For each value assigned to $D$, the Dimension A value for a face was equal to $\cosine(D)$ and the
Dimension B value was \( \sin(D) \). The end result, shown in Fig. 7.7, is a set of faces organized on a circle with no privileged dimensional axes suggested by the set of faces.

With these faces, Goldstone and Steyvers (2001) asked whether the organization of the faces into dimensions could be influenced by the required categorization. Subjects were shown faces, asked to categorize them, and then received feedback on the correctness of their categorization. The categorization rules all involved splitting the 16 faces into two equal piles with straight lines, such as those shown in Fig. 7.7. Each subject was given training with one categorization rule and then was transferred to a second rule. The critical experimental manipulation was whether the final categorization rule involved a rotation of 45° or 90° relative to the initial rule. Given that the initial categorization rules were randomly selected, the only difference between the 45° and 90° rotation conditions was whether the category boundary was shifted by two or four faces. The results from four similar experiments indicate that in the second phase of category learning, there was an advantage for the 90° over 45° rotation condition in all of the conditions that involved integral (psychologically fused) dimensions. This is somewhat surprising given that categorizations related by 90° are completely incompatible as far as their selective attention demands. The dimension that was originally completely irrelevant becomes completely relevant. In the 45° condition, the originally relevant dimension is at least partially relevant later. However, categorizations related by 90° do have an advantage as far as dimensional organization. The relevant dimensions for the two categorizations are compatible with each other in the sense of relying on independent sources of variation. For example, acquiring Dimension A in Fig. 7.6 is compatible with later acquiring Dimension B because both are independent dimensions that can coexist without interference. Learning that Dimension A is relevant and Dimension B is irrelevant should encourage isolating Dimension A from Dimension B to only attend to Dimension A. This is exactly the same isolation of dimensions that is useful for learning that Dimension B is relevant and Dimension A is irrelevant, albeit for opposite reasons. Thus, categorization rules separated by 90° are completely inconsistent with respect to their selective attention demands but are consistent with respect to their dimensional organization of stimuli.

Understanding how dimensions related by 90° are organizationally compatible may be difficult with the abstract dimensions of Fig. 7.7. However, consider representing the rectangles in Fig. 7.8. Categorizing rectangles on the basis of height is compatible with categorizing them on the basis of width because these two dimensions can each be separately registered and do not interfere with each other. Someone who thought about rectangles in terms of height would also be likely to think about them in terms of width. Organizing rectangles in terms of shape (ratio of width to height) and area is an alternative dimensional organization. A person who thinks in terms of rectangle shape might also be expected to think in terms of area because this is the remaining dimension along which rectangles vary.

**Fig. 7.7.** Stimuli from Goldstone and Steyvers (2001). Experiment 3. Using the same dimensions as Fig. 7.6, a set of circularly arranged faces was created by varying degrees from 0° to 360°, assigning the face a value on Dimension A based on the cosine of the degrees, and assigning the face's Dimension B value based on the sine of the degrees. Subjects learned two successive categorizations involving the faces. Each categorization split the faces into two equal groups with a straight dividing line. The two category boundaries given to a subject were either related to each other by 45° or 90°.
7. Learning to Perceive

frequently, with all parts indicating a similar response. Thus, unitization and differentiation are both processes that build appropriate sized representations for the tasks at hand.

Pevtzow and Goldstone (1994; also see Goldstone et al., 2000) reported a series of experiments illustrating the compatibility of unitization and differentiation by showing that both processes are probably involved when learning how to segment pictures into component parts. People often spontaneously segment objects into parts, thereby organizing their world. Palmer conducted several studies on the naturalness of parts within whole objects, exploring factors that make certain parts more natural than others (Palmer, 1977, 1978). Palmer also developed a quantitative model of part naturalness that included a number of objective factors about the parts and whole: how close the line segments within a part are to each other, whether they formed closed objects, whether they have similar orientations, and whether the line segments of a part are similar to line segments within other parts.

In addition to these objective properties that determine how psychologically natural a particular segmentation of an object into parts will be, researchers have found that segmentations also depend on subjective experience. Behrmann, Zemel, and Mozer (1998) found that judgments about whether two parts had the same number of humps were faster when the two parts belonged to the same object rather than different objects. Further work found an influence of experience on subsequent part comparisons. Two stimulus components are interpreted as belonging to the same object if they have cooccurred many times (Zemel, Behrmann, Mozer, & Bavelier, 1999). Pursuing a similar line of inquiry, Pevtzow and Goldstone (1994) explored the influence of category learning on segmentation using materials based on Palmer’s stick figures and shown in Fig. 7.9. Naturalness was measured by how quickly subjects could confirm that a part was contained within a whole object (Palmer, 1978). To test the conjecture that how psychologically natural a part is depends on whether it has been useful for previous categorizations, we gave participants a categorization task, followed by part/whole judgments. During categorization, participants were shown distortions of the objects A, B, C, and D, shown in Fig. 7.9. The objects were distorted by adding a random line segment that connected to the five segments already present. Subjects were given extended training with either a vertical or horizontal categorization rule. For participants who learned that A and C were in one category and B and D were in another (a vertical categorization rule), the two component parts at the bottom of Fig. 7.9 were diagnostic. For participants who learned that A and B belonged in one category and C and D belonged to the other category (a horizontal rule), the components on the right were diagnostic.

During part/whole judgments, participants were shown a whole, and then a part, and were asked whether the part was contained in the whole. Participants were given both present and absent judgments, and examples of these judgments are shown in Fig. 7.10. Note that the two parts shown in Fig. 7.10 were both

![Table](https://via.placeholder.com/150)

**Table:** Compatibility of Dimensions

<table>
<thead>
<tr>
<th>Height</th>
<th>Width</th>
<th>Shape</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1:1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2:1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1:1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1:2</td>
<td>2</td>
</tr>
</tbody>
</table>

*Fig. 7.8.* An illustration of how stimuli can be understood in terms of competing dimensional organizations. Rectangles can be understood in terms of height and width or in terms of their area and shape. An analysis in terms of width is compatible with an analysis in terms of height but is incompatible with an analysis in terms of shape because width and shape are cross-cutting, rather than nonoverlapping, dimensions of variation.

Once shape has been extracted. However, organizing rectangles in terms of height is incompatible with organizing them in terms of area because area is partially dependent on height. In assessing the compatibility of dimensional organizations, two dimensions are compatible if they involve independent sources of variation and are incompatible if they do not. Dimensional axes related by 45° involve cross-cutting and, hence incompatible, dimensions. In summary, there appears to be more to learning than learning to selectively attend to existing dimensions. Perceptual learning also involves establishing which dimensions are available for selective attention. People learn not only appropriate weights for dimensions but also learn how to learn appropriate weights for dimensions.

**Reconciling Unitization and Dimension Differentiation.** The two previous sections described apparently contrary ways in which category learning affects perceptual learning. Unitization involves the construction of a single functional unit out of component parts. Dimension differentiation divides wholes into relatively separable component dimensions. There is an apparent contradiction between experience creating larger chunks via unitization and dividing an object into more clearly delineated components via differentiation. This incongruity can be transformed into a commonality at a more abstract level. Both mechanisms depend on the requirements established by tasks and stimuli. Objects will tend to be decomposed into their parts if the parts reflect independent sources of variation or if the parts differ in their relevancy. Parts will tend to be unitized if they co-occur...
potentially diagnostic during the earlier categorization training. Whether or not a part was diagnostic was independent of the appearance of the part itself, depending only on how the four objects of Fig. 7.9 were grouped into categories.

The major result, shown in Fig. 7.11, was that subjects were faster to correctly respond “present” when the part was diagnostic than when it was nondiagnostic (for a related result, see Lin & Murphy, 1997). To the extent that one can find response-time analogs of signal detection theory sensitivity and bias, this effect seems to be a sensitivity difference rather than a bias difference because absent judgments also tended to be faster for diagnostic than nondiagnostic parts. These results indicate that it is not simply the physical stimulus properties that determine how readily a person can segment an object into a particular set of components; segmentation is also influenced by the learned categorical diagnosticity of the components.

This experiment raises the strong possibility that the processes of composing parts into a single functional unit and differentiating a unit into several functional parts may be aspects of the same mechanism. Acquiring a diagnostic component involves both integrating its parts together into a single unit and separating this unit from other components. Both aspects are evident in the Pevtzow and Goldstone (1994) experiments because the units that are built up during category learning are not the entire five-line segment stimulus that is present on a trial. As such, the constructed units are complex configurations of multiple line segments yet are still only part of the entire stimulus. Both differentiation of a stimulus into parts and unitization could be incorporated in a model that begins with a specific featural description of objects, creates units for conjunctions of features if the features frequently occur together, and divides a feature into subfeatures if independent sources of variation in the original feature are detected. Describing such a model is the task now at hand.

MODELING INTERACTIONS BETWEEN CONCEPTUAL AND PERCEPTUAL LEARNING

In developing a computational model for perceptual learning, I have been drawn to neural networks that possess units that intervene between inputs and outputs and are capable of creating internal representations. For current purposes, these units can be interpreted as learned feature detectors and represent the organism's
acquired perceptual vocabulary. The immediate goal of the modeling described is explaining the qualitative behavior that was observed in the experiments previously described. However, more general attributes of the model are also described and are relevant for information compression and adaptive coding. The specific goals of the modeling are to explain how (1) category learning alters the perceptual vocabulary that is derived for a set of objects, (2) perceptual descriptions are a joint function of category demands and perceptual constraints, (3) category learning can influence the subsequent segmentation of presented objects into parts, (4) detectors are built for complex configurations of parts (unitization), and (5) a perceptually constrained set of building blocks can be created in an online fashion to efficiently represent a set of objects (differentiation).

First, I describe a mutually interacting pair of neural networks that are designed to categorize and segment patterns, with an eye toward accounting for Pevtzow and Goldstone’s (1994) results. As with the experiment, the network is first given categorization training and then is given a subsequent segmentation task, using the same network weights. The goal of the modeling is to show how categorization training can prime the segmentation network so that objects will tend to be segmented into parts that were previously diagnostic for categorization.

The Categorization Network

The categorization network has three layers of units: one representing the input patterns, one representing a bank of learned detectors, and one reflecting the category assignments of the inputs. Both of the weights from the input patterns to the detectors and the weights from the detectors to categories are learned. The categorization task uses a modified unsupervised competitive learning algorithm (O’Reilly & Munakata, 2000; Rumelhart & Zipser, 1985) but includes a top-down influence of category labels that incorporates supervised learning. The network begins with random weights from a two-dimensional input array to a set of detector units and from the detectors to the category units. When an input pattern is presented, the unit with the weight vector that is closest to the input pattern is the winner and will selectively adjust its weights to become even more specialized toward the input. By this mechanism, the originally homogenous detectors will become differentiated over time, splitting the input patterns into categories represented by the detectors. The competitive learning algorithm automatically learns to group input patterns into the clusters that the patterns naturally form. However, given that we want the detectors to reflect the experiment-supplied categories, we need to modify the standard unsupervised algorithm. This is done by including a mechanism that allows detectors that are useful for categorizing an input pattern to become more likely to win the competition to learn the pattern. The usefulness of a detector is assumed to be directly proportional to the weight from the detector to the presented category, which is provided as a label associated with an input pattern. The input-to-detector weights do not have to be set before the weights from detectors to categories are learned.

In addition to modifying the unsupervised development of hidden-layer detectors by considering their usefulness for categorization, a second modification of the standard competitive learning algorithm is required to fix one of its general problems in optimally making use of all detectors to cover a set of input patterns. This problem is that if multiple input patterns are presented that are fairly similar to each other, there will be a tendency for one detector to be the winner for all of the patterns. As a result, the winning detector’s weight vector will eventually become similar to the average of the input patterns’ activations, and the rest of the detectors will not learn at all. This situation is suboptimal because the input patterns are not covered as well as they would be if the unchanging detectors learned something. The standard solution to this problem is called leaky learning and involves adjusting both winning and losing detectors but adjusting losing detectors at a slower rate (Rumelhart & Zipser, 1985). To understand the more subtle problem with this solution, imagine, for example, that four input patterns naturally fall into two groups based on their similarities, and the network is given four detectors. Ideally, each of the detectors would become specialized for one of the input patterns. However, under leaky learning, one detector will tend to become specialized for one cluster, a second will become specialized for the other cluster, and the remaining two detectors will be pulled equally by both clusters, becoming specialized for neither. Note that it does not help to supplement leaky learning by the rule that the closer a detector is to an input pattern, the higher its learning rate should be. There is no guarantee that the two losing units will evenly split so that each is closer to a different cluster.

Other researchers noted related problems with competitive learning and suggested solutions (Grossberg, 1987). Our current solution is to conceptualize competitive learning as not simply a competition among detectors to accommodate a presented input pattern but also as a competition among input patterns to be accommodated by a given detector. Input patterns are presented sequentially to the network, and as they are presented the most similar input pattern to each detector is determined. The learning rate for a detector is set to a higher value for its most similar input pattern than for other inputs. In this manner, detectors that are not the winning detector for a pattern can still become specialized by becoming unequally influenced by different patterns. In addition, the learning rate for a detector when presented with an input pattern will depend on how well the input is currently covered by existing detectors. This dependency is required to allocate detectors to input regions where they are required. Putting these considerations together, the activation of detector $i$ when presented with pattern $p$, is

$$A_{i,p} = \sum_{k=1}^{n} I_{h,k} W_{i,k} + \sum_{j=1}^{c} STW_{j,i},$$

(2)
where $I_{h,p}$ is the activation of input unit $h$ for pattern $p$, $W_{i,h}$ is the weight from input $h$ to detector $i$, $S$ is the strength of the top-down pressure on detector development, $T$ is the teacher signal (if Pattern $p$ belongs to Category $j$ then $T = 1$, otherwise $T = -1$), and $W'_{j,i}$ is the weight from Detector $i$ to Category Unit $j$. The second term increases the activation of a detector to the extent that it is useful for predicting the input pattern's categorization. The detector activation will determine which detector is the winner for an input pattern. As such, detectors that are useful for categorization will tend to become winners, thus increasing their learning rate.

Input-to-detector weights are learned via top-down biased competitive learning using the following equation for changing weights from input pattern $h$ to Detector $i$:

$$
\Delta W_{i,h} = \begin{cases} 
M(I_{h,p} - W_{i,h}) & \text{if } \forall x(A_{i,p} \geq A_{x,p}) \\
N(I_{h,p} - W_{i,h})K_p & \text{if } \forall y(A_{i,p} \geq A_{y,p}) \\
O(I_{h,p} - W_{i,h})K_p & \text{otherwise,}
\end{cases}
$$

(3)

where $M$, $N$, and $O$ are learning rates ($M > N > O$), and $K_p$ is the distance between pattern $p$ and its closest detector. This distance is inversely related to the cosine of the angle between the vector associated with the closest detector and $p$. This set of learning rules may appear to be nonlocal in that all detectors are influenced by the closest detector to a pattern and depend on previous presented inputs. However, the rules can be interpreted as local if the pattern itself transmits a signal to detectors revealing how well covered it is, and if detectors have memories for previously attained matches to patterns. When an input pattern is presented, it will first activate the hidden layer of detectors, and then these detectors will cause the category units to become activated. The activation of the category unit $A_j$ will be

$$
A_j = \sum_{i=1}^{d} A_i W_{j,i},
$$

(4)

where $d$ is the number of detectors. Detector-to-category weights are learned via the delta rule, $\Delta W_{j,i} = L(T - A_j)A_i$, where $L$ is a learning rate and $T$ is the teacher signal previously described.

We formed a network with two detectors units and two category units and presented it with four input patterns. We gave the network four patterns that were used in experiments with human subjects. These patterns are not identical to the patterns shown in Fig. 7.9 but are of the same abstract construction. When the patterns were categorized as shown in Fig. 7.12A, such that the first two patterns belonged to Category 1 and the second two patterns belonged to Category 2, then on virtually every run the detectors that emerged were those reflecting the diagnostic segments—those segments that were reliably present on Category 1 or Category 2 trials. The picture within a detector unit in Fig. 7.12 reflects the entire weight vector from the $15 \times 15$ input array to the detector. The darkness of a pixel is proportional to the strength of the input-to-detector connection. When the same patterns are presented, but are categorized in the orthogonal manner shown in Fig. 7.12B, then different detectors emerge that again reflect the category-diagnostic segments. In both cases, each detector will have a strong association to one and only one of the category units. This is expected given that one of the factors influencing the development of detectors was their categorical diagnosticity. For the results shown here, and the later simulations I report, the following parameter values were chosen: $M = 0.1$, $N = 0.05$, $O = 0.02$, and $S = 0.1$. Activation values were between $-1$ and $+1$. One hundred passes through the input materials were presented to the network.
The Segmentation Network

The basic insight connecting categorization and segmentation tasks is that segmentation can also be modeled using competitive learning, and thus the two tasks can share the same network weights and can consequently have an influence on each other. Competitive learning for categorization sorts complete, whole input patterns into separate groups. Competitive learning for segmentation takes a single input pattern and sorts the pieces of the pattern into separate groups. For segmentation, instead of providing a whole pattern at once, we feed in the pattern one pixel at a time, so instead of grouping patterns, the network groups pixels together. Thus, each detector will compete to cover pixels of an input pattern so that the detector with the pixel-to-detector weight that is closest to the pixel's actual value will adapt its weight toward the pixel's value and inhibit other detectors from so adapting. Panels A–D show the weights from the 15 × 15 input array to each of two detectors and reflect the specializations of the detectors. If the pattern in Fig. 7.13 is presented to the network, the network might segment it in the fashion shown in Panel A. The two segments are complements of each other—if one detector becomes specialized for a pixel, the other detector does not.

Unfortunately, this segmentation is psychologically implausible. No person would decompose the original figure into the parts in Fig. 7.13A. To create psychologically plausible segmentations, we modify the determination of winners. Topological constraints on detector creation are incorporated by two mechanisms: Input-to-detector weights leak to their neighbors in an amount proportional to their proximity in the 15 × 15 array, and input-to-detector weights also spread to each other as a function of their orientation similarity, defined by the inner product of four orientation filters. The first mechanism produces detectors that tend to respond to cohesive, contiguous regions of an input. The second mechanism produces detectors that follow the principle of good continuation, dividing the figure X into two crossing lines rather than two kissing sideways Vs because the two halves of a diagonal line will be linked by their common orientation. Thus, if a detector wins for pixel X (meaning that the detector receives more activation when pixel X is on than any other detector), then the detector will also tend to handle pixels that are close to, and have similar orientations to, pixel X. The segmentation network, augmented by spreading weights according to spatial and orientation similarity, produces segmentations such as the one shown in Panel B of Fig. 7.13. For an alternative approach to segmentation that uses synchronized oscillations rather than architecturally separated detectors to represent segmentations, see Mozer, Zemel, Behrmann, and Williams (1992).

Although the segmentation in Panel B is clearly superior to Panel A’s segmentation, it is still problematic. The pixels are now coherently organized into line segments, but the line segments are not coherently organized into connected parts. Spreading weights according to spatial similarity should ideally create segmentations with connected lines, but such segmentations are often not found because of local minima in the harmony function (the value N defined later). Local minima occur when a detector develops specializations for distantly related pixels, and these specializations develop into local regions of mutually supporting pixels. Adjacent regions will frequently be controlled by different detectors. Each of the detectors may have sufficiently strong specializations for local regions that they will not be likely to lose their specialization due to the local relations of mutual support.

A common solution to local minima is to incorporate simulated annealing, by which randomness is injected into the system, and the amount of randomness decreases as an exponential function of time (Kirkpatrick, Gelatt, & Vecchi, 1983). Simulated annealing allows a system to settle into a good global minimum by ensuring a sufficient amount of randomness early on to explore the coarse solution.
landscape and sufficient regularity later on to ensure that the network will settle on a single solution. One dissatisfaction with simulated annealing is that the amount of randomness is time-locked. That is, it depends on the number of cycles executed by the network rather than the goodness of the network's solution per se. Other researchers have successfully developed annealing approaches that reduce the amount of randomness in the system over time but do so by basing the amount of randomness on the current structural goodness of a solution. For example, in the Copycat project (Hofstadter & Mitchell, 1994), the amount of randomness in the system is inversely proportional to the quality of the solution currently found. The current network uses the same notion to create globally coherent segmentations.

The segmentation network works by fully connecting a 15 × 15 input array of pixel values to a set of N detectors. Although ideally the value of N would be dynamically determined by the input pattern itself, in the current modeling, we assume that each object is to be segmented into two parts (as did Palmer, 1978). When an input pattern is presented, the pixels within it are presented in a random sequence to the detectors, and the activation of detector i which results from presenting pixel p is

\[ A_{i,p} = \sum_{h=1}^{n} I_{h} W_{i,h} S_{h,p}, \]  

where \( I_{h} \) is the activation of pixel h, \( W_{i,h} \) is the weight from pixel h to detector i, and \( S \) is the similarity between pixels h and p. As such, detectors are not only activated directly by presented pixels but are also activated indirectly by pixels that are similar to the presented pixels. Thus, a detector likely will be strongly activated by a certain pixel if it is already activated by other pixels similar to this pixel.

The similarity between two pixels h and p is determined by

\[ S_{h,p} = T \left( \frac{\sum_{i=1}^{n} G_{i,h} L_{i,h,p}}{n} \right) + U e^{-D_{h,p} C}, \]

where \( T \) and \( U \) are weighting factors, \( G_{i,h} \) is the response of orientation filter i to Pixel h, \( L_{i,h,p} \) is the degree to which pixels h and p fall on a single line with an orientation specified by filter i, \( D_{h,p} \) is the Euclidean distance between pixels h and p, and \( C \) is a constant that determines the steepness of the distance function. Four orientation filters were applied, at 0°, 45°, 90°, and 135°. The response of each filter was found by finding the inner product of the image centered around a pixel and a 5 × 5 window with the image of one of the four lines. Thus, the greater the overlap between the line and the image, the greater will be the output of the filter for the line. The alignment of two pixels along a certain direction was found by measuring the displacement, in pixels, between the infinite length lines established by the two pixel/orientation pairs.

Pixel-to-detector weights are learned via competitive learning:

\[ \Delta W_{i,p} = \begin{cases} M(I_{p} - W_{i,p}) + Random(-N, +N) & \text{if } \forall x(A_{i,p} \geq A_{x,p}) \\ Random(-N, +N) & \text{otherwise} \end{cases} \]

where \( M \) is a learning rate, and Random(-N, +N) generates Gaussian random noise scaled by + and -N. The amount of noise, N, in adjusting weights is a function of the harmony across all detectors relative to R, the maximal harmony in the system:

\[ N = R - \sum_{i=1}^{n} \sum_{p=1}^{m} I_{h} W_{i,h} W_{i,p} S_{h,p} \]

As such, if similar pixels in similar states have similar weights to detectors, then the harmony in the system will be high, and the amount of noise will be low. Thus, the amount of randomness in the weight-learning process will be inversely proportional to the coherency of the current segmentation. These learning equations allow the network to regularly create segmentations like the one shown in Panel C of Fig. 7.13.

In the simulations of the segmentation network to be reported, no attempt was made to find optimal parameter values. \( T \) and \( U \) were set at 0.5, \( M \) was set at 0.1, and \( C \) was set to 1.

**Combining the Networks**

Considered separately, the categorization and segmentation networks each can be considered to be models of their respective tasks. However, they were also designed to interact, with the aim of accounting for the results from Pevtzow and Goldstone’s (1994) experiments with human subjects. The segmentation network, because it shares the same input-to-detector weights that were used for the categorization network, can be influenced by previous category learning. Detectors that were diagnostic for categorization will be more likely used to segment a pattern because they have already been primed. Thus, if a particular shape is diagnostic and reasonably natural, the network will segment the whole into this shape most of the time, as shown in Panel D Fig. 7.13. In short, category learning can alter the perceived organization of an object. By establishing multisegment features along a bank of detectors, the segmentation network is biased to parse objects in terms of these features. Thus, two separate cognitive tasks can be viewed as mutually constraining self-organization processes. Categorization can be understood in terms
of the specialization of perceptual detectors for particular input patterns, where the specialization is influenced by the diagnosticity of a segment for categorization. Object segmentation can be viewed as the specialization of detectors for particular parts within a single input pattern. Object segmentation can isolate single parts of an input pattern that are potentially useful for categorization, and categorization can suggest possible ways of parsing an object that would not otherwise have been considered.

To model the results from Pevtzow and Goldstone (1994), the network was first trained on distortions of the patterns A, B, C, and D shown in Fig. 7.9, with either a horizontal or vertical categorization rule. As with the human experiment, the distortions were obtained by adding one random line segment to each pattern in a manner that resulted in a fully contiguous form. Following 30 randomly ordered presentations of distortions of the four patterns, the segmentation network was then presented with the original pattern shown in Fig. 7.13. Segmentations were determined by examining the stable input-to-detector weight matrix for each of the two detector units. Simulation results showed that the segmentation of the ambiguous original object is influenced by category learning. In particular, the original object tended to be segmented into parts that were previously relevant during category learning. As such, the results from Pevtzow and Goldstone (1994) are predicted under the additional assumption that response times in a part/whole task are related to the likelihood of generating a segmentation that includes the probed part.

In a subsequent test of the networks, the actual wholes used by Pevtzow and Goldstone (1994) in their part/whole task were presented to the segmentation network. Each whole was presented 200 times, 100 times preceded by each of the two possible categorization rules. Out of the 24 whole objects tested, segmentations involving categorization-relevant parts were produced more often than segmentations involving irrelevant parts for 19 of the objects. This comparison controls for any intrinsic differences in naturalness between segmentations of a whole object because the parts that are categorization-relevant for half of the simulated subjects are irrelevant for the other half. As such, the reported simulation results generalize to the actual materials used in the experiment.

This modeling shows how learning to group objects into categories affects grouping parts from a single object into segments. Category learning causes detectors to develop, and once these detectors have developed there is a tendency to use these primed detectors when segmenting an object into parts. Future work will be necessary to compare the model to other existing models that allow for experience-dependent visual object segmentation (e.g., Behrmann et al., 1998; Mozzer et al., 1992). Two extensions of the model would clearly be desirable: (1) allowing the model to determine for itself how many segments a pattern should be decomposed into and (2) allowing the computed segmentation of a single pattern to influence its categorization. The latter extension is required to fit human experimental evidence suggesting that not only does category learning influence segmentation, but also the perceived segmentation of an object influences its categorization (Schyns et al., 1998; Schyns & Rodet, 1997; Wisniewski & Medin, 1994).

Creating Building Blocks With the Segmentation Network

The categorization and segmentation networks both use principles derived from competitive learning for creating specialized detectors. Using these principles, detectors become specialized for either important classes of input patterns (with categorization) or for important segments within a single input pattern (with segmentation). In both cases, importance was both a function of the unsupervised statistical and psychophysical information in the stimuli and the task requirements involving the stimuli. The learning rules for the two networks were similar, although there were important differences to allow the categorization network to create optimally specialized detectors and the segmentation network to create psychologically plausible detectors. The main difference between the networks was that the categorization network was presented with multiple patterns, one pattern at a time, whereas the segmentation network was presented with only a single pattern, one pixel at a time. When the segmentation network is presented with multiple patterns, with each pattern still presented one pixel at a time, it shows some remarkable properties. The network is able to discover building blocks that can account for an entire set of patterns (as did Ghaemzadeh, 1995), and it can do so while taking into account perceptual constraints to create coherent building blocks, and with an online (rather than batch) learning process.

To understand these properties of the segmentation network, it is helpful to return to the categorization network with a simple example. In Fig. 7.14A, I present four input patterns to a categorization network with two detectors. Like standard competitive learning algorithms (Rumelhart & Zipser, 1985), the network learns the unsupervised statistical structure of the inputs and will typically divide the inputs into two categories—objects with a vertical bar on the left and objects with a vertical bar on the right. In this figure, the pattern shown in each detector represents the 4 × 4 set of weights going to the detector (0 = white, 1 = black). Superficially, it may appear that this network is extracting the vertical bars as components of the patterns, but these bars are only strongly present because these are the parts shared by the patterns for which a detector is specialized. The network does not isolate parts; it simply superimposes weight vectors from similar patterns. When this categorization network is given four rather than two detectors, as in Fig. 7.14B, then it will almost always develop a specialized detector for each of the patterns.

If the same patterns are presented to the segmentation network, then very different detectors emerge. Now, as shown in Fig. 7.14C, the segmentation network creates parts that, if combined, would generate each of the four patterns. Instead
of creating detectors specialized for entire patterns, it creates detectors for useful parts—parts that consistently appear across the patterns. These useful parts are created because a detector that is the winner for a pixel (is relatively specialized for the pixel) will also tend to be the winner for pixels that cooccur with the pixel. This is because the amount of activation of a detector when presented with a lit pixel is not determined only by the connection strength of the link connecting the pixel to the detector; it is also determined by other lit pixels, modulated by their link to the detector and their similarity to the pixel in question. This simple leakage of activations between similar, cooccurring pixels suffices to create detectors that will win (have the highest activation) for all pixels of a coherent part that repeatedly recurs. The tendency for a single detector to respond to cooccurring pixels can be strengthened even more by making the similarity of two pixels be not only a function of their orientation and spatial similarity but also their cooccurrence similarity. This cooccurrence similarity is simply obtained by creating a side network that connects every pixel with every other pixel and uses Hebbian learning to alter the links connecting pixels. With this network, pixels that are simultaneously present will develop strong positive links connecting them, and these links represent high cooccurrence similarity.

The segmentation network is still influenced by the spatial and orientation similarity of pixels when creating detectors, with the consequence that statistically equivalent input sets are not treated equivalently. This is apparent in Fig. 7.14D. The input patterns have the same logical structure as those in Fig. 7.14C, but one of the pixels (on the top row, third from the left) has been shifted down by two pixels. However, the detectors no longer create four detectors that can combine in different arrangements to exactly recreate each of the input patterns. This would be possible if the network created a detector with two disconnected pixels, but such a detector is unlikely to be created even if the pixels perfectly predict each other’s presence. The previously described EM algorithm does not have this constraint on creating perceptually coherent segments, but the empirical evidence suggests that people do indeed have such a constraint. In other words, using exactly the grid materials shown in Fig. 7.14, we find that people rarely construct parts out of pixels that are separated by a chess knight’s move. Instead, they either ignore one of the pixels, create a contiguous part by adding connecting pixels, or mentally distort the position of one of the pixels to make the part more coherent.

The detectors created in Fig. 7.14, parts B and C, are both reasonable solutions, and each has its place. Fig. 7.14B’s detectors offer a more holistic solution, creating detectors that become imprinted on whole input patterns. The power of Fig. 7.14C’s decompositional solution becomes apparent in Fig. 7.14E. Here, 16 input patterns can be represented in terms of conjunctions of eight acquired detectors. Again, the detectors that emerge combine pixels that are similar in their orientations and locations and reliably cooccur. Each of the 16 original patterns can be reconstructed exactly using only half as many detectors as would be required from a holistic solution. As a side benefit, the network also learns about negative contingencies between stimulus parts, for example, learning that a vertical bar on the top left predicts the absence of a vertical bar on the top right. This is a valuable first step in learning new dimension organizations, given that dimensions can be understood as sets of mutually exclusive alternative values. The feature values Red and Blue belong to the same Color dimension because an elementary object can only possess one of the values. The network implicitly extracts these relations,
coming to represent the stimuli of Fig. 7.14E in terms two dimensions each of which has four mutually exclusive alternatives.

The segmentation network is capable of constructing part-based representations for a set of stimuli, but depending on the statistical structures present in a set of stimuli, it can also create the same kind of holistic representations as the categorization network. For example, in Fig. 7.15A, four patterns that have very little overlap are presented to the network. In this case, the detectors each become specialized for an entire input pattern because there are no useful, frequently recurring parts to extract. More interestingly, in Fig. 7.15B, the input patterns fall into two varieties. Four of the patterns are composed of recurring parts, and one of the patterns is isolated from the others in that it shares no pixels with any other pattern. The segmentation network accommodates these patterns by representing most of the patterns componentially (as the conjunction of two detectors) but one of the patterns holistically. Consistent with the complementary advantages of configural and analytic representations, the network will represent patterns that are composed of factorially combined parts in terms of their parts but will simultaneously represent sparsely distributed and isolated patterns in terms of entire patterns. As such, the network has a natural way of reconciling the empirical evidence for unitization and differentiation. Both representations can be created, depending on the categorical and statistical pressures.

The results from these last figures suggests that the segmentation network can provide an account of learned differentiation into parts, but a final example shows a more specific analog to the 90° and 45° rotation conditions. Recall that participants showed better transfer between categorization rules related by 90°, rather than 45°. I argued that this was because orthogonal dimensions (separated by 90°) are compatible in that their contributions to a stimulus can be independently determined and combined. An analog of this is shown in Fig. 7.15C, where three patterns are presented, and four detectors are assigned to the segmentation network. The learned detectors show that the box pattern on the left is broken down into four bars. The presence of the two horizontal bars should come as no surprise given that they were also presented as input patterns. However, the two vertical bars reveal the extraction of detectors that are complementary to, and hence compatible with, other extracted detectors. Just as a rectangle can be consistently organized in terms of height and width (see Fig. 7.8), the box on the left is consistently organized into horizontal and vertical bars—bars that have no overlapping pieces. Fig. 7.15D shows that this trend toward extracting consistent parts can overcome some preferences in terms of perceptual organization. In Fig. 7.15D, the same box pattern is presented with two diagonal segments, and the network extracts these diagonal segments and the other diagonal segments that can combine to form the box pattern. Similar to the strong transfer found between orthogonal (90° rotated) dimensions, priming the segmentation network with some parts of a pattern makes the network also extract the parts of the pattern that are consistent with these primed parts. The explanation for this behavior is the same competitive learning principle.

![Diagrams showing the segmentation network](image)

**FIG. 7.15.** Like the categorization network, the segmentation network can also develop detectors specialized for entire input patterns if the input patterns possess little overlap (panel A). In fact, it can also develop a hybrid representation, representing some objects in terms of their parts and other objects as configurations. Part-based representations will be used for objects that share several segments in common with other objects from the set, whereas whole-object representations will be used for isolated objects (panel B). In panels C and D, the boxlike input pattern on the left is decomposed into different parts by the segmentation network, depending on other patterns that are presented. Detectors are created for parts that are complements of presented patterns, even though the complements are never presented by themselves. When patterns are presented with category assignments (shown as numbers below input patterns), there is a tendency to extract parts that are diagnostic for the categorization and also to extract the compatible complements of those diagnostic parts. This is shown in panels E and F, as are the learned associations between detectors and category units. The connection strength between detectors and categories is shown by the thickness of the line connecting them.
When the patterns of Fig. 7.15D are presented, detectors will likely emerge for the two pairs of diagonally related pixels that always cooccur. Once these detectors are constructed, they will be likely to win the competition for the diagonal portions of the box pattern when it is presented, leaving the two other diagonal parts left to be claimed. These unclaimed parts will then be claimed by detectors that have not yet become specialized.

Fig. 7.15, Panels C and D, show that actions that promote attending a part (such as presenting it or making it diagnostic for a categorization) also promote attending to the complementary part that is consistent with it. Fig. 7.15, Panels E and F, shows that this occurs even with displays that are closer analogs to the stimuli shown in Fig. 7.9. In these figures, each input pattern is made up of two features, and each feature has two possible values. Unlike previous demonstrations of the segmentation network, the patterns are categorized into two groups so that the first two patterns belong to Category 1, and the second two patterns belong to Category 2. The network creates early detectors for the category-relevant parts and learns strong weights between these detectors and the appropriate category unit, as shown by the relatively thick lines connecting diagnostic detectors to category units. However, the network also learns the complements of these diagnostic parts exactly because the complements are the parts that remain in a stimulus once diagnostic parts have been accommodated.

Fig. 7.15, Panels C–F, are analogs to the dimension differentiation observed by Goldstone and Stuyvers (2001) in that consistent rather than cross-cutting featural interpretations are generated. The simulations are only analogous because the experiments were concerned with the extraction of independent, complementary dimensions that influenced the whole stimulus rather than spatially isolated parts. There seems to be a close connection between organizing an object into dimensions and parts, but I leave it for future work to reveal whether the simulations reported here can be interpreted as more than mere analogies. At least the segmentation network provides a way of intuitively understanding the seemingly paradoxical positive transfer between orthogonal (90° rotated) categorization rules. When considering a physical part of an object, it is natural to notice the complement of the part—the parts that remain in the stimulus when the part is removed. It is easy to see that the complement requires the same parsing of the object into pieces as does the part itself. In the same way, a dimension is compatible with a second dimension that is removed from it by 90°.

CONCLUSIONS

On both empirical and theoretical grounds I have argued that the perceptual descriptions given to objects are influenced by the tasks for which the objects are used. Several types of perceptual change were empirically reported. Imprinting on whole objects happens if they occur frequently or on parts of objects if the parts are shared across several objects. These two types of imprinting can occur in the absence of any feedback. The stimuli themselves, and their statistical properties, can drive the perceptual learning. However, each of these two types of learning also has a parallel when the objects are learned in the context of categorization. Learning by imprinting on whole objects becomes particularly influential when a specific object belongs in an important category. In a process referred to as unification, originally separated parts of an object are combined into a unified and coherent whole. Once constructed, the unit can be efficiently recognized and has properties similar to a photograph-like template. Category learning can also induce the opposite process to occur, separating originally integrated percepts into psychologically differentiated dimensions or parts. Rather than viewing unification and differentiation as contradictory, they are best viewed as aspects of the same process that bundles stimulus components together if they diagnostically cooccur and separates these bundles from other statistically independent bundles. Under this conception, learning a featural or dimensional organization consists in learning how to carve a stimulus into useful components.

The empirical conclusion, generalized to unify differentiation and unification, is that the perceptual building blocks used for representing objects adapt to unsupervised and supervised demands. This conclusion has a consequence for literature on the development of perception with age or expertise. Rather than trying to characterize development in terms of a single integral-to-separable or component-to-configural trend, this conclusion suggests that both of these trends should be expected and that which is found will depend on the nature of the stimuli and the feedback. Both trends potentially can be explained by a single process of creating useful building blocks. There are real and important differences between perceiving an object by breaking it down into parts and by matching it to a configurational representation. However, in both cases units are formed so that the physical segments within the units interact and interfere with each other and are likely to be perceived together, whereas segments from different units are relatively independent. For example, regardless of whether a unit of perception exists at the line segment, letter, syllable, word, phrase, or sentence level (Wheeler, 1970), the same factors will probably determine the strength of the unit: How often do the unit's pieces cooccur, how important is that cooccurrence for a linguistic category, how often do the unit's pieces cooccur with pieces from other units, and what perceptual grouping principles support the unit?

The major theoretical contribution of the research has been to specify a model for how perceptual and conceptual learning might interact. In the neural network presented, feature detectors are developed that represent the network's set of acquired vocabulary elements. The network begins with homogenous, undifferentiated detectors that become specialized for different patterns or pattern parts over time. Furthermore, the model has a mechanism by which detector-to-category associations modify the nature of the detectors. It is not necessary to first develop detectors and then build associations between detectors and categories. These two types of learning can and should go on simultaneously (Goldstone, 1994b). The concepts to be learned can reach down and influence the very features that ground
the concepts. This is done without requiring dynamic, online recurrent activation passing, as is found in McClelland and Rumelhart's (1981) interactive activation model. That is, when processing patterns, the network is purely feedforward, with input patterns triggering detectors, and detectors triggering categories. Categories influence the detectors that are created by making useful detectors more likely to be constructed. However, once constructed, categories do not need to send activation to detectors nor do detectors affect input representations (although these are interesting possibilities to explore). The feedforward nature of the network is advantageous with respect to speed and efficiency. The network's detectors can be quickly triggered before category-level information is available on a particular occasion, even though their nature has been adapted over longer time courses to categorization needs.

Like a good pair of Birkenstock shoes that provide support by flexibly conforming to the foot, perceptual object descriptions support our concepts by conforming to these concepts. Ironically, more stable shoes and perceptions would actually be less adequate as foundations. In these and other situations, flexible, not rigid, foundations make strong foundations. If flexibility in perceptual foundations is desirable, it is important to also remember that this flexibility is constrained by perceptual grouping principles (Kellman & Arterberry, 1998), informational requirements, patterns of repetition and cooccurrence, and task usefulness. Within these constraints, the building blocks of perception can show a surprising degree of adaptability. Human creativity requires a certain amount of conceptual block busting to be sure (Adams, 1974) but also benefits from a healthy dose of perceptual block building.

ACKNOWLEDGMENTS

The research reported in this chapter has benefited greatly from comments and suggestions by Marlene Behrmann, Rutie Kimchi, John Kuschke, Douglas Medin, Amy Needham, Robert Nosofsky, Mary Peterson, and Richard Shiffrin. This research was supported by National Institute of Health Grant R01 MH56871 and National Science Foundation Grant 0125287. The author can be reached by electronic mail at rgoldsto@indiana.edu and further information about the laboratory can be found at http://cognition.psych.indiana.edu/.

REFERENCES


Cattell, J. M. (1886). The time it takes to see and name objects. Mind, 11, 63–65.

7. LEARNING TO PERCEIVE


