9 Learning to See and Conceive

Robert L. Goldstone, Alexander Gerganov, David Landy, and Michael E. Roberts

Human concept learning depends upon perception. Our concept of "car" is built out of perceptual features such as "engine," "tire," and "bumper." However, recent research indicates that the dependency works both ways. We see bumpers and engines in part because we have acquired "car" concepts and detected examples of them. Perception both influences and is influenced by the concepts that we learn. We have been exploring the psychological mechanisms by which concepts and perception mutually influence one another, and building computational models to show that the circle of influences is benign rather than vicious.

Perceptual Learning Is "Early" Neurologically, Functionally, and Developmentally

An initial suggestion that concept learning influences perception comes from a consideration of the differences between novices and experts. Experts in many domains, including radiologists, wine tasters, and Olympic judges, develop specialized perceptual tools for analyzing the objects in their domains of expertise. Much of training and expertise involves not only developing a database of cases or explicit strategies for dealing with the world but also tailoring perceptual processes to more efficiently represent the world (Gibson 1991). Tuning one's perceptual representation to the environment is a risky proposition. Once a perceptual representation has been altered, it affects all "downstream" processes that act as consumers of this altered representation. It makes sense to adapt perceptual systems slowly and conservatively. However, the payoffs for perceptual flexibility are also too enticing to forego. They allow an organism to respond quickly, efficiently, and effectively to stimuli without dedicating on-line attentional resources. Instead of strategically determining how to use an unbiased perceptual representation to fit one's needs, it is often easier to rig up a perceptual system to give task-relevant representations, and then simply leave this rigging in place without strategic control. Perceptual learning is early in several senses: neurological, functional, and developmental.

Neurological Evidence

Several sources of evidence point to expertise influencing perceptual processing at a relatively early stage of processing. First, electrophysiological recordings show enhanced electrical activity at about 164 milliseconds after the presentation of dog or bird pictures to dog and bird experts, but only when they categorized objects within their domain of expertise (Tanaka and Curran 2001). A similar early electrophysiological signature of expertise is found with fingerprint experts when they are shown upright fingerprints, but is delayed when the fingerprints are inverted (Busey and Vanderkolk 2005). Interestingly, the timing and form of this expertise-related activity is similar to the pattern found when people are presented with faces, a stimulus domain in which, arguably, almost all people are experts (Gauthier et al. 2003).

Second, prolonged practice with a subtle visual categorization results in much improved discrimination, but the improvements are highly specific to the trained orientation (Notman et al. 2005). This profile of high specificity of training is usually associated with changes to early visual cortex (Fahle and Poggio 2002). Practice in discriminating small motions in different directions significantly alters electrical brain potentials that occur within 100 milliseconds of the stimulus onset (Fahle 1994). These electrical changes are centered over the primary visual cortex, suggesting plasticity in early visual processing. Karni and Sagi (1993) 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. In the auditory modality, training in a selective attention task produces differential responses as early in the sensory processing stream as the cochlea (Puel et al. 1988). This amazing degree of top-down modulation of a peripheral neural system is mediated by descending pathways of neurons that project from the auditory cortex all the way back to olivocochlear neurons, which in turn project to outer hair cells within the cochlea (Suga and Ma 2003).

Third, expertise can lead to improvements in the discrimination of low-level simple features, as with the documented sensitivity advantage that radiologists have over novices in detecting low-contrast dots in X-rays (Sowden et al. 2000). Fourth, imaging techniques have succeeded in identifying brain regions associated with the acquisition of expertise. Expertise for visual stimuli as eclectic as butterflies, cars, chess positions, dogs, and birds has been associated with an area of the temporal lobe known as the fusiform face area (Bukach et al. 2006). The identification of a common brain area implicated in many domains of visual expertise suggests the promise of developing general theories and models of perceptual learning. This is the main purpose of our work.

Several other pieces of auxiliary evidence point to experience having early effects on perception, where "early" is operationalized neurologically in terms of a relatively small number of intervening synapses connecting a critical brain region to the external world. Experience making fine tactile discriminations influences primary somatosensory cortices. Monkeys trained to make discriminations between slightly different sound frequencies develop somatosensory cortex representations for the presented frequencies than control monkeys (Recanzone et al. 1993). Similarly, monkeys learning to make a tactile discrimination with one hand develop a larger cortical representation for that hand than for the other hand (Recanzone et al. 1992). Elbert and colleagues (1995) measured brain activity in the somatosensory cortex of violinists as their fingers were lightly touched. There was greater activity in the sensory cortex for the left hand than the right hand, consistent with the observation that violinists use their left-hand fingers considerably more than their right-hand fingers.

Functional Evidence

In terms of functional evidence, experience often exerts an influence before other putatively early perceptual processes have been completed. Peterson and Gibson (Peterson 1994; Peterson and Gibson 1994; Peterson et al. 1991) found that the organization of a scene into figure and ground is influenced by the visual familiarity of the contours. Their participants were more likely to respond that familiar, compared to unfamiliar, forms were figural elements occluding the background. This effect was not found when flipping the scenes upside down eliminated familiarity, but was found regardless of whether the familiar object was black or white. Interpreting the familiar region as a figure was found even when the unfamiliar regions had the strong Gestalt organization cue of symmetry. Peterson and Lampignano (2003) found direct evidence that the acquired familiarity of a shape successfully competes against Gestalt cues such as partial closure to determine the organization of a scene into figure and ground.

Consistent with an influence of training that occurs relatively early in the information-processing stream, perceptual organizations that are natural according to Gestalt laws of perception can be overlooked in favor of perceptual organizations that involve familiarized materials. Behrmann and colleagues (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 fragments were interpreted as belonging to the same object if they had co-occurred many times in a single shape (Zemel et al. 2002). As shown in figure 9.1, object fragments that are not naturally grouped together because they do not follow the Gestalt law of good continuation (according to which there is an inherent tendency to see a line continuing its established direction) can nonetheless be perceptually joined if participants are familiarized with an object that unifies the fragments.

Developmental Evidence

Perceptual learning is also “early” in the developmental sense. Many of the most striking changes to our perceptual systems occur in the first two years of life. Infants are surprisingly adept at adapting their perceptual systems to statistical regularities in their environment. Needham and Baillargeon (1998, see also Needham 1999; Needham et al. 2005)

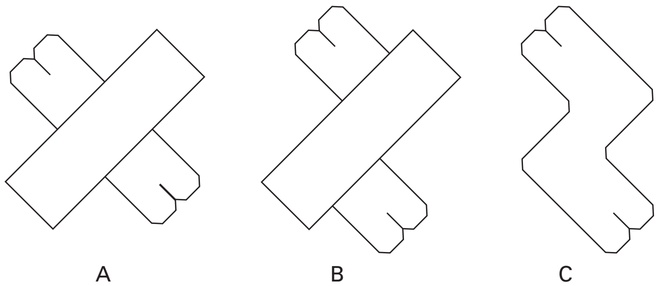

Figure 9.1
Images used by Zemel et al. (2002). Object fragments that are not naturally grouped together because they do not follow the Gestalt law of good continuation (panel B) can nonetheless be perceptually joined if participants are familiarized with an object that unifies the fragments (panel C). Figure courtesy of Zemel and colleagues.

found that exposing infants to single or paired objects tended to lead the infants to parse subsequent events in terms of these familiarized configurations. As shown in figure 9.2, infants initially exposed to a cylinder abutting a rectangular box showed relatively long looking times, suggesting surprise, if one of the objects subsequently moved separately from the other. This surprise occurred even though the natural perceptual cue of minima of curvature (which would segment a scene into parts at negative minima of curvature on silhouette edges) (Hoffman and Richards 1984) would suggest a plausible division between the cylinder and box.

Paul Quinn and his colleagues (Quinn and Schyns 2003; Quinn et al. 2006) were interested in further pursuing the question of whether infants, like adults, can perceive objects in terms of familiarized parts rather than the parts given by default perceptual organizations. They contrasted familiarity-based segmentations with segmentations derived from one of the Gestalt perceptual laws of organization, good continuation. In figure 9.3, the shapes in the "familiarization" set are all ambiguous, interpretable as either a polygon combined with an overlapping circle or as a closed figure consisting of both straight lines and curves combined with a three-quarter-circle "Pac-man" shape.

The former interpretation is consistent with good continuation. Consistent with this law, three-to-four-month-old infants tend to see the shapes in the "familiarization" set as containing a circle and a polygon. The evidence for this is that when the infants are subsequently presented with the full and three-quarters circle, they look at the full circles 39 percent of the time and three-quarters circles 61 percent of the time (Quinn et al. 2006), shown as path A in figure 9.3. Prior base-line experiments showed that this looking preference was not due to a general preference for the Pac-man shape; when infants were not first shown either the prefamiliarized or familiarized shapes, there was no reliable tendency for infants to preferentially look at the Pac-man shape. Together with many other experiments on visual shape perception, this first result suggests that infants have a novelty preference—a preference to look at unfamiliar objects—and that the Pac-man shape seems

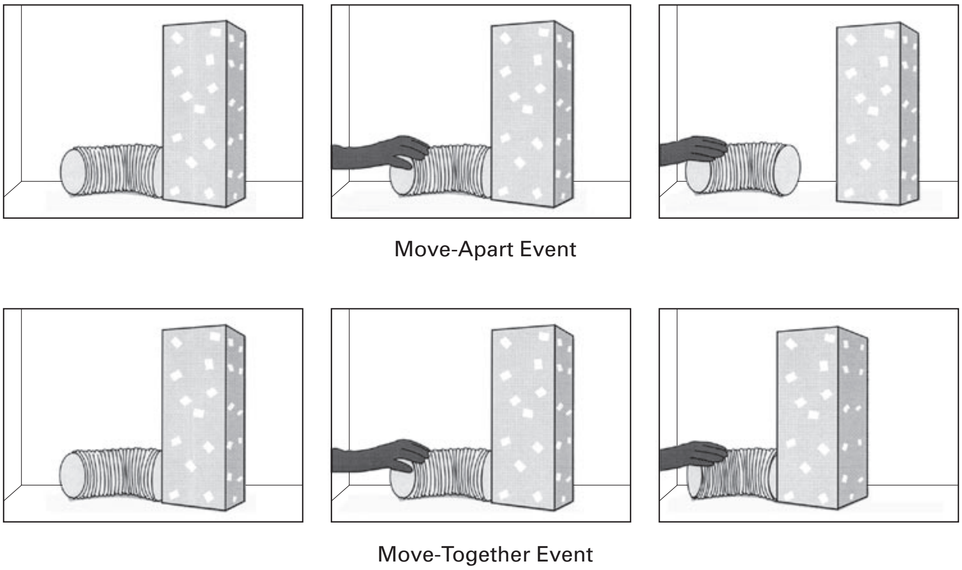

Figure 9.2
Stimuli from Needham and Baillargeon (1998). Infants exposed to a cylinder juxtaposed with a rectangular box showed relatively long looking times (suggesting surprise) if one of the objects moved separately from the other, as depicted in the move-apart event in the top panel. Figure courtesy of Needham and Baillargeon.

novel because, although it is present in the familiarized shapes, it is not the natural segmentation for infants to make. Their natural segmentation, like that of adults, is to obey follow the law of good continuation and interpret the ambiguous complexes as containing circles.

A second condition suggests that the infants' segmentations can be altered by prior learning. For some infants, looking at the "familiarization" shapes were preceded by looking at the "prefamiliarization" shapes shown in figure 9.3. These "prefamiliarization" shapes consist of the Pac-man shape combined with a polygon. Habituation trials directly after infants saw the "prefamiliarization" shapes show that the infants interpreted these forms as containing the Pac-man shape, rather than a partially covered circle, as indicated by their tendency to look at the novel-seeming circle 56 percent of the time when it was presented next to a Pac-man shape. This tendency to preferentially look at the circle continued to be found even after the "familiarization" stimuli were shown to infants (path B in figure 9.3). This strongly suggests that the "familiarization" shapes are now interpreted as containing Pac-men rather than circles, and thus circles seem novel and worth more extended scrutiny. Taken together, these results indicate that early in development, children are predisposed to learn shapes from their environments and then interpret their environment in terms of these learned shapes.

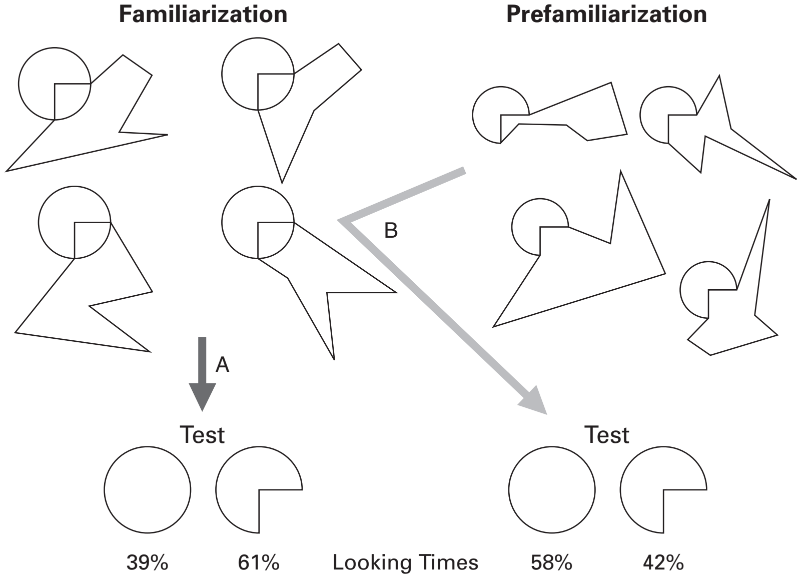

Figure 9.3
When three- to four-month-old infants are familiarized with figures consisting of a complex polygon superimposed on a circle, they tend to look more at the three-quarter-circle "Pac-man" type of shape than at the full circle. However, when they are prefamiliarized with polygons superimposed with the three-quarters "Pac-man," they tend to look more at the full circle.

Perceptual Learning Via Unitization

Perceptual learning is powerful because it is not only early in the above three senses, but also, unlike a reflex, is task-dependent. This combination of properties allows perception to be both fast and useful. The nature and degree of perceptual learning is typically closely tied to the task, goals, and knowledge of the observer. Although perceptual learning may occur without awareness (Watanabe et al. 2001), it is more common for researchers to report learning that depends upon both the objective frequency and subjective importance of the physical feature (Sagi and Tanne 1994; Shiu and Pashler 1992). For example, altering the color of target objects in a visual search paradigm from training to transfer tasks does not influence performance unless the training task requires encoding of color (Logan et al. 1996). Our empirical research has been focused on the particular mechanisms by which perceptual processes are modified by experience and tasks. One result of category learning is to create perceptual units that combine stimulus components useful for the categorization. Such a process is one variety of the more general phenomenon of unitization, by which single functional units are constructed that are triggered when a complex configuration arises (Goldstone 1998, 2000). In the next section, the complementary process, dimension differentiation, will be described. Although unitization and dimension differentiation seem to be contradictory processes, we will argue on computational grounds that they reflect the same mechanism of determining useful perceptual building blocks for representing patterns.

Cattell (1886) invoked the notion of perceptual unitization to account for the advantage that he found for tachistoscopically presented words relative to nonwords. Unitization has also been posited in the field of attention, where researchers have claimed that shape components of often-presented stimuli with practice become processed as a single functional unit (LaBerge 1973). Shiffrin and Lightfoot (1997) report evidence from the slopes relating the number of distracter 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 twenty hour-long sessions. These prolonged decreases were not observed for a simple search task requiring attention to only one component. The authors concluded that conjunctive training leads to the unitization of the set of diagnostic line segments, resulting in fewer required comparisons.

Our own experiments (Goldstone 2000) have explored unitization from a complementary perspective. First, our experiments reflect our primary interest in the influence of category learning on unitization, under the hypothesis that a unit will tend to be created if the parts that make up the unit frequently co-occur, and if the unit is useful for determining a categorization. Second, we use a new method for analyzing response-time distributions to assess the presence of unitization.

Whenever the claim for the construction of new units is made, two objections must be addressed. First, perhaps the unit existed in people's vocabulary before categorization training. Our stimuli are designed to make this explanation unlikely. Each unit to be sensitized is constructed by connecting five randomly chosen curves. There are ten curves that can be sampled, yielding 10^5 possible different units. As such, if it can be shown that a subject can be sensitized to any randomly selected unit, then an implausibly large number of vocabulary items would be required under the constraint that all vocabulary items are fixed and a priori. 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 will be untenable if participants, at the end of extended training, are faster at categorizing the units than would be expected by the analytic approach.

In our experiments the categorization task was designed so that evidence for five components must be received before certain categorization responses are made. That is, it was a conjunctive categorization task. The stimuli and their category memberships are shown in figure 9.4.

Each of the letters refers to a particular segment of one "doodle." Each doodle was composed of five segments, with a semicircle below the segments added to create a closed figure. To correctly 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

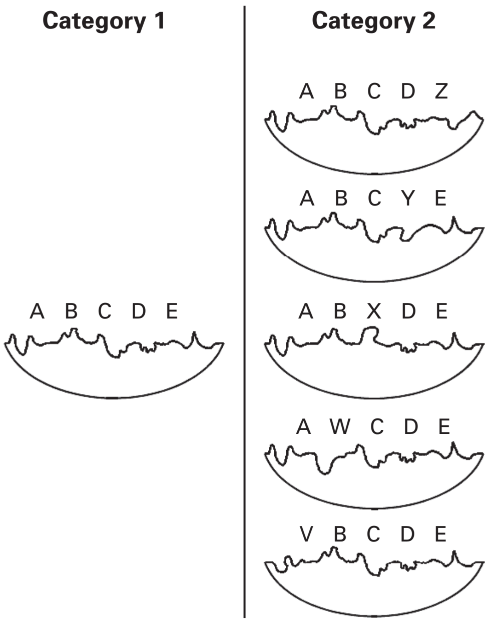

Figure 9.4
Stimuli used by Goldstone (2000). Each letter represents a particular stimulus segment, and each stimulus is composed of five segments. To categorize the item represented by "ABCDE" as belonging to Category 1, it is necessary to process information associated with each of the segments.

component was not attended, then "ABCDE" could not be distinguished from "ABCDZ," which belongs in category 2. Only the complete five-way conjunction suffices to accurately categorize "ABCDE." If unitization occurs during categorization, then, with training, the stimulus "ABCDE" may become treated functionally like a single component. If this occurs, then participants should be able to quickly respond that this stimulus belongs to category 1. So a pronounced decrease in the time required to categorize the conjunctively defined stimulus "ABCDE" was taken as initial evidence for unitization.

For improvement in the conjunctive task to be taken as evidence for unitization, two important control conditions are necessary. First, it is important to show that tasks that do not require unitization do not show comparable improvements. To this end, a control task was included that allows participants to categorize the item "ABCDE" by attending to only a single component rather than a five-way conjunction. This was done by having category 2 contain only one of the five category 2 doodles shown in figure 9.4, randomly selected for each participant. This "One" (component) condition should not result in the same speed-up over training as the "All" (components) condition where five components must be attended. If it does, then the speed-up can be attributed to a simple practice effect rather than unitization. Second, it is important to show that stimuli that cannot be unitized also do not show comparable speed-ups. For this control condition, it was necessary to attend a five-way conjunction of components, but the ordering of the components within the stimulus was randomized. That is, "ABCDE" and "CEBDA" were treated as equivalent. In this "Random" condition, a single template cannot serve to categorize the "ABCDE" stimulus and unitization should therefore not be possible.

The results from the experiment were suggestive of unitization. The results in figure 9.5 reflect only the correct responses to the category 1 doodle "ABCDE."

The horizontal axis shows the amount of practice over a 1.5-hour experiment. The condition where all components were necessary for categorization and where they were combined in a consistent manner to create a coherent image showed far greater practice effects than the others. This dramatic improvement suggests that the components are joined to create a single functional unit to serve categorization. Particularly impressive speed-ups were found when and only when unitization was possible and advantageous.

This paradigm also provides stronger evidence for unitization. The alternative to the unitization hypothesis is that responses in the "All" task are obtained by integrating evidence from five separate judgments of the type required in the "One" task. In arguing against this analytic account, a highly efficient version of the analytic account was devised so that it could be observed whether it still predicted response times that were too slow. The first advantage given to the analytic model was fully parallel processing; "All" responses were made by combining five "One" responses, but evidence for these five "One" responses was assumed to be obtained simultaneously. Second, the analytic model was given unlimited capacity; identifying one component was not slowed by the need to identify another component. In obtaining predictions from this charitably interpreted

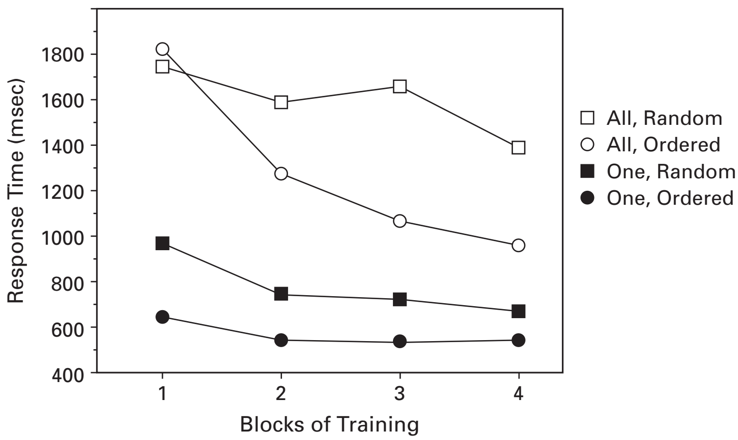

Figure 9.5
Results from Goldstone (2000). The most pronounced improvement was observed when all components were required for a categorization ("All"), and the components were always in the same positions ("Ordered").

analytic model, it is important to remember that the "All" task is a conjunctive task; to categorize ABCDE as a category 1 item with the required 95 percent categorization accuracy, all five components must be identified. Also, there is intrinsic variability in response times, even in the simple task where only one component must be identified. An analytic model of response times can be developed that predicts what the "All" task response-time distribution should be, given the "One" task distribution. After training, a distribution of response times in the "One" task can be empirically determined. To derive the analytic model's predictions, one can randomly sample five response times from this distribution. The maximum of these five times, rather than the average, is selected because no response can be made to the conjunction until all components have been recognized. We can repeat this selection process several times to create a distribution of the maximums, and this yields the predicted response-time distribution for the "All" task. Fortunately, there is an easier, more formal way of obtaining the predicted distribution. The "One" task response-time distribution is converted to a cumulative response-time distribution, and each point on this distribution is raised to the fifth power. If the probability of one component's being recognized in less than 400 msec is 0.2, then the probability of all five components' being recognized in less than 400 msec is 0.2 raised to the fifth power, assuming sampling independence.

A replication of the experiment shown in figure 9.4 was conducted that included the "Ordered All" and "One" tasks. Only four research assistants participated as participants, but unlike in the 1.5-hour experiment described previously, each participant was given fourteen hour-long training sessions. The results, shown in figure 9.6, are only for category 1 responses on the final day of the experiment.

These results indicate violations of the analytic model. Naturally, the "One" task was the fastest (most shifted to the left) according to the cumulative response-time distributions. The analytic model's predictions are shown by the curve labeled "One^5," which is obtained simply by raising each point on the "One" curve to the fifth power. For two of the four participants, the actual "All" distribution was faster than the analytic model's predictions for all regions of the distribution. For all four participants, the fastest 30 percent of response times for the "All" task were faster than predicted by the analytic model, even though all participants were achieving accuracies greater than 95 percent. Although the advantage of the "All" over the "One^5" distribution may not look impressive, the entire distributions were significantly different by a Kolmogorov-Smirnoff test for all participants except the participant C.H. Why were the violations of the analytic model restricted to the fast response times? A likely possibility is that a range of strategies was used for placing ABCDE into category 1 in the "All" task. On trials where a participant used the analytic strategy, the charitably interpreted analytic model would be expected to underestimate observed response times, given the implausibility of pure parallel, unlimited capacity processing. However, on trials where participants used the single constructed unit to categorize ABCDE, violations of the analytic model are predicted. On average, the unit-

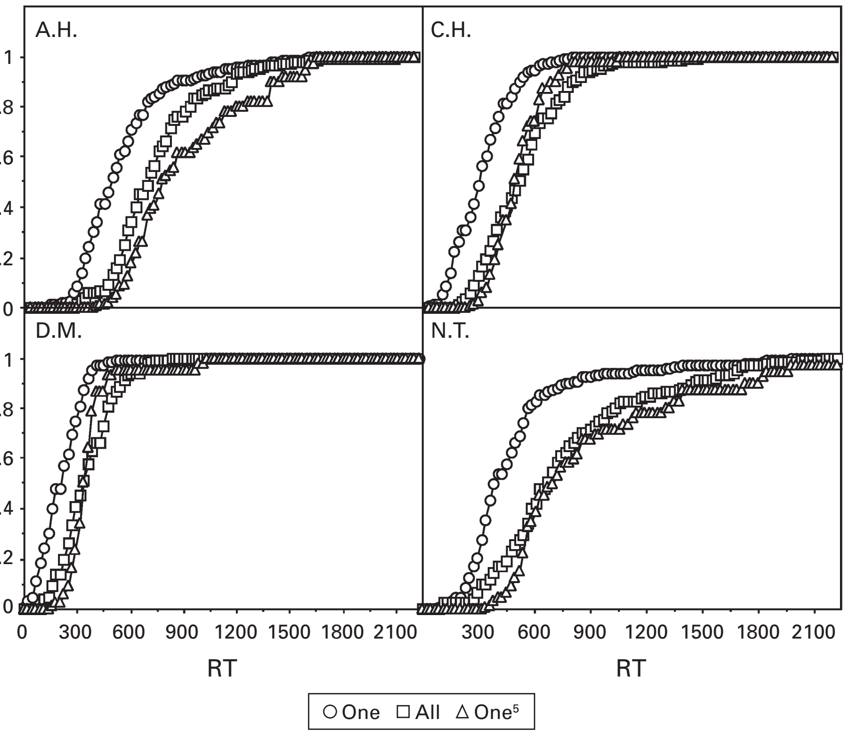

Figure 9.6
The cumulative response time distributions for the four participants taken from the last session of the experiment. The "One" and "All" distributions were empirically obtained. The "One^5" distribution is obtained by raising each point along the "One" distribution to the fifth power, and represents the analytic model's predicted cumulative distribution for the "All" task. Violations of this analytic model occur when the "All" task's distribution is shifted to the left of the analytic model's distribution. Such violations occur for the fastest half of response times for all four participants (significantly so for all participants except C.H.).

based trials will be faster than the analytic trials. That is, if participants successfully use a single unit to categorize ABCDE then they will tend to do so quickly. If they cannot use this route, their response time will tend to be slower. Thus, if the fast and slow response times tend to be based on single units and analytic integration, respectively, then we would predict violations of the analytic model to be limited to, or more pronounced for, the fast response times.

In light of these results, we concluded that category learning probably created new perceptual units. Large practice effects were found if and only if stimuli were unitizable (the first experiment), and responses after fourteen hours of training were faster for conjunctively defined categories than predicted by a charitably interpreted analytic model. The results shown in figure 9.6 only violate the analytic model if negative dependencies or independence is assumed between the five sampled response times that make up one

"All" judgment. Although it is beyond the scope of this chapter, we also have evidence for violations of the analytic model for classes of positive dependencies, using Fourier transformations to deconvolve shared input/output processes from the One task response-time distribution (Goldstone 2000; Smith 1990).

There is still a remaining question: Exactly how do people become so fast at categorizing ABCDE in the "All" task? Two qualitatively different mechanisms could account for the pronounced speed-up of the conjunctive categorization: a genuinely holistic match process to a constructed unit, or an analytic model that incorporates interactive facilitation among the component detectors. According to a holistic match process, a conjunctive categorization is made by comparing the image of the presented item to an image that has been stored over prolonged practice. The stored image may have parts, but either these parts are arbitrarily small or they do not play a functional role in the recognition of the image.

There is evidence supporting the gradual development of configural features. Neurophysiological findings suggest that some individual neurons represent familiar conjunctions of features (Perrett and Oram 1993; Perrett et al. 1984), and that prolonged training can produce neurons that respond to configural patterns (Logothetis et al. 1995). However, our results could also arise if detecting one component of "ABCDE" facilitates detection of other components (Townsend and Wenger 2004). In either case, the process is appropriately labeled "unitization" in that the percepts associated with different components are closely coupled. In fact, an interactive facilitation mechanism could be seen as the mechanism that implements holistic unit detection at a higher functional level of description.

Perceptual Learning via Differentiation

New perceptual representations can be created by chunking together elements that were previously psychologically separated in a process of unitization, and the converse process also occurs. This second process is dimension differentiation, according to which dimensions that are originally psychologically fused become separated and isolated. It is useful to contrast dimension differentiation from the more basic learning process of learning to selectively attend to one psychological dimension of stimulus. Selective attention assumes that the different dimensions that make up a stimulus can actually be individually attended. In his classic research on stimulus integrality and separability, Garner (1976, 1978) argues that stimulus dimensions differ in how easily they can be isolated or extracted from each other. 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 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 and Shepp 1988; Melara et al. 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 for four-year-old children (Kemler and Smith 1978; Smith and Kemler 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). 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 to nouns (Smith et al. 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 make up color than are nonexperts (Burns and Shepp 1988). Goldstone (1994) has shown 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 a selectively heightened sensitivity at making discriminations of saturation, relative to brightness. 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 in order to know how integral two dimensions are, one has to know something about the observer's history.

Goldstone and Steyvers (2001) used a category learning and transfer paradigm to explore whether genuinely arbitrary dimensions can become isolated from each other. Our subjects first learned to categorize a set of sixteen faces into two groups by receiving feedback from a computer, and then were transferred to a second categorization task. The stimuli varied along two arbitrary dimensions, A and B, that were created by morphing between randomly paired faces (Steyvers 1999). We created a set of stimuli with no preferred dimensional axes by assigning sixteen faces to coordinates that fall on a circle in the abstract space defined by dimensions A and B. To this end, A variable D was created that was assigned sixteen different values, from 0 to 360, in 22.5-degree 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 sine(D). The end result, shown in figure 9.7, is a set of faces that are organized on a circle with no privileged dimensional axes suggested by the set of faces.

With these faces, Goldstone and Steyvers 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 sixteen faces into two

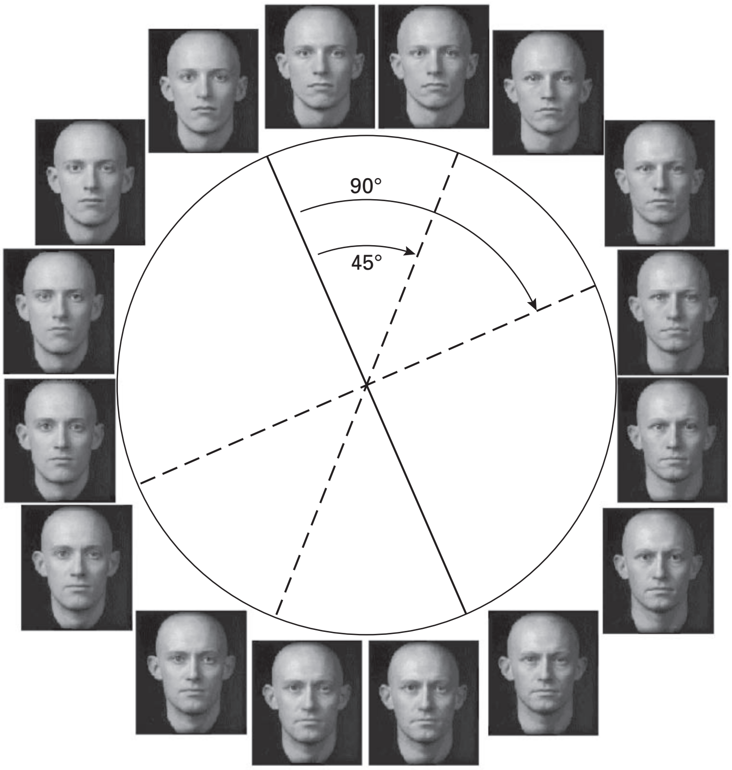

Figure 9.7
Stimuli from Goldstone and Steyvers (2001), experiment 3. The proportions of two randomly paired faces are negatively correlated such that the more of face 1 is present, the less of face 2 there will be. This negative correlation establishes the dimension on the X-axis, and a similar negative correlation between two other faces establishes the Y-axis dimension. A set of circularly arranged faces was created by varying degrees from 0 to 360, and assigning the face a value on the X-axis dimension based on the cosine of the degrees, and assigning the face's Y-axis 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 line dividing them. The two category boundaries given to a subject were related to each other by either 45 or 90 degrees.

equal piles using straight lines such as those shown in figure 9.7. Each subject was given categorization training with one classification rule and then was transferred to a second categorization task governed by a different rule. The critical experimental manipulation was whether the final categorization rule involved a rotation of 45 or 90 degrees relative to the initial rule. Given that the initial categorization rules were randomly selected, the only difference in the 45- and 90-degree rotation conditions was whether the category boundary was shifted by two or four faces. When the category boundary was shifted by two faces in the 45-degree condition, the labels could either be preserved (six faces assigned to the same category) or reversed (two faces assigned to the same category), but the category boundary itself was the same for these two conditions. The results, shown in "Whole Face Dimensions" columns of figure 9.8, indicate that in the second phase of category learning, there was an advantage for the 90- over 45-degree rotation condition in these integral dimension conditions.

This is somewhat surprising, given that categorizations related by 90 degrees are completely incompatible in regard to their selective-attention demands. The dimension that was originally completely irrelevant becomes completely relevant. In the 45-degree condition, the originally relevant dimension is at least partially relevant later. However, categorizations related by 90 degrees do have an advantage as far as dimensional organization. The relevant dimensions for the two categorizations are compatible with each other in the

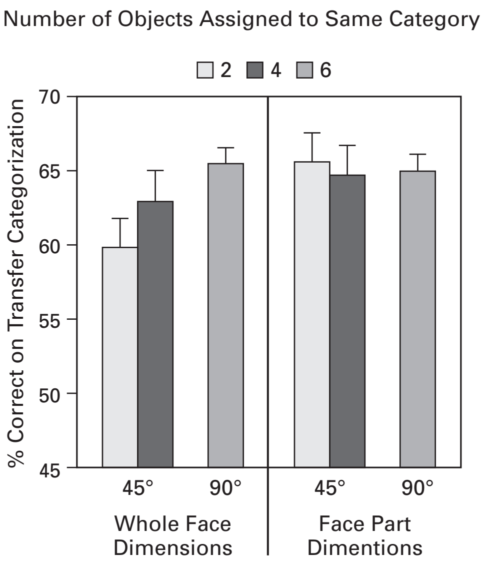

Figure 9.8
The 90-degree-rotation condition produced better transfer on the final categorization than the 45-degree condition, but only for overlapping face dimensions such as those in figure 9.7, and not spatially separated dimensions such as eyes and mouth.

sense of relying on independent sources of variation. For example, acquiring one dimension in figure 9.7 is compatible with later acquiring the 90-degree rotation of this dimension, because both are independent dimensions that can coexist without interference.

By analogy, 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 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. Thus, categorization rules separated by 90 degrees are inconsistent with respect to their selective attention demands because the dimension that was originally attended must be ignored and vice versa. However, the rules are consistent with respect to their dimensional organization of stimuli.

Our account for the advantage of a 90- over 45-degree rule rotation is that only the former rotation maintains a compatible dimensional interpretation of the stimuli across the two categorizations. If the relatively good transfer in the 90-degree rotation condition is because the two categorizations encourage the same differentiation of the faces into dimensions rather than crosscutting dimensions, then we should not expect the 90-degree rotation condition to produce better performance when dimension differentiation is not required—that is, with more separable stimuli. This is exactly what was found when we created dimensions by morphing select face parts rather than entire dimensions. One dimension morphed the eyes from one face to the eyes of another face, and the other dimension morphed mouths. These dimensions are more separable than those shown in figure 9.7 because people can attend to one dimension without showing much interference due to irrelevant variation on the other dimension (Garner 1976). With these more separable dimensions, participants should be able to selectively attend one dimension without as much need to differentiate fused dimensions. As shown by the right panel of figure 9.8, with these dimensions, the advantage for the 90-degree rotation was no longer found. This again suggests that the 90-degree advantage is due to participants' learning to isolate two originally fused dimensions from each other, and transferring this knowledge.

This conclusion is controversial. Op de Beeck and his colleagues (2003) created stimuli composed of novel, spatially overlapping dimensions. The shapes in figure 9.9 were created by combining seven sinusoidal functions (each with three parameters: frequency, phase, and amplitude), referred to individually as radial frequency components (RFCs), into a single, complex curve and then bending these to create closed contours.

While five of the seven RFCs remained fixed, two were chosen to have their amplitudes varied to define a two-dimensional space. Op de Beeck et al. (2003) showed that these

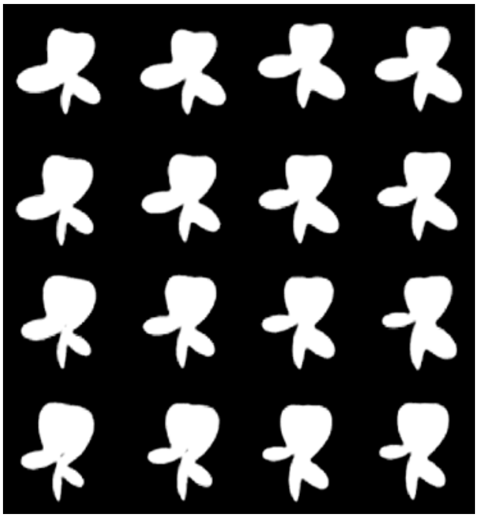

Figure 9.9
Stimuli used by Hockema et al. (2005), based on stimuli first developed by Op de Beeck et al. (2003), composed of two overlapping and integral dimensions. Looking from left to right, the amplitude of one radial frequency component (RFC) of the shapes is increased. Looking from top to bottom, a second RFC component is varied. These stimuli were purposefully blurred so that participants could not use local pixel features on the contour edges that might be correlated with the diagnostic dimensions. With these blurred stimuli participants needed to pay attention to relatively global aspects of the shape-based dimensions.

dimensions are relatively integral. With these stimuli there was no evidence that categorization via a horizontal or vertical boundary led to greater improvement in discriminability along the category-relevant dimension relative to the irrelevant dimension. On the basis of this result, they argued that category learning is only capable of altering weights to already separable dimensions, but not of making integral dimensions more separable.

Our dimension-differentiation account is not forced to predict that differentiation will occur for any set of dimensions within any set length of categorization training. Our account holds that some relatively integral dimensions can become differentiated, not that every arbitrary pair of indistinguishable dimensions can be well separated by categorization learning. Still, we found the RFC stimuli compelling, and were consequently interested in whether an improved training regime could lead to the differentiation of the RFC components (Hockema et al. 2005). In particular, we controlled the order of training trials to start with easy categorizations—shapes far from the category boundary—and gradually increase the difficulty to include shapes nearer to the boundary. Echoing classic results showing highly efficient learning with easy-to-hard training regimes (Mackintosh 1974), this training allowed our participants to eventually make categorizations that would have proved too difficult to learn in initial training. This training effectively challenged the perceptual system to adapt in order to support the categorization. After training, participants were better able to discriminate between stimuli near the categorization boundary that varied along the relevant, compared to irrelevant, dimension. This result indicates the kind of selective sensitization of a single dimension for which Op de Beeck and colleagues (2003) failed to find evidence. Further experiments showed that dynamic animations, varying on single dimensions, also successfully trained selective sensitization for an originally integral pair of dimensions. In light of these results, our current claim is that not all categorization training that would benefit from dimension differentiation will result in the desired differentiation. However, if care is taken to create training situations that push the perceptual system beyond its original capacity, then dimensions that were originally psychologically fused are not necessarily doomed to remain that way. Perceptual learning involves learning to selectively attend to relevant dimensions, but can also involve establishing what dimensions are available for selective attention in the first place. People not only learn appropriate weights for dimensions but also learn how to learn attentional weights for dimensions.

A Computational Reconciliation

Unitization involves the construction of a single functional unit out of component parts. Dimension differentiation divides wholes into separate 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 frequently, with all parts indicating a similar response. Thus, unitization and differentiation are both processes that build appropriately sized representations for the tasks at hand.

We have developed computational models to show how concept learning can lead to learning new perceptual organizations via unitization and differentiation (Gerganov et al. 2007; Goldstone 2003). In this pursuit, we have been drawn to neural networks that possess units that intervene between inputs and outputs and are capable of creating internal representations. For the current purposes, these intervening units can be interpreted as learned-feature detectors, and represent an organism's acquired perceptual vocabulary. Just as we perceive the world through the filter of our perceptual system, so the neural network does not have direct access to the input patterns, but rather only has access to the detectors that it develops.

The conceptual and perceptual learning by unitization and segmentation model, or CPLUS, is given a set of pictures as inputs and produces as output a categorization of each picture. Along the way to this categorization, the model comes up with a description of how the picture is segmented into pieces. The segmentation that CPLUS creates will tend to involve parts that (1) obey the Gestalt laws of perceptual organization by connecting object parts that have similar locations and orientations, (2) occur frequently in the set of presented pictures, and (3) are diagnostic for the categorization. For example, if the five input pictures of figure 9.10 are presented to the network and labeled as belonging to category A or category B, then originally random detectors typically become differentiated as shown.

This adaptation of the detectors reveals three important behavioral tendencies. First, detectors are created for parts that recur across the five objects, such as the lower square and upper rectangular antenna. Thus, the first input picture on the left will be represented by combining responses of the square and rectangular antenna detectors. Second, single, holistic detectors are created for objects such as the rightmost input picture that do not share any large pieces with other inputs. In this way, the model can explain how the same learning process unitizes complex configurations and differentiates other inputs into pieces. Third, the detectors act as filters that lie between the actual inputs and the categories. The

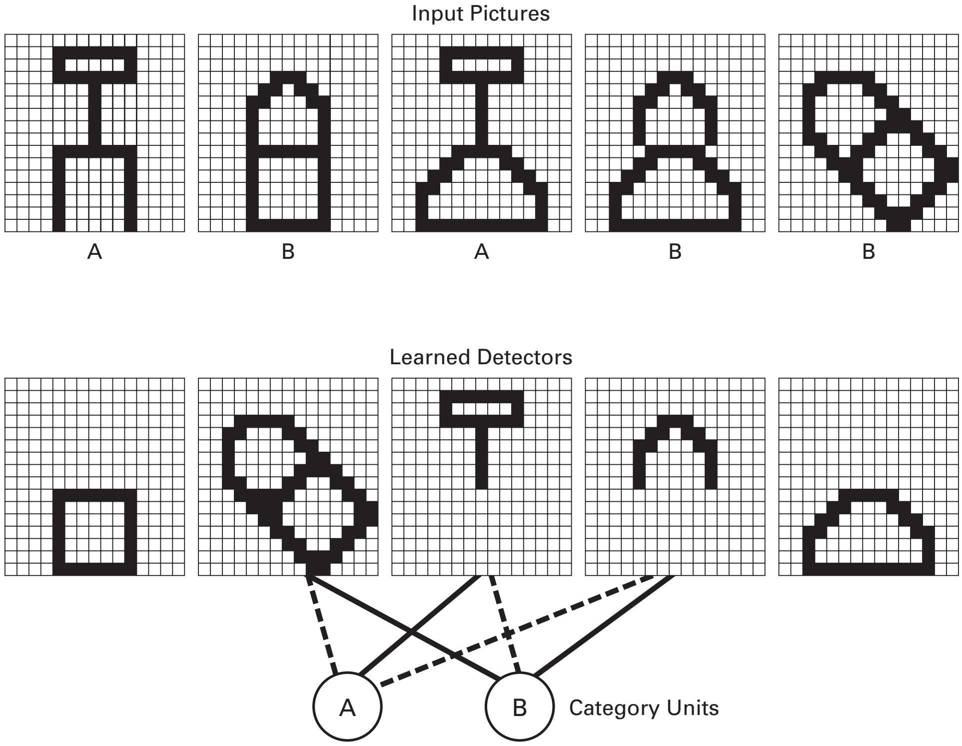

Figure 9.10
Sample output from the CPLUS (conceptual and perceptual learning by unitization and segmentation) model. After being exposed to the input pictures and their categorizations, the neural network creates detectors that can be assembled, like building blocks, to recreate the inputs. The detectors are learned at the same time that they are associated to categories. Solid lines represent excitatory connections; dashed lines represent inhibitory connections.

learned connections between the acquired detectors and the categories are shown by thick solid lines for positive connections and dashed lines for negative connections. The network learns to decompose the leftmost input picture into a square and rectangular antenna, but also learns that only the rectangular antenna is diagnostic for categorization, predicting that category A is present and that category B is not.

Although the mathematical details of the model are described elsewhere (Goldstone 2003), it is possible to give a functional description of the basic workings of the model that allow it to both separate stimuli into useful parts and create larger units. The heart of CPLUS is a competitive learning process (Rumelhart and Zipser 1985) in which detectors compete for the "right" to adapt toward randomly presented input patterns. In competitive learning algorithms, detectors start out as homogenous and undifferentiated. As inputs are presented, the detector that is most similar to a presented input adjusts its weights so that it is even more similar to the input, and inhibits the other detectors from learning to adapt toward the input. This leaves the other detectors available to become specialized for a different class of patterns. The originally homogenous detectors will be differentiated over time, and will split the input patterns into two categories. For another example of how small physical differences between detectors can developmentally snowball into substantially different modules, see Johnson (2000; see also chapter 15, this volume).

Competitive learning is typically applied to unsupervised categorization, in which case it takes entire patterns as input, and sorts these complete, whole input patterns into separate categories. However, CPLUS applies competitive learning to the problem of segmentation, by taking a single input pattern and sorting the pieces of the pattern into separate groups. Consistent with competitive learning, 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. This technique, by itself, segments a pattern into complementary parts. If one detector becomes specialized for a pixel, the other detector does not. Unfortunately, each detector can become specialized for a random set of pixels, rather than a coherent, psychologically plausible segmentation.

To create psychologically plausible segmentations, we modify the determination of winners. Topological constraints on detector creation are incorporated by two mechanisms: (1) input-to-detector weights "leak" to their neighbors by an amount proportional to their proximity in space, and (2) 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 Vs, because the two halves of a diagonal line will be linked by their common orientation. Thus, if a detector wins for pixel Y (meaning that the detector receives more activation when pixel Y 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 Y. For an alternative approach to segmentation that uses synchronized oscillations rather than architecturally separated detectors to represent segmentations, see Mozer et al. (1992).

As described thus far, the algorithm is completely unsupervised, creating detectors as a function of statistical dependencies and bottom-up perceptual properties of the set of input images. However, much of the experimental evidence previously reviewed indicates a strong influence of learned categories on acquired perceptual encodings. This influence is incorporated into CPLUS by biasing diagnostic detectors to win the competition to learn the pattern. The diagnosticity of a detector is assumed to be directly proportional to the weight from the detector to the category 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 fact, in the actual operation of CPLUS, the detectors adapt at the same time that they become associated with categories to be learned. This core assumption of CPLUS allows it to create detectors only when they are useful for an important categorization rather than having to postulate a large initial set of detectors just in case one is needed for a future categorization task.

CPLUS differs from most other models of categorization in that it prominently features a perceptual segmentation process. Its ability to flexibly organize a pattern into learned parts is a large advantage to the extent that the world consists of objects that have parts that recur many times across many objects. CPLUS is inherently a componential model in which objects are broken down into parts during perception, and these parts are differentially associated with categories. Admittedly, there is little advantage to creating compositional representations for the five input pictures in figure 9.10. They can be represented by five holistic detectors just as effectively as by the shown componential representation. However, if we had a set of objects that each had five pieces, and each of those pieces had four variants, then the holistic strategy would require 1,024 (4^5) detectors whereas CPLUS requires only 20. In a world where objects are built from elementary blocks, a perceptual system that can take advantage of this fact stands to benefit considerably, both in terms of representational efficiency and generative flexibility (see also Griffiths and Ghahramani, 2006). A categorization advantage is also accrued when the building blocks are diagnostic for needed categories.

Recently we have incorporated even greater plasticity in the creation of feature detectors in CPLUS. The original version of CPLUS incorporated a hard-wired pressure to create detectors for stimulus elements that are close and create smoothly varying curves. Gerganov and colleagues (2007) have explored the possibility that these constraints themselves may be learned rather than hard-wired. If a set of input patterns mostly contains connected and smooth elements, then a neural network can internalize these regularities as it develops. Other researchers have proposed that visual detectors can be created by a system that simply internalizes statistical regularities extracted from a large set of natural photographic images (Olshausen and Field 1996); however, their detectors were assumed to adapt on evolutionary time scales, and hence be built in for any modern individual. In contrast, the newer version of CPLUS assumes that the process of learning constraints from an environment can take place in an individual's own lifetime. The empirical basis for this contention stems from developmental studies suggesting that some perceptual constraints appear to be learned rather than innate (Quinn and Bhatt 2006; Sheya and Smith 2006). Learning without constraints is impossible, but the exciting possibility still remains that constraints themselves can be learned.

Conclusion

The previously described neural network builds detectors at the same time that it builds connections between the detectors and categories. The psychological implication is that our perceptual systems do not have to be set in place before we start to use them. The concepts we need can and should influence the perceptual units we create. The influence of these concepts comes in at least two forms: unitizing originally individuated elements, and differentiating originally fused elements. Rather than viewing unitization and differentiation as contradictory, they are best viewed as aspects of the same process that bundles stimulus components if they diagnostically co-occur and separates these bundles from other statistically independent bundles. Under this conception, learning new features or detectors consists in learning how to carve a stimulus into useful components.

One of the most powerful ideas in cognitive science has been the notion that flexible cognition works by assembling a fixed set of building blocks into novel arrangements. Many of the most notable discoveries in cognitive science have involved finding these kinds of compositional encodings. In linguistics, phonemes have been represented by the presence or absence of fewer than twelve features such as *voiced*, *nasal*, and *strident* (Jakobson et al. 1963). Scenarios, such as ordering food in a restaurant, have been represented by Schank (1972) in terms of a set of twenty-three primitive concepts such as *physical-transfer, propel, grasp,* and *ingest*. In the field of object recognition, Biederman (1987) proposed a set of thirty-six geometric shapes such as *wedge* and *cylinder* to be used for representing objects such as telephones and flashlights. We are in complete agreement with these proposals in terms of the cognitive and computational advantage of creating representations by composing elements. The only difference, but a critical one, is that we believe these elements can be flexibly created during experience with an environment, rather than being fixed.

We have argued that the concepts we learn can reach down and influence the very perceptual descriptions that ground the concepts. This interactive cycle is figuratively shown in figure 9.11.

A person creates perceptual building blocks from his or her experiences in the world. Then, the person's subsequent experience of this same world is influenced by these learned building blocks. Naturally, cases of experience-induced hallucinations of figure 9.11's extremity are rare (but possible; Grossberg 2000), but this is a graphic, degenerate case

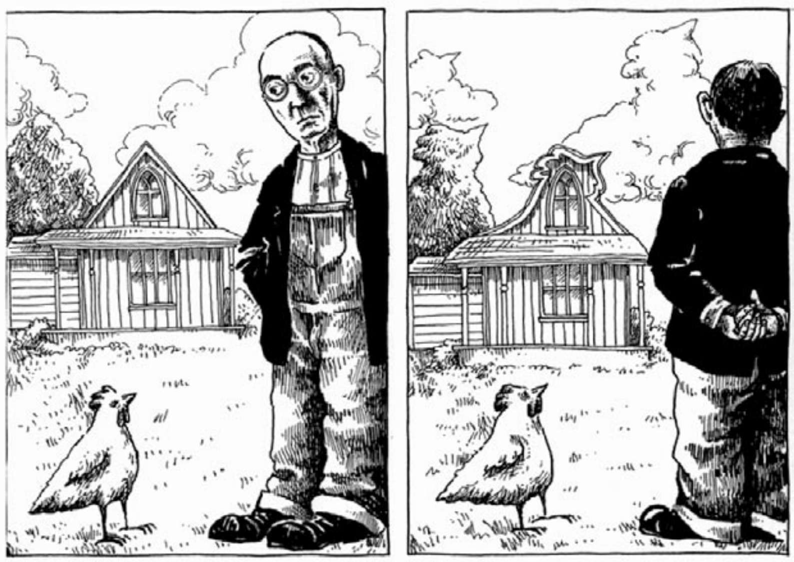

Figure 9.11
We learn about chickens from the world (left panel), and then turn around and interpret the world in terms of the chickens that we have learned (right panel). (Idea: Robert Goldstone; artwork: Joe Lee.)

of the everyday phenomenon in which our perceptions and conceptions are tightly coupled (Wisniewski and Medin 1994). Hopefully, the CPLUS model provides some reassurance that this interactive loop between perception and conception need not be viciously circular. 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.

Acknowledgments

The research reported in this chapter has benefited greatly from comments and suggestions by Drew Hendrickson, Amy Needham, Mary Peterson, Paul Quinn, Richard Shiffrin, and Linda Smith. This research was funded by Department of Education, Institute of Education Sciences grant R305H050116, and National Science Foundation ROLE grant 0527920.

References

Behrmann M, Zemel RS, Mozer MC (1998) Object-based attention and occlusion: Evidence from normal participants and a computational model. J Exp Psychol Hum Percept Perform 24: 1011–1036.

Biederman I (1987) Recognition-by-components: A theory of human image understanding. Psychol Rev 94: 115–147.

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

Burns B, Shepp BE (1988) Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. Percept Psychophys 43: 494–507.

Busey TA, Vanderkolk JR (2005) Behavioral and electrophysiological evidence for configural processing in fingerprint experts. Vision Res 45: 431–448.

Cattell JM (1886) The time it takes to see and name objects. Mind 11: 63–65.

Elbert T, Pantev C, Wienbruch C, Rockstro, B, Taub E (1995) Increased cortical representation of the fingers of the left hand in string players. Science 270: 305–307.

Fahle M (1994) Human pattern recognition: Parallel processing and perceptual learning. Perception 23: 411–427.

Fahle M, Poggio T (2002) Perceptual learning. Cambridge, MA: MIT Press.

Garner WR (1976) Interaction of stimulus dimensions in concept and choice processes. Cognitive Psychol 8: 98–123.

Garner WR (1978) Selective attention to attributes and to stimuli. J Exp Psychology Gen 107: 287–308.

Gauthier I, Curran T, Curby KM, Collins D (2003) Perceptual interference evidence for a nonmodular account of face processing. Nat Neurosc 6: 428–432.

Gerganov A, Grinberg M, Quinn PC, Goldstone RL (2007) Simulating conceptually guided perceptual learning. In: Proceedings of the twenty-ninth Annual Conference of the Cognitive Science Society (McNamara DS, Trafton JG, eds), 287–292. Nashville, TN: Cognitive Science Society.

Gibson EJ (1991) An odyssey in learning and perception. Cambridge, MA: MIT Press.

Goldstone RL (1994) Influences of categorization on perceptual discrimination. J Exp Psychol Gen 123: 178–200.

Goldstone RL (2003) Learning to perceive while perceiving to learn. In: Perceptual organization in vision: Behavioral and neural perspectives (Kimchi R, Behrmann M, Olson C, eds), 233–278. Hillsdale, NJ: Lawrence Erlbaum.

Goldstone RL (1998) Perceptual learning. Annu Rev Psychol 49: 585–612.

Goldstone RL (2000) Unitization during category learning. J Exp Psychol Hum Percept Perform 26: 86–112.

Goldstone RL, Stevyers M (2001) The sensitization and differentiation of dimensions during category learning. J Exp Psychol Gen 130: 116–139.

Griffiths TL, Ghahramani Z (2006) Infinite latent feature models and the Indian buffet process. Adv Neural Info Proc Sys 18.

Grossberg S (2000) How hallucinations may arise from brain mechanisms of learning, attention, and volition. J Int Neuropsych Soc 6: 579–588.

Hockema SA, Blair MR, Goldstone RL (2005) Differentiation for novel dimensions. In: Proceedings of the twenty-seventh Annual Conference of the Cognitive Science Society (Bara BG, Barsalou L, Bucciarelli M, eds), 953–958. Hillsdale, NJ: Lawrence Erlbaum.

Hoffman DD, Richards WA (1984) Parts of recognition. Cognition 18: 65–96.

Jakobson R, Fant G, Halle M (1963) Preliminaries to speech analysis: The distinctive features and their correlates. Cambridge, MA: MIT Press.

Johnson MH (2000) Functional brain development in infants: Elements of an interactive specialization framework. Child Dev 71: 75–81.

Karni A, Sagi D (1993) The time course of learning a visual skill. Nature 365: 250–252.

Kemler DG, Smith LB (1978) Is there a developmental trend from integrality to separability in perception? J Exp Child Psychol 26: 498–507.

LaBerge D (1973) Attention and the measurement of perceptual learning. Mem Cognition 1: 268–276.

Logan GD, Taylor SE, Etherton JL (1996) Attention in the acquisition and expression of automaticity. J Exp Psychol Learn 22: 620–638.

Logothetis NK, Pauls J, Poggio T (1995) Shape representation in the inferior temporal cortex of monkeys. Curr Biol 5: 552–563.

Mackintosh NJ (1974) The psychology of animal learning. London: Academic Press.

Melara RD, Marks LE, Potts BC (1993) Primacy of dimensions in color perception. J Exp Psychol Hum Perc Perform 19: 1082–1104.

Mozer MC, Zemel RS, Behrmann M, Williams CKI (1992) Learning to segment images using dynamic feature binding. Neural Comput 4: 650–665.

Needham A (1999) The role of shape in 4-month-old infants' segregation of adjacent objects. Infant Behav Dev 22*:* 161–178.

Needham A, Baillargeon R (1998) Effects of prior experience in 4.5-month-old infants' object segregation. Infant Behav Dev 21*:* 1–24.

Needham A, Dueker G, Lockhead G (2005) Infants' formation and use of categories to segregate objects. Cognition 94: 215–240.

Notman LA, Sowden PT, Özgen E (2005) The nature of learned categorical perception effects: A psychophysical approach. Cognition 95: B1–B14.

Olshausen BA, Field DJ (1996) Emergence of simple cell receptive field properties by learning a sparse code for natural images. Nature 381: 607–609.

Op de Beeck H, Wagemans J, Vogels R (2003) The effect of category learning on the representation of shape: Dimensions can be biased but not differentiated. J Exp Psychol Gen 132: 491–511.

Perrett DI, Oram MW (1993) Neurophysiology of shape processing. Image Vision Comput 11: 317–333.

Perrett DI, Smith PAJ, Potter DD, Mistlin AJ, Head AD, Jeeves MA (1984) Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. Hum Neurobiol 3: 197–208.

Peterson MA (1994) Object recognition processes can and do operate before figure-ground organization. Curr Dir Psychol Sci 3: 105–111.

Peterson MA, Harvey EH, Weidenbacher HL (1991) Shape recognition inputs to figure-ground organization: Which route counts? J Exp Psychol Hum Perc Perform 17: 1075–1089.

Peterson MA, Gibson BS (1994) Must figure-ground organization precede object recognition? An assumption in peril. Psychol Sci 5: 253–259.

Peterson MA, Lampignano DW (2003) Implicit memory of novel figure-ground displays includes a history of cross-border competition. J Exp Psychol Hum Perc Perform 29: 808–822.

Puel JL, Bonfils P, Pujol R (1988) Selective attention modifies the active micromechanical properties of the cochlea. Brain Res 447: 380–383.

Quinn PC, Bhatt RS (2006) Are some Gestalt principles deployed more readily than others during early development? The case of lightness versus form similarity. J Exp Psychol Hum Perc Perform 32: 1221–1230.

Quinn PC, Schyns P (2003) What goes up may come down: Perceptual process and knowledge access in the organization of complex visual patterns by young infants. Cognitive Sci 27: 923–935.

Quinn PC, Schyns PG, Goldstone RL (2006) The interplay between perceptual organization and categorization in the representation of complex visual patterns by young infants. J Exp Child Psychol 95: 117–127.

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13: 87–103.

Recanzone GH, Merzenich MM, Jenkins WM (1992) Frequency discrimination training engaging a restricted skin surface results in an emergence of a cutaneous response zone in cortical area 3a. J Neurophysiol 67: 1057–1070.

Rumelhart DE, Zipser D (1985) Feature discovery by competitive learning. Cognitive Sci 9: 75–112.

Sagi D, Tanne D (1994) Perceptual learning: Learning to see. Curr Opin Neurobiol 4: 195–199.

Schank R (1972) Conceptual dependency: A theory of natural language understanding. Cognitive Psychol 3: 552–631.

Sheya A, Smith LB (2006) Perceptual features and the development of conceptual knowledge. Journal Cogn Dev 7: 455–476.

Shiffrin RM, Lightfoot N (1997) Perceptual learning of alphanumeric-like characters. In: The psychology of learning and motivation, volume 36 (Goldstone RL, Schyns PG, Medin DL, eds), 45–82. San Diego: Academic Press.

Shiu L, Pashler H (1992) Improvement in line orientation discrimination is retinally local but dependent on cognitive set. Percept Psychophys 52: 582–588.

Smith LB (1989) From global similarity to kinds of similarity: The construction of dimensions in development. In: Similarity and analogical reasoning (Vosniadou S, Ortony A, eds), 146–178. Cambridge: Cambridge University Press.

Smith LB, Gasser M, Sandhofer C (1997) Learning to talk about the properties of objects: A network model of the development of dimensions. In: The psychology of learning and motivation, volume 36 (Goldstone RL, Schyns PG, Medin DL, eds), 219–255. San Diego: Academic Press.

Smith LB, Kemler DG (1978) Levels of experienced dimensionality in children and adults. Cognitive Psychol 10: 502–532.

Smith PL (1990) Obtaining meaningful results from Fourier deconvolution of reaction time data. Psychol Bull 108: 533–550.

Sowden PT, Davies IRL, Roling P (2000) Perceptual learning of the detection of features in X-ray images: A functional role for improvements in adults' visual sensitivity? J Exp Psychol Hum Perc Perform 26: 379–390.

Steyvers M (1999) Morphing techniques for generating and manipulating face images. Behav Res Meth Ins C 31: 359–369.

Suga N, Ma X (2003) Multiparametric corticofugal modulation and plasticity of the auditory system. Nat Rev Neurosci 4: 783–794.

Tanaka JW, Curran T (2001) A neural basis for expert object recognition. Psychol Sci 12: 43–47.

Townsend JT, Wenger MJ (2004) A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. Psychol Rev 111: 1003–1035.

Watanabe T, Náñez J, Sasaki Y (2001). Perceptual learning without perception. Nature 413: 844–848.

Wisniewski EJ, Medin DL (1994). On the interaction of theory and data in concept learning. Cognitive Sci 18: 221–281.

Zemel RS, Behrmann M, Mozer MC, Bavelier D (2002) Experience-dependent perceptual grouping and object-based attention. J Exp Psychol Hum Perc Perform 28: 202–217.