Perception & Psychophysics
2001, 63 (7), 1183-1200

Separating perceptual processes from decisional processes in identification and categorization

W. TODD MADDOX
University of Texas, Austin, Texas

Four observers completed perceptual matching, identification, and categorization tasks using separable-dimension stimuli. A unified quantitative approach relating perceptual matching, identification, and categorization was proposed and tested. The approach derives from general recognition theory (Ashby & Townsend, 1986) and provides a powerful method for quantifying the separate influences of perceptual processes and decisional processes within and across tasks. Good accounts of the identification data were obtained from an initial perceptual representation derived from perceptual matching. The same perceptual representation provided a good account of the categorization data, except when selective attention to one stimulus dimension was required. Selective attention altered the perceptual representation by decreasing the perceptual variance along the attended dimension. These findings suggest that a complete understanding of identification and categorization performance requires an understanding of perceptual and decisional processes. Implications for other psychological tasks are discussed.

An important goal of psychological inquiry is to understand how behavior is influenced by the environmental stimulation and the task at hand. Information about the environment is obtained through the sensory systems, and the task at hand determines which information is relevant for making a decision. A detailed understanding of this complex system requires knowledge of the interplay between perceptual and decisional processes and thus requires a theory that acknowledges their separate and unique influences. The importance of separating perceptual from decisional processes in simple tasks such as detection has been well known since the early days of signal detection theory (Green & Swets, 1967). Unfortunately, the importance of this distinction has been less well understood in predicting performance in more complex tasks such as identification and categorization (however, see Ashby & Lee, 1991; see also Geisler & Chou, 1995, for an application to visual search). The most common approach is to downplay the importance of perceptual processes by essentially ignoring them or by making overly simplified assumptions, such as assuming that each stimulus is represented by a point in some multidimensional psychological space (e.g., Nosofsky, 1986). This provides an inadequate model of perceptual processing for real-world stimuli that are characterized by noisy perceptual effects (Ashby & Lee, 1993; Green & Swets, 1967) generally exaggerating the importance of decision processes while assuming (often implicitly) that perceptual processes remain unchanged. Even when models that assume noisy perceptual effects are utilized, often the perceptual noise assumptions are overly simplistic (e.g., equal, uncorrelated perceptual noise across dimensions and stimuli).

In certain situations, perceptual processes and the resulting perceptual representation likely remain unchanged across different types of tasks. In these cases, different tasks should require different decision strategies while leaving the perceptual representation unchanged. In other cases, though, the task at hand might alter the nature of the perceptual representation. For example, when only some aspects of the environment (or stimulus) are relevant to solving a particular problem, it may be the case that perceptual processing is altered to fine-tune or sharpen the perceptual representation along the relevant dimension. Given the large amount of stimulation from the environment at any one time and the limited capacity of our attention system, this type of selective attention is likely commonplace.

The overriding goal of the present study was to examine the separate effects of perceptual and decisional processes on performance in two complex tasks: identification and categorization. Identification and categorization were chosen because they are primary components of many behaviors of all organisms and thus are of fundamental importance. Two specific hypotheses were of interest. First, a rigorous test of the (often implicit) assumption that changes in task demand affect decisional processes while leaving the perceptual representation unchanged was undertaken. Second, the hypothesis that decisional selective attention tasks (i.e., tasks that require the observer to place all weight in their decision on one aspect of the stimulus, while ig-

This research was supported in part by National Science Foundation Grant SBR-9796206 and NIH Grant R01 MH59196. The author thanks Greg Ashby, Sergei Bogdanov, Corey Bohil, Leslie Cohen, Randy Diehl, Jeffrey Dodd, Wilson Geisler, and Arthur Markman for discussions that influenced this work and Greg Ashby, Anne-Marie Bonnel, Robert Nosofsky, and two anonymous reviewers for comments on an earlier version of this manuscript. The author is especially grateful to Sergei Bogdanov for developing the computer programs used to test the observers. Correspondence should be addressed to W. T. Maddox, Department of Psychology, Mezes Hall 330, Mail Code B3800, University of Texas, Austin, TX 78712 (e-mail: maddox@psy.utexas.edu).

noring all other aspects of the stimulus; Maddox & Ashby, 1998) affect both decisional and perceptual processes was examined. Selective attention problems of this sort are ubiquitous in everyday life and thus demand a better understanding. To achieve these goals requires a theory that makes reasonable assumptions about the nature of the perceptual representation and that contains parameters associated with perceptual and decisional processes that are separately identifiable. General recognition theory (Ashby & Townsend, 1986), a multidimensional generalization of signal detection theory, provides such a theory.

The next two (second and third) sections introduce briefly general recognition theory and review the relevant literature. The details of the theory can be found in numerous other articles (Ashby & Lee, 1991; Ashby & Perrin, 1988; Ashby & Townsend, 1986). The fourth section introduces the experiment and details the method. The fifth section reports the results and theoretical analyses. The sixth section provides a summary and discussion.

GENERAL RECOGNITION THEORY

Perceptual Representation Assumptions

General recognition theory acknowledges that a fundamental characteristic of all perceptual systems is that repeated presentations of the same stimulus yield different perceptual effects—that is, that perceptual noise exists (e.g., Ashby & Lee, 1993; Ashby & Townsend, 1986; Geisler, 1989; Green & Swets, 1967). General recognition theory assumes that a single multidimensional stimulus can be represented perceptually by a multivariate probability distribution (Ashby & Lee, 1993). For a two-dimensional stimulus, a bivariate normal distribution is assumed to describe the set of percepts. A bivariate normal distribution is described by a mean and variance along each dimension, as well as a covariance term, μ_x, μ_y, σ_x^2, σ_y^2, cov_{xy}, where the subscripts x and y denote dimensions x and y. Figure 1a depicts hypothetical equal likelihood contours for four stimuli constructed from the factorial combination of two levels along two dimensions x and y. With bivariate normal distributions, the equal likelihood contours are always circular or elliptical. In general recognition theory, *perceptual independence* holds for a single stimulus if and only if the perceptual effects for dimensions x and y are statistically independent (see Ashby, 1988; Ashby & Maddox, 1991; Ashby & Townsend, 1986; Perrin & Ashby, 1991). With bivariate normal distributions, perceptual independence holds when the major and minor axes of the contour are parallel to the coordinate axes (i.e., when the covariance is zero). Perceptual independence holds for all four distributions in Figure 1a. A positive slope for the major axis implies a positive perceptual dependence, and a negative slope implies a negative perceptual dependence. *Perceptual separability* holds when the distribution of percepts associated with one component is unaffected by the level of the other component (Ashby & Townsend, 1986; Maddox, 1992). In Figure 1a, dimension y is perceptually separable from dimension x, but dimension x is not perceptually separable from dimension y.

Decision Process Assumptions

In general recognition theory, the experienced observer learns to divide the perceptual space into response regions and assigns a response to each region. The partitions between response regions are called *decision bounds.* On each trial, the observer determines the location of the perceptual effect and gives the response associated with that region of the perceptual space. Decision bounds and the

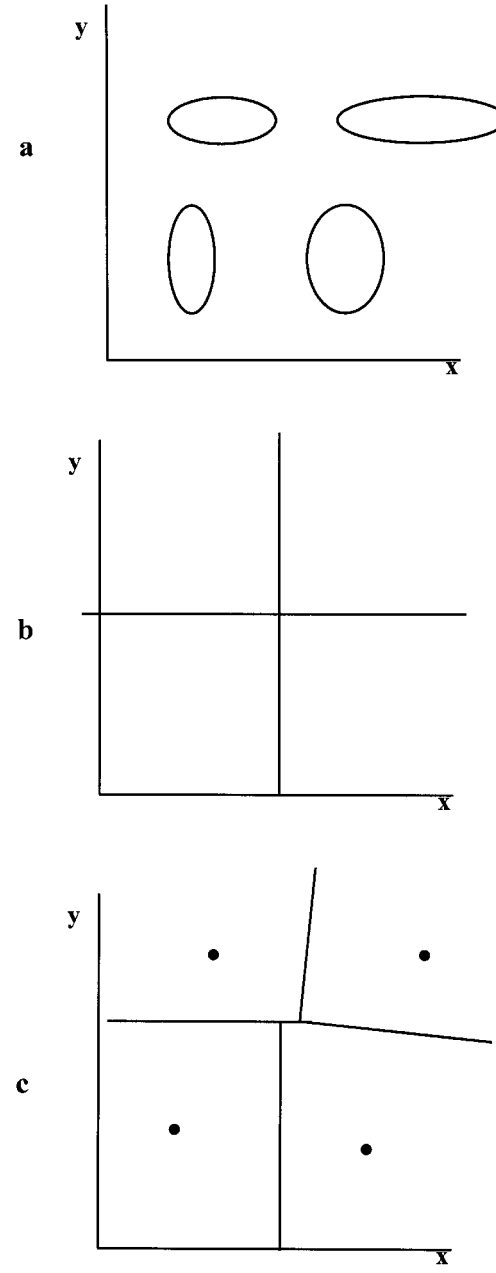

Figure 1. (a) Hypothetical contours of equal likelihood for four stimuli constructed from the factorial combination of two levels along two stimulus components. (b) Hypothetical response regions and decisional separability bounds for the same four stimuli. (c) Hypothetical response regions and striatal units for the same four stimuli.

resulting response regions can come in many forms (Ashby, Waldron, Lee, & Berkman, 2001). Decision bounds that are parallel to the coordinate axes are of particular interest because they result when the observer sets a criterion along one stimulus dimension and ignores the other dimension. This decision strategy is referred to as *decisional separability* in general recognition theory and is generally the optimal strategy in decisional selective attention tasks (Ashby & Townsend, 1986; Maddox, 1992). Hypothetical response regions and decisional separability bounds for the four stimuli in Figure 1a are displayed in Figure 1b. Details of the response regions examined in this article will be reserved for the Results and Theoretical Analyses section.

REVIEW OF THE RELEVANT LITERATURE

Only a few studies in the literature have tested rigorously the hypothesis that changing task demands affect decision processes while leaving perceptual processes unchanged. Ashby and Lee (1991; see also Maddox & Ashby, 1996) tested this hypothesis with data from an identification task and several categorization tasks. They applied a model derived from general recognition theory to the identification confusions to estimate the perceptual representation. Holding this perceptual representation fixed and allowing changes in the decision process across categorization conditions provided a good quantitative account of the categorization data. Interestingly, though, they found that a perceptual representation that satisfied perceptual separability provided a better account of the categorization data than did the perceptual representation estimated from the identification data. Unfortunately, Ashby and Lee (1991) did not examine the possibility that the perceptual representation might be altered differently by different categorization decision rules. They assumed that the perceptual representation was the same across all categorization conditions.

Maddox and Bogdanov (2000) followed up on this work by investigating the possibility that decisional selective attention categorization problems alter both perceptual and decisional processes, whereas categorization problems that require attention to all stimulus dimensions (i.e., decisional integration tasks) alter only decisional processes. The possibility that decisional selective attention tasks might have a unique influence on perceptual and decisional processes is suggested by the work of Shepard, Hovland, and Jenkins (1961; Shepard & Chang, 1963) and Nosofsky (1986), who found that the similarity relations among exemplars in a multidimensional scaling (MDS) psychological space were altered in a specific way in the decisional selective attention conditions, but not in the decisional integration conditions. Unfortunately, within their theoretical framework, it is not possible to determine whether this selective attention effect is perceptual, decisional, or both.

To achieve their objective, Maddox and Bogdanov (2000) had observers complete two tasks using the same set of simple line stimuli. Instead of using an identification task, they used a matching task (described in detail in the Method section; Alfonso-Reese, 1996, 1997) to derive an initial perceptual representation. They then had observers complete decisional integration categorization conditions and a decisional selective attention categorization condition. Maddox and Bogdanov used the perceptual representation derived from matching to predict performance across conditions. They found support for a model that assumed that the decisional selective attention categorization condition altered both perceptual and decisional processes, whereas the decisional integration categorization conditions altered only decisional processes. The model parameters indicated that decisional selective attention led to a decrease in perceptual variability along the decisionally attended dimension. The notion that selective attention can reduce perceptual variability has a long history in signal detection theory (e.g., Braida & Durlach, 1972; Luce & Green, 1978; Macmillan, Goldberg, & Braida, 1988; see also Bonnel & Hafter, 1998).

The present research builds on the previous body of research in a number of ways. Prior to Maddox and Bogdanov (2000), all studies that examined the identification–categorization relationship used the identification data to derive an initial MDS psychological space (e.g., Nosofsky, 1986, 1987, 1989) or an initial perceptual representation (e.g., Ashby & Lee, 1991; Maddox & Ashby, 1996). This initial psychological space or perceptual representation was then used to predict categorization performance. To my knowledge, no attempt has been made previously to predict identification performance from an initial psychological space or perceptual representation estimated from another task. The present study used general recognition theory to predict both identification and categorization performance from a perceptual representation derived from the matching task. Specifically, the hypothesis that the perceptual representation is identical in both identification and categorization was tested. In addition, general recognition theory was used to predict categorization performance from a perceptual representation derived from the identification data. Of particular interest was a comparison of the matching task perceptual representation with the identification task perceptual representation in their ability to account for categorization performance. The hypothesis that decisional selective attention tasks alter the decision strategy and reduce the perceptual variability along the attended dimension was tested. Finally, a preliminary investigation of possible perceptual representation and decision process changes that might occur with categorization training was undertaken by partitioning the categorization data into two phases: early and later training.

PERCEPTUAL MATCHING, IDENTIFICATION, AND CATEGORIZATION EXPERIMENT

In the present experiment, 4 observers completed a large number of sessions in each of three tasks: perceptual matching, identification, and categorization. The same stimulus

ensemble was used in each task. The stimuli were 18 lines (see Figure 2a) that varied in length and orientation. The 4 observers first completed ten 360-trial sessions of perceptual matching, followed by twelve 480-trial sessions of identification. The categorization task followed and consisted of four separate conditions. The 4 observers completed fifteen 800-trial sessions in each of four categorization conditions. Thus, each observer completed a total of 72 sessions and 57,360 trials in the complete experiment. Two decisional selective attention categorization conditions were included: one with line length relevant (DSAL; Figure 2b), and the other with line orientation relevant (DSAO; Figure 2c). Two decisional integration conditions were included: One required a linear integration strategy (LI; Figure 2d), and the other required a nonlinear integration strategy (NLI; Figure 2e).

The approach was to use the perceptual matching data to estimate an initial perceptual representation. The resulting perceptual representation was then held fixed in an attempt to model identification and categorization. Specifi-

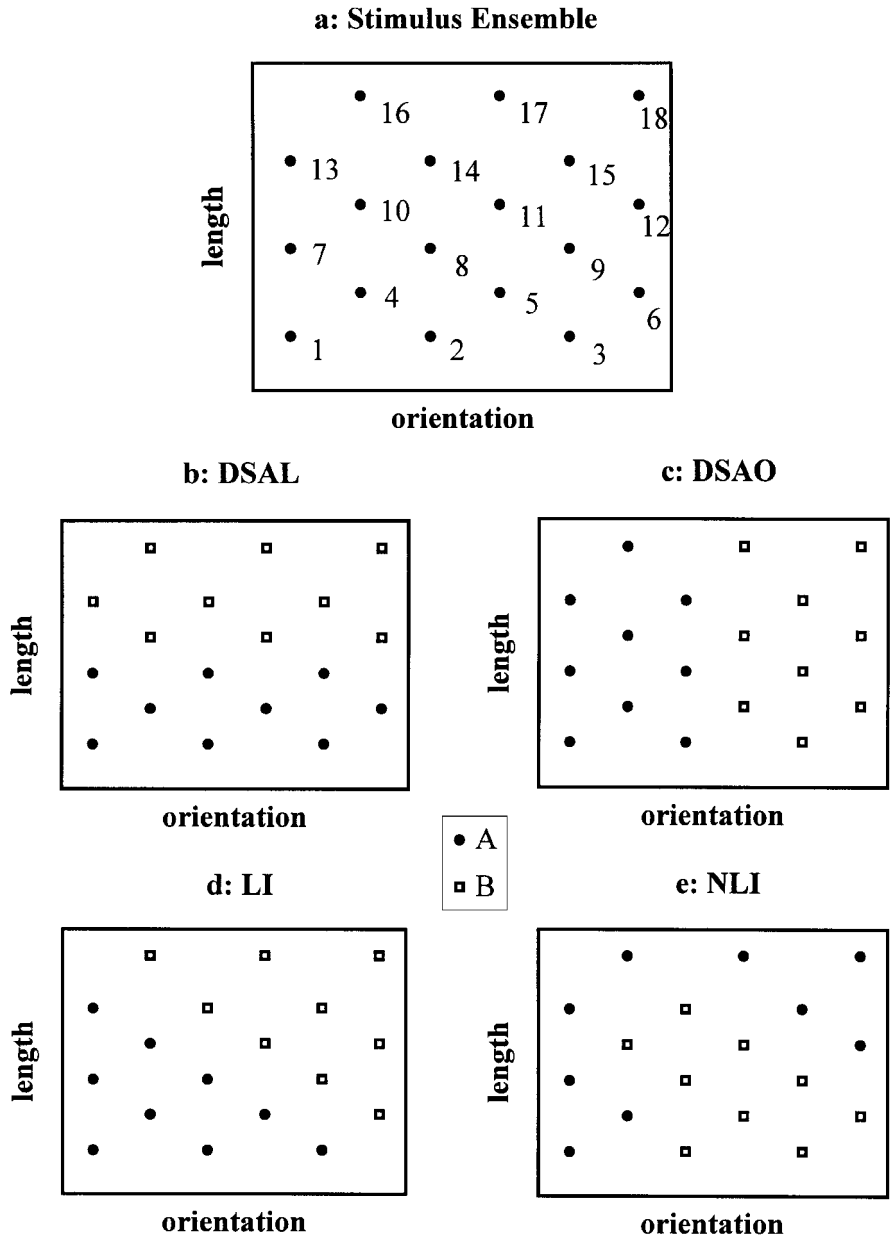

Figure 2. (a) Stimulus ensemble and numbering scheme for the 18 stimuli used in all tasks. Stimulus structure and stimulus-to-category assignments for the (b) decisional selective attention to length (DSAL), (c) decisional selective attention to orientation (DSAO), (d) linear integration (LI), and (e) nonlinear integration (NLI) categorization conditions.

cally, both the identification response frequencies (i.e., confusion matrix entries) and the categorization response frequencies for each stimulus were predicted quantitatively from the matching task perceptual representation. In categorization, a test of the hypothesis that decisional selective attention reduced the perceptual variability along the attended dimension, relative to the perceptual variability along the unattended dimension, was undertaken. Several hypotheses regarding the nature of the response regions were also examined. All analyses were performed at the level of the individual observer. Data were not collapsed across observers. Maddox (1999; see also Maddox & Ashby, 1998) showed that averaging often alters the structure of the data in such a way that the correct model of individual observer performance might provide a poor account of the aggregate data (see also Ashby, Maddox, & Lee, 1994; Estes, 1956; Smith & Minda, 1998).

Method

Observers

Four observers were solicited from the University of Texas community and were paid for their participation. All observers had prior experience with similar experiments and stimuli. All observers had 20/20 vision or vision corrected to 20/20. The same 4 observers participated in all three tasks.

Stimuli

The stimulus set consisted of 18 lines. These stimuli represent a subset taken from a larger group of 36 stimuli that were constructed from six levels of line length and six levels of line orientation. The six line lengths were 140, 144, 148, 152, 156, and 160 pixels. The six line orientations were .377, .427, .477, .527, .577, and .627 radians. The 18 stimuli and a numbering scheme are depicted in Figure 2a. Previous research suggests that length and orientation are separable (e.g., Garner, 1974; Garner & Felfoldy, 1970; Shepard, 1964; however, see Ashby & Lee, 1991, Ashby & Maddox, 1990, and Maddox, 1992). The observers were seated approximately 40 in. from the computer screen, and each stimulus subtended a visual angle of approximately 1°. The stimuli were computer-generated and were displayed on a 21-in. monitor with 1,360 × 1,024 resolution in a dimly lit room. Each line was presented in white on a black background. To minimize line jaggedness, Alfonso-Reese's (1997) antialiasing routine, developed for use with Brainard's (1997) Psychophysics Toolbox, was applied.

Procedure

Matching task. The first session was considered as practice and was excluded from further analyses. On each trial, one of the 18 stimuli was selected at random with equal probability. This served as the standard stimulus on that trial. The center of the standard stimulus was presented 150 pixels to the left of the horizontal center of the computer screen. The vertical position of the center of the standard stimulus was normally distributed with a mean of zero and a standard deviation of 5.5 pixels. The center of the comparison stimulus was presented 150 pixels to the right of the horizontal center of the computer screen, and the vertical center was normally distributed with a mean of zero and a standard deviation of 5.5 pixels. The stimuli were offset vertically to minimize the possibility that the observer would use a matching strategy based solely on the endpoints of the stimuli. At the start of each trial, the comparison stimulus was a dot. The observer's task was to adjust the length and orientation of the comparison stimulus until it was perceived (judged) to "match" the standard stimulus. The mouse controlled the adjustments. Once the observer perceived the comparison and standard stimuli to match, he or she pressed the mouse button. The observer then pressed a button to "accept" the match or to "reject" the match and begin again. The observers were instructed "to be as accurate as possible." Each session consisted of four blocks of 90 trials, with breaks between each block.

Identification task. The first session was considered as practice and was excluded from further analyses. On each trial, one of the 18 stimuli was selected at random with equal probability. The stimulus was presented centered on the computer monitor for 100 msec, followed by a pattern mask and then a 6 × 6 checkerboard response grid. The response grid was constructed to mimic the stimulus space with 18 valid locations, one for each stimulus, and 18 invalid locations that were blacked out. To respond, the observer moved the mouse to one of the 18 valid locations and clicked on the mouse button. The selected location was then filled with a gray square, and the location associated with the correct response was brightened to provide the observer with corrective feedback. Corrective feedback was presented for 500 msec, followed by a 1,000-msec blank screen and initiation of the next trial.

Categorization task. Each observer completed four separate categorization conditions. The first session of each categorization condition was considered as practice and was excluded from further analyses. The stimulus-to-category response mappings for the DSAL, DSAO, LI, and NLI conditions are depicted in Figures 2b, 2c, 2d, and 2e, respectively. The condition orders were as follows: for Observer 1, NLI, LI, DSAO, DSAL; for Observer 2, LI, NLI, DSAO, DSAL; for Observer 3, NLI, LI, DSAL, DSAO; and for Observer 4, LI, NLI, DSAL, DSAO. The trial-by-trial procedure was identical to that for the identification experiment except that the checkerboard response display was removed, and the observers responded by pressing one of two keys on the keyboard. In addition, the checkerboard feedback display was removed, and the correct category "A" or "B" was displayed on the screen.

Results and Theoretical Analyses

Presentation of the results and theoretical analyses is organized as follows. First, the matching data are summarized and are used to derive an initial perceptual representation. Second, the matching task perceptual representation is used along with general recognition theory to predict quantitatively identification performance. Third, the matching task perceptual representation is used along with general recognition theory to predict quantitatively early and later training categorization performance. The aim was to determine whether the matching task perceptual representation provides a good account of both identification and categorization performance. A second aim was to test the hypothesis that decisional selective attention tasks lead to changes in both the perceptual representation and the response regions, whereas decisional integration tasks lead to changes only in the response regions. Finally, identification and categorization are compared directly by deriving an initial perceptual representation from identification and then using the resulting representation to predict categorization performance.

Matching Experiment

A perceptual representation was derived separately for each observer by computing the sample orientation and length means (mean vector entries) and the sample orienta-

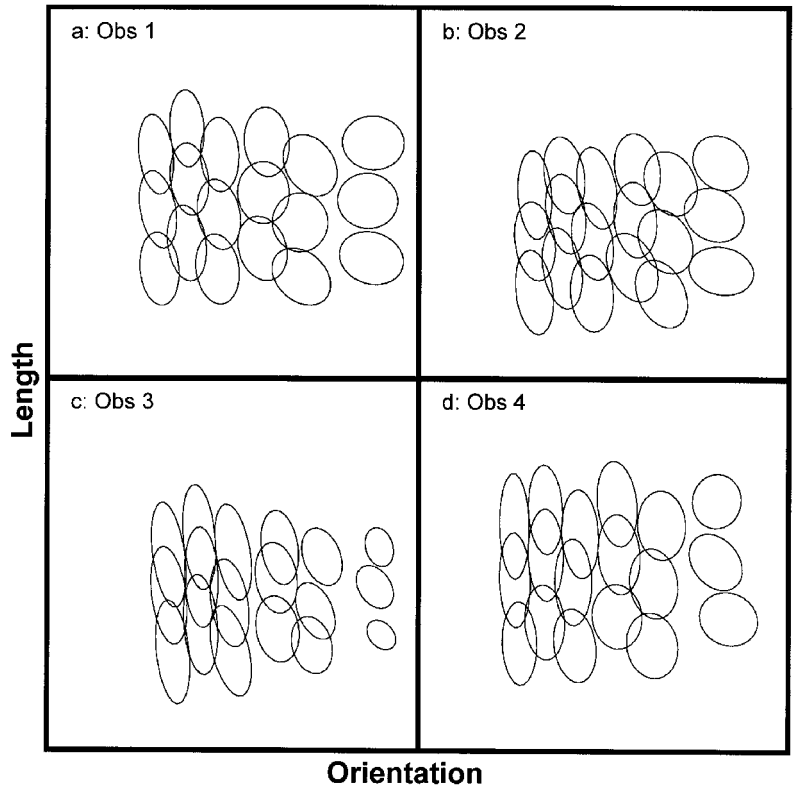

Figure 3. Equal likelihood contours that describe the perceptual representation derived from the matching task for (a) Observer 1, (b) Observer 2, (c) Observer 3, and (d) Observer 4.

tion and length variances and the sample length–orientation covariance (covariance matrix entries). Because two stimuli were presented simultaneously in the matching task, these variance and covariance estimates represent the sum of the standard and comparison stimulus variances and covariances (Alfonso-Reese, 1997; Maddox & Bogdanov, 2000). Assuming statistical independence between the percepts of the standard and comparison stimuli, each entry in the covariance matrix was divided by two.[1] Each mean vector and covariance matrix provided an estimate of the population parameters for each bivariate normal distribution of perceptual effects. Because length and orientation are measured in different units, a method was needed to equate these two measures. To facilitate this goal, all orientation values (measured in radians) were transformed via a linear transformation with slope = 80, and intercept = 109.84.

Figures 3a–3d display the 18 equal likelihood contours (for likelihood = .003) for Observers 1–4, respectively. Several aspects of these plots are of interest. First, the representations are nearly perceptually separable, although some violations exist. Second, the contours are generally perpendicular to the coordinate axis; this result suggests perceptual independence. Third, variability in perceived length showed a general (although not always monotonic) increase as physical length increased. Fourth, the orientation dimension appears more discriminable than the length dimension. As we will see shortly, this resulted in better performance in the DSAO condition than in the DSAL condition. Finally, notice that the resulting perceptual representations are similar across observers. The similarities among the equal likelihood contours and to those from previous research (e.g., Ashby & Lee, 1991; Maddox & Bogdanov, 2000) suggest that the matching task provided a reasonable approximation of the true perceptual distributions.

It is important to note that these analyses provide only a rough estimate of perceptual noise. These analyses ignored the possibility that there was noise in the decision process (termed *criterial noise*). Criterial noise was minimized by (1) using highly trained observers, (2) presenting the stimuli simultaneously, (3) using high-contrast displays, and (4) allowing the observer unlimited time to perform each trial. However, there was probably some criterial noise that is not being accounted for. Even so, one advantage of estimating the perceptual representation from one task (such as the matching task) and then using that representation to predict performance in another task (such as identification or categorization) is that any changes in perceptual and criterial noise across tasks can be estimated. While it would be best if "true" perceptual noise

estimates could be garnered from the matching task, this multiple-task approach should allow one to adequately tease apart these two sources of noise. Of course, the strongest test of the claim that the matching task yields reasonable estimates of the perceptual representation will come from actually applying this representation to the identification and categorization data.

Predicting Identification From Perceptual Matching

Overall identification accuracies were 37.3%, 23.3%, 22.2%, and 34.9% for Observers 1, 2, 3, and 4, respectively. The models were fit to the data using an iterative search routine that minimized the sum of the squared errors between the predicted and observed frequencies for each of the 324 (18 × 18) cells in the confusion matrix.[2] Several of the models have a nested structure, and so it is possible to test whether the extra parameters of a more general models, lead to a significant improvement in fit over the more restricted model. Let SSE_r and SSE_g denote the sum of squared errors for the restricted and more general models, respectively. In addition, let df_r and df_g denote the degrees of freedom for these models. Then, under the null hypothesis that the restricted model is the correct model, the statistic

$$F_{obs} = \frac{(SSE_r - SSE_g)/(df_r - df_g)}{SSE_g/df_g}$$

has an approximately F distribution with $df_r - df_g$ degrees of freedom in the numerator and df_g degrees of freedom in the denominator (e.g., Khuri & Cornell, 1987).[3]

Perceptual representation assumption. The perceptual representation derived from the matching task was assumed to be the correct perceptual representation for identification, with one caveat. Recall that during the perceptual matching task, the stimulus exposure duration was unlimited, whereas during identification, the stimulus exposure duration was 100 msec. Thus, the observer could obtain more perceptual information on each trial during the matching task, resulting in less perceptual variability (Green & Swets, 1967; Luce, 1986; Nosofsky, 1983). In addition, the decision problem is different in the two tasks, which will likely lead to differences in criterial noise. Although perceptual and criterial noise are nonidentifiable in matching and identification, it was important to accommodate changes across these two tasks. Following Maddox and Bogdanov (2000; see also Maddox & Ashby, 1996), the perceptual noise estimates from the matching tasks were scaled by adding a fixed constant to each of the 18 length variances and by adding a fixed constant to each of the 18 orientation variances. These constants were assumed to provide an estimate of the increase in perceptual noise and the change in the magnitude of criterial noise in identification. The value of the two constants was estimated from the identification data, requiring two free parameters. Notice that although the perceptual noise was allowed to be different across matching and identification, the "global" structure of the perceptual representation derived from matching was held fixed in identification. Specifically, the perceptual means and perceptual correlations were unchanged, and the overall relations among perceptual noise values remained constant. This is referred to as the *matching perceptual representation.*

Response region assumptions. Response regions can be constructed in a number of ways (Ashby et al., 2001). For example, the decision bounds could satisfy decisional separability or be linear or quadratic in form. To implement the model, some assumptions need to be made about the nature of the response regions. Ashby et al. offered a neuropsychological theory of this process as applied to identification. The anatomical details and empirical support for the theory are provided in other articles (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Ashby & Waldron, 1999; Ashby et al., 2001). In short, visual stimuli are represented perceptually in higher level visual areas, such as inferotemporal cortex (IT). Since as many as 10,000 cells in visual cortex project to the same cell in the striatum (Wilson, 1995), it is assumed that a low-resolution map of the perceptual space is represented among the striatal units. As the observer gains experience with the task, each unit becomes associated with a particular response. On each trial, the observer determines which unit is closest to the perceptual effect and gives the associated response. Notice that this is equivalent to minimum distance classification (Ashby & Maddox, 1993). This is referred to as the *striatal pattern classifier* (SPC). Hypothetical response regions and units for the four stimuli in Figure 1a are displayed in Figure 1c. Regardless of whether one accepts the neuropsychological underpinnings of the SPC, it has been found to provide a good computational model of the observer's response regions in identification.

Two versions of the SPC were applied to the data. Both assumed that a single striatal unit was associated with each stimulus. The first, SPC(fixed), assumed that the units corresponded to the perceptual means and required no free parameters. The second, SPC(free), assumed that the units were free parameters and required that the location of each be estimated from the data (36 parameters total).

The number of free parameters, *SSE*, and percent of variance accounted for by each model for each observer are presented in Table 1. For now, ignore the rightmost columns headed "ID Perceptual Representation." Recall that the models allowed for differences in perceptual and criterial noise across matching and identification by estimating two additive scalars. An F test was conducted to determine whether the addition of these two parameters provided a significant improvement in fit. In all cases, allowing for differences in perceptual and criterial noise across matching and identification provided a significant improvement in fit [Fs(2,304) = 1,190.73, 421.71, 612.69, and 698.69, for Observers 1, 2, 3, and 4, respectively; ps < .0001]. As predicted, the parameter estimates suggested that there was an increase in perceptual and criterial noise

Table 1
Goodness of Fit (*SSE*) and Percent of Variance (% Var) Accounted for by Several General Recognition Theory Models as Applied to the Identification Confusion Matrices

	Perceptual Representation and Response Region Assumptions							
	Matching				ID			
	SPC(fixed)		SPC(free)		SPC(fixed)		SPC(free)	
Observer	*SSE*	% Var	*SSE*	% Var	*SSE*	% Var	*SSE*	% Var
1	38,853	87.05	9,302	96.74	17,566	93.87	4,413	98.52
2	110,334	47.21	15,289	91.66	41,780	78.63	8,881	95.17
3	42,077	75.14	11,961	91.50	13,943	90.32	4,691	96.66
4	56,776	80.37	10,738	96.13	22,416	92.31	4,952	98.32

Note—SPC, striatal pattern classifier. For matching task perceptual representation, the numbers of parameters were 2 and 38 for SPC(fixed) and SPC(free), respectively. For ID perceptual representation, the numbers of parameters were 88 and 120 for SPC(fixed) and SPC(free), respectively.

from the matching to the identification task. An F test was also conducted to determine whether the additional free parameters of the SPC(free) model provided a significant improvement in fit over the SPC(fixed) model. For all 4 observers, the SPC(free) model provided a significant improvement in fit [Fs(36,268) = 23.65, 46.28, 18.74, and 31.92, for Observers 1, 2, 3, and 4, respectively; ps < .0001]. This finding mirrors that of Ashby et al. (2001). Notice that the *SSE* improvement was quite large, ranging from threefold to sevenfold. In addition, the improvement in percent of variance accounted for was large, ranging from 5% to 45%.

The units and associated response regions from the SPC(free) model for each observer are displayed in Figure 4. Scatterplots of the observed confusion matrix frequencies against the predicted confusion matrix frequencies for the same model for each observer are shown in Figure 5.

Brief summary. The analyses presented in this section represent an attempt to predict quantitatively identification confusions from an initial perceptual representation estimated from the perceptual matching task, along with reasonable assumptions regarding the response regions. A good description of the data was provided by a model that assumed that (1) the "global" structure of the perceptual

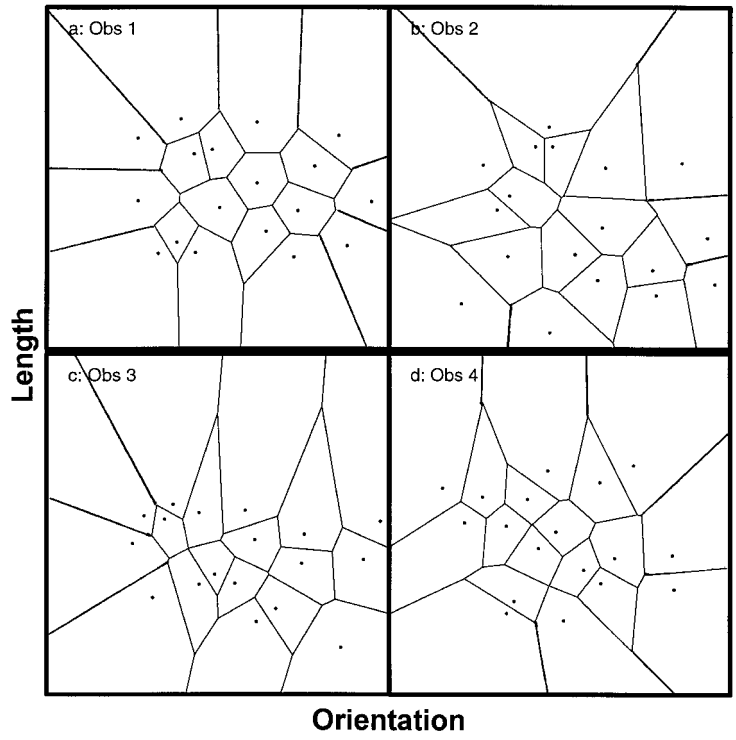

Figure 4. Response regions and units from the SPC(free) model of the decision process derived from a fit of the matching task perceptual representation to the identification data for (a) Observer 1, (b) Observer 2, (c) Observer 3, and (d) Observer 4.

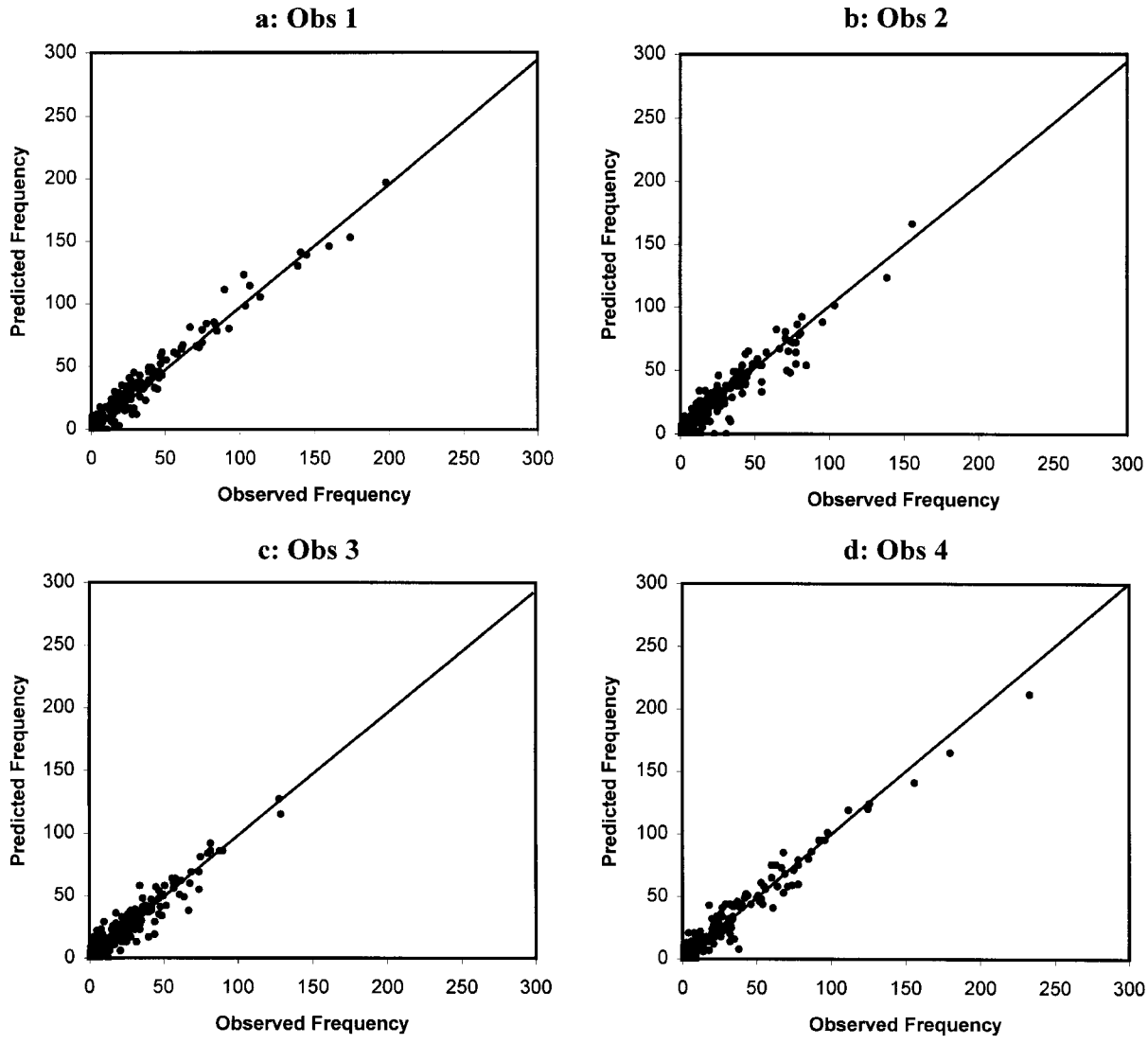

Figure 5. Scatterplot of the observed and predicted identification frequencies from the identification model that assumed the matching task perceptual representation and the SPC(free) response regions for (a) Observer 1, (b) Observer 2, (c) Observer 3, and (d) Observer 4.

representation remained unchanged across tasks, except for a global increase in perceptual noise, and (2) the response regions were characterized by the SPC. This fairly simple model accounted (on average) for 94% of the variance in the data.

Predicting Categorization From Perceptual Matching

Accuracy rates. The accuracy rates for the DSAL, DSAO, LI, and NLI conditions for each observer are presented in Table 2. In an attempt to provide a preliminary examination of categorization learning, the categorization data were divided into two phases: early learning (Sessions 2–8) and late learning (Sessions 9–15). The results can be summarized as follows. First, performance generally improved with learning. Second, performance was poorest in the NLI condition. Finally, DSAO performance was consistently better than DSAL performance. This finding is interesting because it is predicted from the matching task perceptual representation (see Figure 3). Recall that the levels of orientation were more discriminable than the levels of length for all 4 observers.

Model-based analyses. The data to be modeled were the observed response frequencies (for Categories A and B) for each of the 18 stimuli in each of four categorization conditions, resulting in 72 degrees of freedom in the data. The technical difficulties that precluded the use of maximum likelihood parameter estimation in identification are not present in the present categorization conditions. Thus, the model parameters were estimated using maximum

Table 2
Accuracy Rates for Each Categorization Condition

	Categorization Condition			
Training Phase	DSAL	DSAO	LI	NLI
	Observer 1			
Early	71.9	76.9	75.4	68.9
Late	74.7	81.0	76.6	69.6
	Observer 2			
Early	68.9	76.1	72.4	53.7
Late	72.7	73.5	71.0	57.3
	Observer 3			
Early	83.3	91.9	85.3	75.1
Late	85.2	93.0	85.5	76.3
	Observer 4			
Early	81.9	81.4	81.5	70.3
Late	80.2	83.1	80.5	70.3
	Average			
Early	76.5	81.6	78.7	67.0
Late	78.2	82.7	78.4	68.4

Note—DSAL, decisional selective attention to length; DSAO, decisional selective attention to orientation; LI, linear integration; NLI, nonlinear integration.

likelihood procedures (Ashby, 1992b; Wickens, 1982), and the goodness-of-fit statistic was

$$\mathrm{AIC} = 2r - 2\ln L,$$

where r is the number of free parameters and L is the likelihood of the model given the data from the four categorization conditions (Akaike, 1974; Takane & Shibayama, 1992). The AIC statistic penalizes a model for extra free parameters in such a way that the smaller the AIC, the closer a model is to the "true model," regardless of the number of free parameters. Thus, to find the best model among a given set of competitors, one simply computes an AIC value for each model and chooses the model associated with the smallest AIC value.

Perceptual representation assumptions. Two perceptual representation assumptions were tested. The first assumed that the initial perceptual representation estimated from the matching task is the correct perceptual representation for categorization and that this perceptual representation is unaffected by different categorization decision rules—that is, that there is no effect of different categorization decision rules on the perceptual representation. The model does allow for the possibility of greater perceptual noise in categorization by estimating the length and orientation scalars that were described earlier. This is referred to as the *categorization condition invariance* hypothesis because the perceptual representation is assumed to be invariant across categorization conditions (Maddox & Bogdanov, 2000).

The second tested the hypothesis that (1) decisional selective attention conditions alter the perceptual representation by reducing the perceptual variance along the relevant dimension, relative to the perceptual variance along the irrelevant dimension, and (2) decisional integration conditions do not alter the perceptual representation. To instantiate this hypothesis, two additive scalars were applied to the perceptual noise estimates only from the data from the two decisional selective attention conditions. Specifically, one scalar was applied to the relevant dimension (i.e., the length perceptual variance in the DSAL condition and the orientation perceptual variance in the DSAO condition), and the other was applied to the irrelevant dimension (i.e., the orientation perceptual variance in the DSAL condition and the length perceptual variance in the DSAO condition). This is referred to as the *decisional integration/decisional selective attention* hypothesis (Maddox & Bogdanov, 2000). If this hypothesis is correct, then the scalar along the relevant dimension should be smaller in magni-

Table 3
Goodness of Fit (AIC) and Percent of Variance (% Var) Accounted for by Several General Recognition Theory Models as Applied to the Early and Late Training Phases of Categorization Performance

	Matching				ID			
	CI		DI/DSA		CI		DI/DSA	
Training Phase	AIC	% Var	AIC	% Var	AIC	% Var	AIC	% Var
				Observer 1				
Early	1,024	94.57	876	95.74	1,050	94.61	924	95.63
Late	1,432	91.56	1,026	94.74	1,354	92.61	1,006	94.90
				Observer 2				
Early	676	96.82	636	97.53	694	96.37	648	97.19
Late	648	97.48	612	97.84	740	95.59	724	95.90
				Observer 3				
Early	760	98.70	638	99.00	1,074	97.24	1,008	97.52
Late	594	99.34	548	99.52	1,072	97.66	1,060	97.73
				Observer 4				
Early	740	98.25	690	98.52	926	97.06	854	97.41
Late	712	98.43	686	98.52	954	96.82	886	97.20

Note—CI, categorization condition invariance; DI/DSA, decisional integration/decisional selective attention. The numbers of parameters were 19 and 21 for CI and DI/DSA, respectively.

Table 4
Perceptual Noise Scalar Values Along the Relevant and Irrelevant Dimension for the Decisional Integration/Decisional Selective Attention Model Assuming the Matching Task Perceptual Representation

	Dimension	
Training Phase	Relevant	Irrelevant
	Observer 1	
Early	12.91	18.01
Late	5.50	33.48
	Observer 2	
Early	15.46	18.69
Late	7.77	52.17
	Observer 3	
Early	4.99	22.44
Late	22.44	48.93
	Observer 4	
Early	2.16	26.83
Late	8.06	22.45

tude than the scalar along the irrelevant dimension. Because these scalars are assumed to measure perceptual processes, it is imperative that one obtain a separate, and unique, estimate of criterial noise. A series of decision bound models derived from general recognition theory has been developed that allows one to estimate separately the effects of perceptual and criterial noise (for details, see Ashby, 1992a, and Maddox & Ashby, 1993).

Response region assumptions. Decision bound models assume that there will be trial-by-trial fluctuations in the memory for the location of the decision bounds, and, in many cases, this criterial noise is identifiable. Although the decision bound could take any mathematical form, most applications of the model have assumed either linear or quadratic decision bounds. In many cases, the response regions predicted by decision bound models are identical or are very similar to the response regions predicted by the SPC. For example, the SPC with one unit assigned to each category is mathematically equivalent to a linear decision bound model (Ashby & Waldron, 1999; Ashby et al., 2001). In addition, with only a few units assigned to each category, the SPC can approximate a wide range of quadratic decision bounds.

Two sets of response region assumptions were examined. The first assumed that linear decision bounds were used in the two decisional selective attention conditions (DSAL and DSAO) and that quadratic decision bounds were used in the two dimensional integration conditions (LI and NLI). Each linear bound required that a slope and intercept be estimated, and each quadratic decision bound required that five quadratic equation coefficients be estimated for a total of 14 decision bound parameters. In addition, three criterial noise parameters were estimated: one for the NLI condition, one for the LI condition, and a single criterial noise parameter was estimated simultaneously for both DSAL and DSAO conditions. The criterial noise parameter had to be "yoked" across the two DSA conditions to make it identifiable from perceptual noise. The second assumed that quadratic decision bounds were used in all four conditions. This did not lead to a significant improvement in fit and will not be discussed further.

The number of free parameters, AIC, and percent of variance accounted for by the categorization condition invariance and decisional integration/decisional selective attention models are presented in the two leftmost columns of Table 3. The decisional integration/decisional selective attention model was superior to the categorization condition invariance model for all 4 observers during early and late learning. This replicates the finding of Maddox and Bogdanov (2000) but extends it in two ways. First, Maddox and Bogdanov examined a decisional selective attention to length condition, whereas the present study examined a decisional selective attention to length and decisional selective attention to orientation conditions. Second, the present study provided a preliminary investigation of the decisional selective attention effect on the perceptual representation across early and late training, which will be discussed shortly. Importantly, the fit of the model was quite good, with the percent of variance accounted for ranging from 95% to 99% during both early and late learning. Thus, it appears that the matching task perceptual representation provides a good method for deriving an initial perceptual representation for both identification and categorization.

Figures 6 and 7 display scatterplots of the observed and predicted categorization probabilities from the decisional integration/decisional selective attention model for each stimulus by condition and observer for early and late training, respectively.

The decisional integration/decisional selective attention model assumed that decisional selective attention reduces the perceptual variance along the decisionally relevant dimension, relative to the decisionally irrelevant dimension. Table 4 displays the perceptual noise scalar values for the relevant and irrelevant dimensions by training phase (early vs. late) for each observer. In support of the hypothesis, the scalar along the relevant dimension was always smaller than the scalar along the irrelevant dimension, except for Observer 3 during early learning. One might reasonably postulate that the magnitude of this effect would increase with training—that is, the difference between the relevant and irrelevant scalar values might increase with training. This pattern held for Observers 1, 2, and 3, but it did not hold for Observer 4.

The model assumed also that decisional selective attention categorization problems are solved by applying linear decision bounds. Given the need for decisional selective attention, the linear decision bounds should be nearly orthogonal to the relevant dimension axis. Table 5 displays the slope of the linear bound for the DSAL and DSAO conditions separately for each training phase by observer. In the DSAL condition, a zero slope implies perfect decisional selective attention. As the absolute value of the slope approaches zero, decisional selective attention improves. In the DSAO condition, an infinite slope implies perfect decisional selective attention. As the absolute

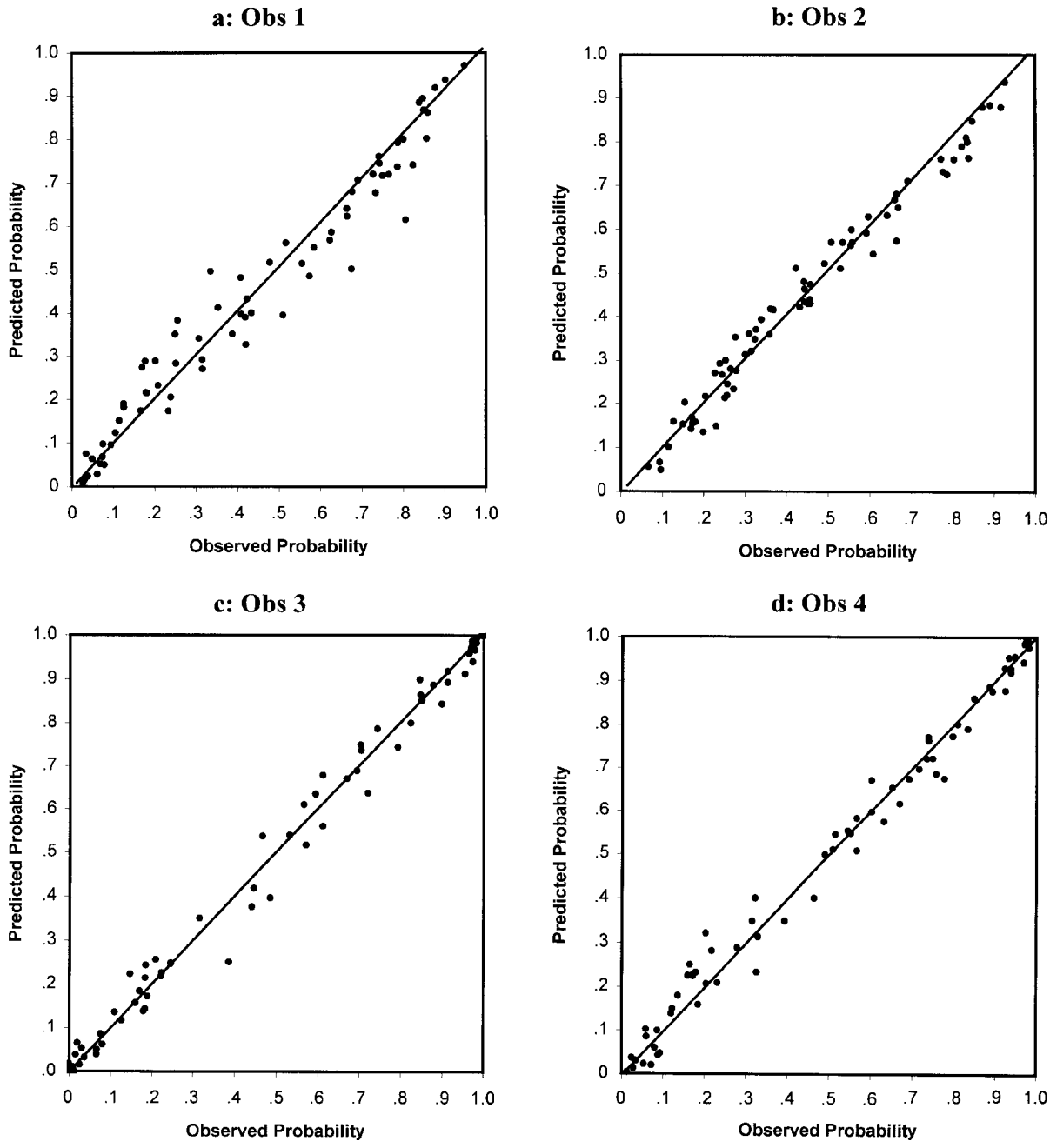

Figure 6. Scatterplot of the observed and predicted early training categorization probabilities from the decisional integration/decisional selective attention model assuming the matching task perceptual representation for (a) Observer 1, (b) Observer 2, (c) Observer 3, and (d) Observer 4.

value of the slope approaches infinity, decisional selective attention improves. A slope of 1 implies equal decisional attention to both dimensions. Two results stand out. First, in every case, more decisional selective attention is being paid to the relevant dimension. In other words, the absolute value of the DSAL slopes are all less than 1, and the absolute value of the DSAO slopes are all greater than 1. Second, there was general support for the hypothesis that the magnitude of decisional selective attention improved with training. Observers 1, 2, and 3 showed better decisional selective attention late in training for both DSAL and DSAO categorization conditions. Observer 4 showed the same pattern, but only for the DSAO condition. The learning effects are clearly preliminary and need to be fol-

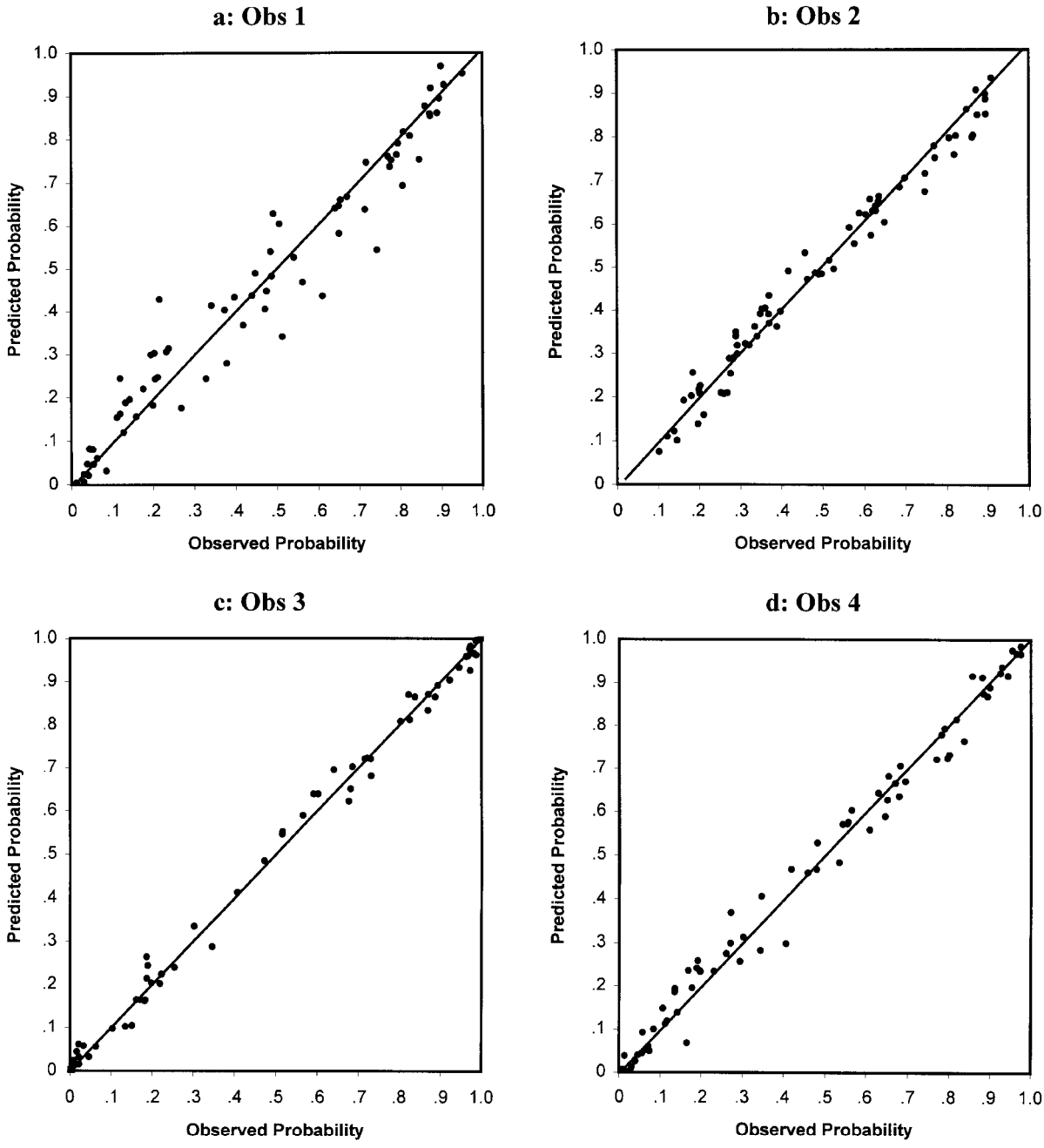

Figure 7. Scatterplot of the observed and predicted late training categorization probabilities from the decisional integration/decisional selective attention model assuming the matching task perceptual representation for (a) Observer 1, (b) Observer 2, (c) Observer 3, and (d) Observer 4.

lowed up with future studies. Even so, they are interesting and suggest that experience has consistent and quantifiable effects on perceptual and decisional processes.

Brief summary. The analyses presented in this section represent an attempt to predict quantitatively the data from a number of categorization conditions from an initial perceptual representation derived from perceptual matching data. The analyses suggested that decisional selective attention categorization problems reduce the perceptual variance along the relevant dimension, relative to the perceptual variance along the irrelevant dimension. This model provided a good account of the data, accounting for 97% of the variance in the data on average. In addition, the model suggested that experience had some interesting ef-

Table 5
Slope of the Linear Decision Bound From the Decisional Integration/Decisional Selective Attention Model Assuming the Matching Task Perceptual Representation for the DSAL and DSAO Conditions

	Condition	
Training Phase	DSAL	DSAO
	Observer 1	
Early	−0.51	−1.90
Late	−0.35	−2.71
	Observer 2	
Early	−0.38	−2.11
Late	−0.24	−2.92
	Observer 3	
Early	−0.16	−6.97
Late	−0.12	−8.74
	Observer 4	
Early	−0.21	−2.19
Late	−0.23	−2.86

Note—DSAL, decisional selective attention to length; DSAO, decisional selective attention to orientation.

fects on perceptual and decisional processes. In particular, as the observer gained experience with the task, the magnitude of the perceptual variance reduction and decisional selective attention generally increased.

Predicting Categorization From Identification

The previous two sections outlined attempts to predict identification and categorization from an initial perceptual representation estimated from the matching task. As an additional test of the validity of the matching task as a method for estimating the perceptual representation and to more fully examine the identification–categorization relationship, in this section an attempt is made to predict quantitatively categorization performance from a perceptual representation estimated from identification.

Modeling identification. We begin by deriving an initial perceptual representation from identification. Because the goal was to derive the best perceptual representation, we fit a version of general recognition theory that allowed all of the perceptual distribution parameters to be freely estimated. This required 36 perceptual distribution mean parameters, 36 perceptual variance parameters, and 18 perceptual covariance parameters. This is referred to as the *identification perceptual representation.* The SPC(fixed) and SPC(free) decision process models were examined. The number of parameters, *SSE*, and percent of variance accounted for are displayed in the two rightmost columns of Table 1. For all 4 observers, the SPC(free) provided a significant improvement in fit over the SPC(fixed) [$F(32,186)$= 17.32, 21.53, 11.46, and 20.50, for Observers 1, 2, 3, and 4, respectively; ps < .00001]. Notice that the *SSE* improvement was quite large, ranging from threefold to fivefold. In addition, the improvement in percent of variance accounted for was large, ranging from 5% to 17%. These values are in line with those for the SPC(free) and SPC(fixed) comparison when the matching perceptual representation was assumed.

Because the SPC(free) that assumed the matching task perceptual representation is nested under the SPC(free) that assumed the identification perceptual representation, F tests can be performed. For all 4 observers, estimating the perceptual representation from the identification data provided a significant improvement in fit over the case in which the matching task perceptual representation was used to predict identification performance [Fs(82,186) = 2.51, 1.64, 3.52, and 2.65, for Observers 1, 2, 3, and 4, respectively; ps < .01]. Although statistically significant, it is worth noting that the improvement in fit was relatively small (approximately a twofold *SSE* improvement and a 3% improvement in percent of variance accounted for). Perhaps more importantly, this small improvement in fit was obtained by adding 86 parameters, whereas the much larger improvement in fit for the SPC(free) over the SPC(fixed) required only 36 additional response region parameters. In addition, the absolute deviations between the perceptual representation parameters from the ID model and those for the matching model were small. Specifically, averaged across the 4 observers, the absolute deviations between the orientation means and standard deviations, measured in radians, were .036 and .031, respectively. The average absolute deviations between the length means and standard deviations, measured in pixels, were 1.963 and 1.604, respectively. Finally, the average absolute deviations between the correlations were .044. Taken together, these analyses suggest that the matching task does provide a good initial perceptual representation for identification, nearly as good as the perceptual representation freely estimated from identification confusions.[4]

Predicting categorization from the identification perceptual representation. In this section, we attempt to predict early and late categorization performance in each of the four categorization conditions from the identification perceptual representation. Both the categorization condition invariance hypothesis and the decisional integration/decisional selective attention hypothesis were examined. In addition, both hypotheses regarding the response regions were examined (recall that one assumed linear decision bounds in the DSA conditions, whereas the other assumed quadratic decision bounds). In line with the analyses presented earlier, quadratic decision bounds in the DSA conditions did not provide a significant improvement in fit.

The number of free parameters, AIC, and percent of variance accounted for by the categorization condition invariance and decisional integration/decisional selective attention models for each observer are presented in the two rightmost columns of Table 3. Several results stand out. First, the decisional integration/decisional selective attention models provided a better account of the data than the categorization condition invariance model for all 4 observers during early and late learning. This replicates the finding of Maddox and Bogdanov (2000) and the previous

analyses and extends this finding to the identification–categorization relationship. The perceptual noise scalar estimates and linear decision bound slopes from this model were very similar and showed a similar pattern across learning to those outlined in Tables 4 and 5 that were estimated assuming the matching task perceptual representation. Second, the matching task perceptual representation generally provided a better account of categorization performance than did the identification task perceptual representation. Specifically, for the decisional integration/decisional selective attention model, the matching task perceptual representation yielded a better account of the data for all 4 observers during early and late training, with one exception (Observer 1, late training). Even so, both perceptual representations provided good accounts of the data accounting (on average) for 97.68% and 96.69% of the variance in the categorization data for the matching and identification perceptual representations, respectively. Thus, both tasks provide good methods for deriving an initial perceptual representation, although there appears to be a slight advantage for the matching task.

DISCUSSION

The present study applied general recognition theory to account simultaneously for perceptual matching, identification, and categorization performance. General recognition theory assumes that perceptual processing is noisy and probabilistic and is defined by parameters that are separately identifiable from those associated with decision processes. Two specific hypotheses were tested. First, the (often implicit) assumption that changing task demands affect decision processes while leaving the perceptual representation unchanged was examined. Second, the hypothesis that decisional selective attention tasks alter both decision processes and perceptual processes by reducing the perceptual variance along the attended dimension whereas decisional integration tasks alter only decision processes was tested.

A perceptual representation for 18 line stimuli was derived from the perceptual matching data. This representation was then held fixed in an attempt to model data from the identification and categorization tasks. The matching perceptual representation was compared with a perceptual representation that was freely estimated from the identification data in their ability to account for the identification and categorization data. Not surprisingly, the identification perceptual representation was superior to the matching perceptual representation in accounting for the identification confusions. On average, the identification perceptual representation accounted for approximately 97% of variance in the data, as compared with the matching task perceptual representation that accounted for approximately 94% of the variance in the data. However, this relatively small improvement in fit for the identification perceptual representation came about at the expense of 86 additional free parameters. In addition, the parameters of the resulting perceptual representations were quite similar. Clearly, the matching task perceptual representation is providing a good account of the data. More remarkable, though, is the excellent fit of the matching task perceptual representation to the categorization data. In fact, in seven of eight cases (4 observers $\times$ 2 training phases), the matching task perceptual representation provided a better account of the categorization data than did the identification perceptual representation. This is likely the strongest evidence to date that the perceptual representation remains relatively unchanged across these different, but related, tasks.

The hypothesis that decisional selective attention tasks alter both perceptual and decisional processes whereas decisional integration tasks affect only decisional processes found support. Specifically, decisional selective attention problems led to the use of decision bounds that placed more weight in the decision on the relevant dimension and decreased the perceptual variance along the attended dimension, relative to the unattended dimension. The magnitude of both effects was found to increase as the observer gained experience with the tasks. This is an important finding because it suggests that some decision problems—in particular, those that require selective attention—can alter both the decision strategy applied and the nature of the underlying perceptual representation. This also adds to the growing body of empirical support for the notion that attention can operate at both perceptual and decisional levels (Johnston, McCann, & Remington, 1995; Pashler, 1989, 1991, 1993; Posner, 1993; Posner, Sandson, Dhawan, & Shulman, 1989) and extends these results to categorization (see also Maddox, Ashby, & Waldron, 2000).

These results converge with current thinking in the cognitive neurosciences, and a recently proposed neuropsychological theory of categorization. Ashby, Alfonso-Reese, et al. (1998) proposed a neuropsychological theory of categorization that postulates two separate and competing categorization systems: one dominated by explicit categorization rules, and the other dominated by implicit categorization rules. The implicit system bases its decision on a perceptual representation computed in some visual area (IT or lower), whereas the verbal system presumably uses a representation from object-based working memory (thought to be in ventrolateral prefrontal cortex; Baddeley, 1995). Current theories of visual selective attention postulate separate attentional effects in visual areas, in IT, and in ventrolateral prefrontal cortex (e.g., LaBerge, 1995). Thus, the perceptual representation used by the explicit system receives different (but overlapping) attentional processing than the representation used by the implicit system. In the present study, it is reasonable to suppose that the decisional selective attention conditions are each dominated by an explicit rule (e.g., in the DSAL condition, the observer should respond "A" if the length is small and respond "B" if the length is long), whereas the linear and nonlinear integration conditions are dominated by the implicit system (since no explicit rule exists that could adequately solve these categorization problems). Because

the perceptual representations for the explicit and implicit systems receive different (but overlapping) attention processing, it is possible that the representations would be affected differently. Clearly, this reasoning is tentative and requires more rigorous empirical testing. Even so, these data and the present approach provide an interesting crossroads between traditional studies of categorization and the growing interest in cognitive neuroscience.

Future Directions

Despite the wealth of information gained from the present empirical and theoretical investigation, this study represents only a first step along a potentially fruitful path of inquiry. An obvious next step would be to replicate the present study with highly integral-dimension stimuli, such as color patches that vary in saturation and brightness. A large body of research suggests that the efficiency of attention is different for integral- and separable-dimension stimuli (e.g., Garner, 1974). In particular, it is argued that selective attention is easy with separable-dimension stimuli but is difficult with integral-dimension stimuli (Garner, 1974; McKinley & Nosofsky, 1996; Nosofsky, 1987; however, see Ashby & Maddox, 1994, and Maddox, 1992). It is possible that the attention effects observed in the present study might not obtain (or might operate less efficiently) with integral-dimension stimuli. Another, perhaps more intriguing, possibility is that decisional processes will be relatively unchanged across separable- and integral-dimension stimuli but that the effects of decisional selective attention on the nature of the perceptual representation will be absent with integral-dimension stimuli. Another interesting line of future research would be to investigate learning effect more fully. The results for the early and late phases of categorization learning are suggestive, but future research should examine the time course of learning across more than two learning phases. In addition, an examination of possible perceptual and decisional learning effects in identification could also be fruitful. Finally, applications to other complex tasks are conceivable. For example, the formal model of feature binding in object perception developed by Ashby, Prinzmetal, Ivry, and Maddox (1998; see also Maddox, Prinzmetal, Ivry, & Ashby, 1994) could be generalized by relaxing the all-or-none feature perception and feature sampling independence assumptions (see Ashby, Prinzmetal, et al., 1998, for details). In addition, perceptual matching and/or identification data could be collected to provide an initial estimate of the perceptual representation. Several other applications are possible. For example, Geisler and Chou (1995) took a similar approach to the study of the visual search task.

In conclusion, the present study applied a unified theoretical framework—general recognition theory—to the relations among perceptual matching, identification, and categorization. General recognition theory separates perceptual from decisional processes. Excellent accounts of the identification and decisional integration categorization data were obtained by assuming changes in the response regions across tasks but a fixed perceptual representation derived from perceptual matching. Decisional selective attention categorization tasks led to the use of a decision strategy that placed more weight in the decision on the relevant dimension and decreased the perceptual variance along the attended dimension, relative to the unattended dimension. Both of these effects generally increased in magnitude with learning.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.

Alfonso-Reese, L. A. (1996). *Dynamics of category learning.* Doctoral dissertation, University of California, Santa Barbara.

Alfonso-Reese, L. A. (1997). *On the dangers of ignoring noise in high-level perception experiments* (Technical report). Bloomington: Indiana University.

Ashby, F. G. (1988). Estimating the parameters of multidimensional signal detection theory from simultaneous ratings on separate stimulus components. *Perception & Psychophysics*, **44**, 195-204.

Ashby, F. G. (1992a). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449-484). Hillsdale, NJ: Erlbaum.

Ashby, F. G. (1992b). Multivariate probability distributions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 1-34). Hillsdale, NJ: Erlbaum.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, **105**, 442-481.

Ashby, F. G., & Lee, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General,* **120**, 150-172.

Ashby, F. G., & Lee, W. W. (1992). On the relationship among identification, similarity, and categorization: Reply to Nosofsky and Smith (1992). *Journal of Experimental Psychology: General*, **121**, 385-393.

Ashby, F. G., & Lee, W. W. (1993). Perceptual variability as a fundamental axiom of perceptual science. In S.C. Masin (Ed.), *Foundations of perceptual theory* (pp. 369-399). Amsterdam: North-Holland.

Ashby, F. G., & Maddox, W. T. (1990). Integrating information from separable psychological dimensions. *Journal of Experimental Psychology: Human Perception & Performance,* **16**, 598-612.

Ashby, F. G., & Maddox, W. T. (1991). A response time theory of perceptual independence. In J. P. Doignon & J. C. Falmagne (Eds.), *Mathematical psychology: Current developments* (pp. 389-414). New York: Springer-Verlag.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology,* **37**, 372-400.

Ashby, F. G., & Maddox, W. T. (1994). A response time theory of perceptual separability and perceptual integrality in speeded classification. *Journal of Mathematical Psychology,* **33**, 423-466.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science*, **5**, 144-150.

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review,* **95**, 124-150.

Ashby, F. G., Prinzmetal, W., Ivry, R. B., & Maddox, W. T. (1998). A formal theory of feature binding in object perception. *Psychological Review*, **103**, 165-192.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review,* **93**, 154-179.

Ashby, F. G., & Waldron, E. M. (1999). On the nature of implicit categorization. *Psychonomic Bulletin & Review*, **6**, 363-378.

Ashby, F. G., Waldron, E. M., Lee, W. W., & Berkman, A. (2001). Suboptimality in categorization and identification, *Journal of Experimental Psychology: General*, **130**, 77-96.

Baddeley, A. (1995). Working memory. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 755-764). Cambridge, MA: MIT Press.

Bonnel, A.-M., & Hafter, E. R. (1998). Divided attention between simultaneous auditory and visual signals. *Perception & Psychophysics*, **60**, 179-190.

Braida, L. D., & Durlach, N. I. (1972). Intensity perception: II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, **51**, 483-502.

Brainard, D. H. (1997). Psychophysics software for use with MATLAB. *Spatial Vision*, **10**, 433-436.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, **53**, 134-140.

Garner, W. R. (1974). *The processing of information and structure*. New York: Wiley.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, **1**, 225-241.

Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review,* **96**, 267-341.

Geisler, W. S., & Chou, K. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review*, **102**, 356-378.

Green, D. M., & Swets, J. A. (1967). *Signal detection theory and psychophysics*. New York: Wiley.

Johnston, J. C., McCann, R. S., & Remington, R. W. (1995). Chronometric evidence for two types of attention. *Psychological Science*, **6**, 365-369.

Khuri, A. I., & Cornell, J. A. (1987). *Response surfaces: Designs and analyses*. New York: Marcel Dekker.

LaBerge, D. (1995). Computational and anatomical models of selective attention in object identification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 649-664). Cambridge, MA: MIT Press.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.

Luce, R. D., & Green, D. M. (1978). Two tests of a neural attention hypothesis for auditory psychophysics. *Perception & Psychophysics*, **23**, 363-371.

Macmillan, N. A., Goldberg, R. F., & Braida, L. D. (1988). Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America*, **84**, 1262-1280.

Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147-180). Hillsdale, NJ: Erlbaum.

Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & Psychophysics*, **61**, 354-374.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics,* **53**, 49-70.

Maddox, W. T., & Ashby, F. G. (1996). Perceptual separability, decisional separability, and the identification–speeded classification relationship. *Journal of Experimental Psychology: Human Perception & Performance,* **22**, 795-817.

Maddox, W. T., & Ashby, F. G. (1998). Selective attention and the formation of linear decision boundaries: Comment on McKinley and Nosofsky (1996). *Journal of Experimental Psychology: Human Perception & Performance,* **24**, 301-321.

Maddox, W. T., Ashby, F. G., & Waldron, E. M. (2000). *Multiple attention systems in perceptual categorization.* Manuscript submitted for publication.

Maddox, W. T., & Bogdanov, S. V. (2000). On the relation between decision rules and perceptual representation in multidimensional perceptual categorization. *Perception & Psychophysics*, **62**, 984-997.

Maddox, W. T., Prinzmetal, W., Ivry, R. B., & Ashby, F. G. (1994). A probabilistic multidimensional model of location discrimination. *Psychological Research*, **56**, 66-77.

McKinley, S. C., & Nosofsky, R. M. (1996). Selective attention and the formation of linear decision boundaries. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 294-317.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review*, **4**, 79-95.

Nosofsky, R. M. (1983). Information integration and the identification of stimulus noise and criterial noise in absolute judgment. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 299-309.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, **115**, 39-57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* **13**, 87-109.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics*, **45**, 279-290.

Pashler, H. (1989). Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology,* **21**, 469-514.

Pashler, H. (1991). Shifting visual attention and selecting motor responses: Distinct attentional mechanisms. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 1023-1040.

Pashler, H. (1993). Dual-task interference and elementary mental mechanisms. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 245-263). Cambridge, MA: MIT Press.

Perrin, N. A., & Ashby, F. G. (1991). A test of perceptual independence with dissimilarity data. *Applied Psychological Measurement*, **15**, 79-93.

Posner, M. I. (1993). Attention before and during the decade of the brain. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 342-350). Cambridge, MA: MIT Press.

Posner, M. I., Sandson, J., Dhawan, M., & Shulman, G. L. (1989). Is word recognition automatic? A cognitive-anatomical approach. *Journal of Cognitive Neuroscience*, **1**, 50-60.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, **22**, 325-345.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, **1**, 54-87.

Shepard, R. N., & Chang, J. J. (1963). Stimulus generalization in the learning of classification. *Journal of Experimental Psychology*, **65**, 94-102.

Shepard, R. N., Hovland, C. I., & Jenkins, J. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, **75**(13, Whole No. 517).

Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1411-1436.

Takane, Y., & Shibayama, T. (1992). Structures in stimulus identification data. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 335-362). Hillsdale, NJ: Erlbaum.

Wickens, T. D. (1982). *Models for behavior: Stochastic processes in psychology*. San Francisco: Freeman.

Wilson, C. J. (1995). The contribution of cortical neurons to the firing pattern of striatal spiny neurons. In J. C. Houk (Ed.), *Models of information processing in the basal ganglia* (pp. 29-50). Cambridge, MA: MIT Press.

NOTES

1. The independence assumption is likely incorrect, but, in the present situation, it should lead to a constant error across stimuli. Since the model-based analyses of the identification and categorization data will allow for an additive constant to be applied to the perceptual variances, this "error" should not adversely affect our ability to interpret the data.

2. The preferred method of parameter estimation is maximum likelihood. Unfortunately, for technical reasons, this was not possible in the present study (or in other applications of general recognition theory to identification confusions; Ashby & Lee, 1991; Ashby et al., 2001; Maddox & Ashby, 1996). To estimate the probability of responding j when stimulus i is presented, one must integrate the stimulus i perceptual dis-

tribution over the response region assigned to *j*. Theoretically, this probability can never be zero with normal distributions, but, when fitting the models, numerical integration is required. When the response *j* region is far from the stimulus *i* perceptual distribution, these numerical integration procedures will yield a predicted response probability equal to zero. Predicted probabilities equal to zero yield infinite fit values from the method of maximum likelihood. This problem could be alleviated by replacing predicted response probabilities of zero with some small value such as .001 or .0001, but the specific value selected would change drastically the fit value. Instead of selecting some arbitrary value, the least squares approach taken by previous researchers was utilized (Ashby & Lee, 1991; Ashby et al., 2001; Maddox & Ashby, 1996). In addition, a major problem in model fitting is to avoid "local minima." Local minima result when the parameter estimation algorithm identifies a set of parameters that are "locally" superior (i.e., provide the best fit) but are not "globally" superior. Although one can never be certain that the global minima have been obtained, procedures should be followed to minimize this possibility. Along these lines, multiple starting parameter values were utilized. In addition, since many of the models were nested, the parameters from a more restricted model were used as starting values for a more general model. These procedures likely improve one's chances of identifying the global minima.

3. The results of this test must be interpreted with caution, since it assumed that the confusion matrix frequencies are independent, which is likely not the case. Even so, Ashby and Lee (1991, 1992) showed that these *F* tests generally agree with tests based on a maximum likelihood approach that does not assume independence.

4. One might argue that it is no surprise that models of this complexity with so many free parameters provide good accounts of the identification data. Unfortunately, model complexity is a difficult issue that has not been adequately defined (however, see Myung & Pitt, 1997). However, two points are worth mentioning. First, these models are all falsifiable—that is, they have fewer free parameters than degrees of freedom in the data, which is not a property of many popular models (e.g., ANOVA). Second, these models have fewer parameters than the benchmark model of identification—namely, the Luce–Shepard similarity (biased) choice model (Luce, 1963; Shepard, 1957).

(Manuscript received April 10, 2000;
revision accepted for publication January 24, 2001.)