Same–Different Texture Discrimination in Pigeons: Testing Competing Models of Discrimination and Stimulus Integration

Robert G. Cook
Tufts University

John T. Wixted
University of California, San Diego

The choice behavior of 6 pigeons performing a multidimensional same–different texture discrimination was examined. On each trial, they had to choose among 2 choice hoppers depending on whether a color, shape, or redundant (color and shape) target signal was present or not in a textured stimulus. Receiver operating characteristic (ROC) curves were produced by variations in the a priori signal presentation probabilities across conditions. Quantitative analyses of these ROC curves were used to evaluate different competing theories of discrimination (signal detection vs. high-threshold–default response models) and information integration (independent observations, additive integration, unidimensional models). The results suggested the structure of the pigeons' choice behavior in this same–different discrimination was best described by an unequal variance signal detection model involving a unidimensional evidence variable (e.g., degree of difference).

As birds forage through the forest canopy or on the ground, they encounter a wide variety of visual cues about the location and types of food that might be available. How birds, and animals in general, internally represent such information to guide behavior is an important issue in animal cognition. This article focuses on the discrimination and integration of simple dimensional information by pigeons. In a same–different discrimination task (Cook, Cavoto, & Cavoto, 1995), the pigeons in our experiment were required to respond *same* whenever all of the elements of a multielement textured stimulus were identical in form and color, and respond *different* whenever a small contrasting *target* region was present. These targets could differ from the surrounding *distractor* region in either color, shape, or a compound of both of these dimensions. Our interest in this discrimination was derived in part from recent transfer evidence suggesting that pigeons solve this task by using a generalized same–different rule (Cook, Cavoto, et al., 1995; Wasserman, Hugart, & Kirkpatrick-Steger, 1995) and from the potential asymmetries associated with this form of discrimination due to, for example, the inherently greater number of possible Different stimuli than Same stimuli. The

present research was designed to illuminate in more detail the structure of how pigeons perform this complex stimulus discrimination.

Several issues that we have separately addressed in the past figure prominently in the work described next, with each playing a key role in our attempt to better understand the larger question of how pigeons perform the present same–different discrimination. The first issue concerns how to best conceptualize the nature of the decision strategy used by animals in making choice discriminations, especially those using psychologically asymmetrical stimuli (Dougherty & Wixted, 1996; Wixted, 1993). The main question of interest has been whether the pigeons' decision strategy is more accurately characterized by a signal detection or a default response–asymmetrical coding model. The second issue concerns how to best conceptualize the mechanisms underlying the processing by animals of dimensional information presented in elemental and compound stimulus configurations (Cook, 1992; Cook, Cavoto, & Cavoto, 1996; Cook, Riley, & Brown, 1992; Riley, Cook, & Lamb, 1981). In the current experiment, we were interested in how pigeons discriminate and integrate elemental (color or shape) and compound (color and shape) target information during the performance of a multidimensional same–different task. As described next, these questions about the nature of the pigeons' decision strategy and their rules for the discrimination and integration of dimensional information are more closely related than they might seem at first glance.

Decision Strategies in Avian Choice Behavior

Presence-Versus-Absence Discriminations

With respect to the pigeons' decision strategy, our starting point was based on prior work examining delayed presence-versus-absence discriminations in pigeons. Consider, for example, pigeons' delayed discrimination of food versus

no-food samples (e.g., Grant, 1991). In this procedure, a sample consists of the brief presentation of food or the brief presentation of an empty food hopper. Following a delay, red and green choice stimuli are presented. A response to one choice stimulus is reinforced on food trials (with more food), whereas a response to the other choice stimulus is reinforced on no-food trials. In most cases, choice performance declines as a function of delay interval following food samples, but declines little, if at all, following no-food samples (Colwill, 1984; Colwill & Dickinson, 1980; Grant, 1991; Sherburne & Zentall, 1993a, 1993b; Wilson & Boakes, 1985). Exactly the same pattern of results occurs when the procedure involves a delayed discrimination between the presence-versus-absence of a keylight instead of a food sample (Grant, 1991; Dougherty & Wixted, 1996; Wixted, 1993).

One popular account of this asymmetrical forgetting in presence-versus-absence discriminations is the default response hypothesis. According to this account, the presentation of a food sample produces a memory code of that event, but no-food samples are basically treated as nonevents (and, as such, produce no memory code). Whenever memory for a food sample exists, the pigeon correctly responds to the appropriate choice alternative. In the absence of a memory code, the other alternative is selected by default. Note that a memory code is absent on all no-food trials (leading to correct choices) and on long-delay food trials in which the sample has been forgotten (leading to incorrect choices). Thus, performance on food sample trials declines with an increasing retention interval as the memory trace fades away, but performance on no-food sample trials remains constant because memory is not involved. Instead, the no-food choice alternative is selected by default on no-food trials regardless of the size of the retention interval.

Wixted (1993) pointed out the formal equivalence of the default response–asymmetrical coding account to an earlier theory of human psychophysical discrimination known as high-threshold theory (HTT; Blackwell, 1963; Swets, Tanner, & Birdsall, 1961; see Link, 1992, for a review). HTT proposes that when detecting any dimensional signal from a background of noise, subjects respond yes (signal) if the strength of the signal exceeds a fixed threshold or criterion. On all other trials (noise alone trials and subthreshold signal + noise trials), HTT proposes that observers make a biased response among the yes and no alternatives. As such, in terms of this model a bird's default response is a manifestation of a strong bias to choose the no-food alternative when no memory for a food sample is present. The diagram in Figure 1 illustrates the different choice pathways of this particular theory.

The major competitor to HTT models of performance has been signal detection theory (SDT; Peterson, Birdsall, & Fox, 1954; Tanner & Swets, 1954). According to SDT, the subject's decision is guided by information derived from the stimulus and the relative placement of a response or decision criterion. In conjunction, these two factors determine whether the subject will respond yes or no on signal + noise and noise trials. The diagram in Figure 2 illustrates the different
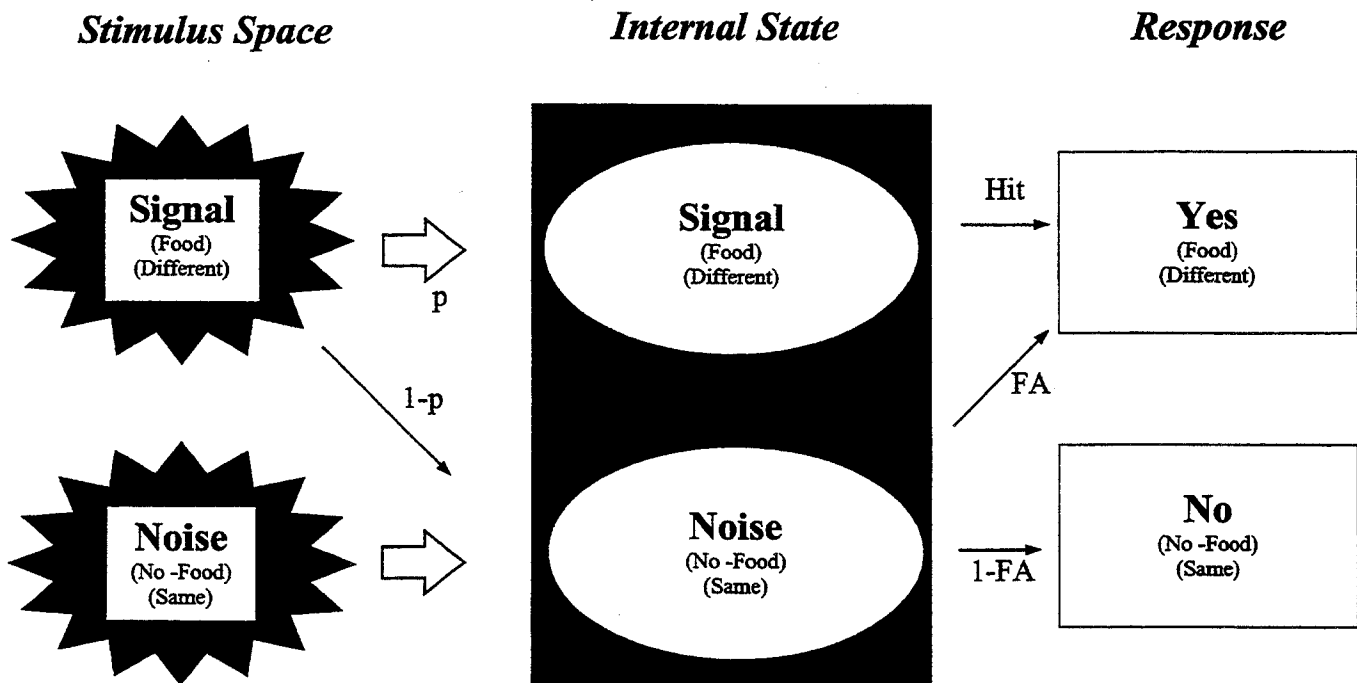

Figure 1. The hypothetical pathways associated with making a decision based on high-threshold theory (see text for details). FA = false alarm; p = proportion of signal trials exceeding signal threshold.

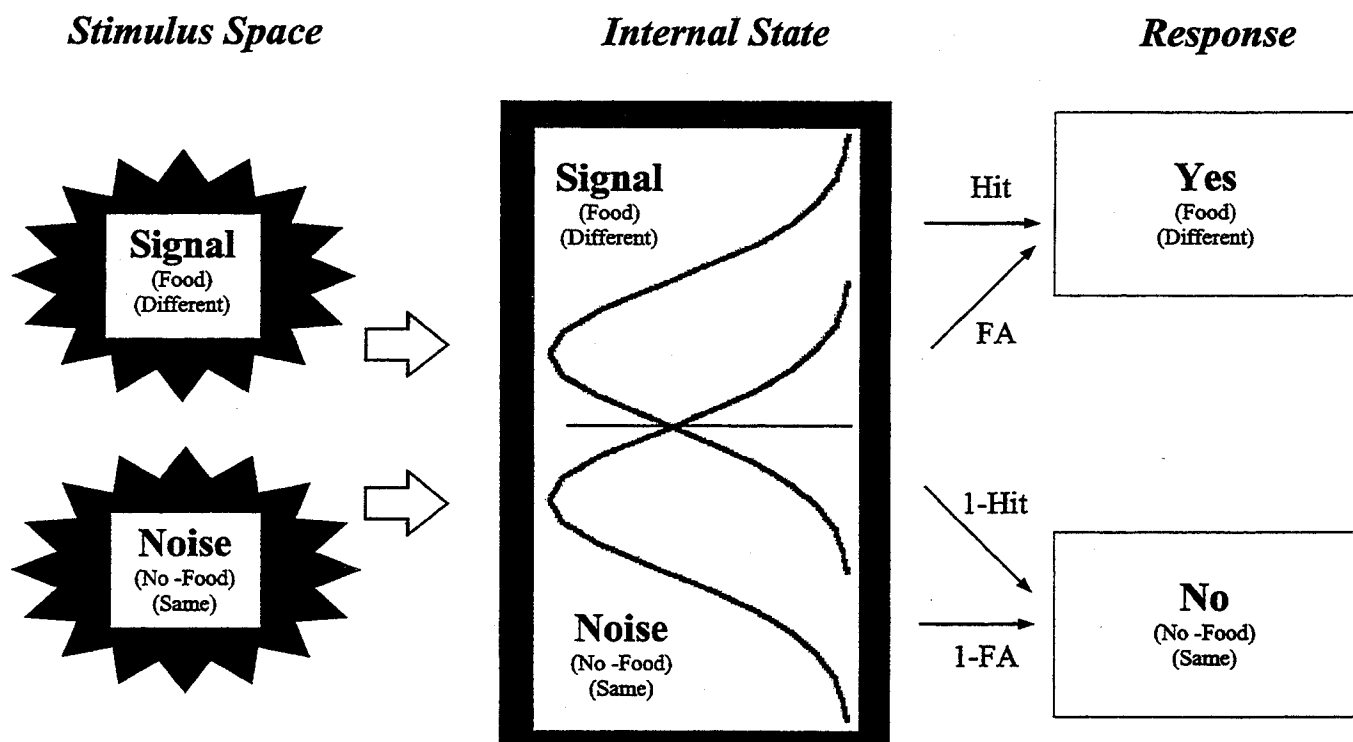## Stimulus Space   Internal State   Response

Figure 2. The hypothetical pathways associated with making a decision based on signal detection theory (see text for details). FA = false alarm.

choice pathways of this particular theory and how they contrast with those of HTT.

Figures 1 and 2 show that an important difference between HTT and SDT concerns the nature of the errors made on noise-only trials. Because in HTT no internal representation of a signal is possible on noise trials (i.e., there is no path between the *no* signal stimulus and the internal state of a signal), this model says that *false alarms* (choosing *yes* on noise trials) come from an incomplete bias on the part of the observer to choose one of the response alternatives. Returning to our original example, the pigeons on no-food trials have a very strong, but not complete, bias to choose the no-food alternative, resulting in the occasional choice of the food alternative. On the other hand, SDT proposes that false alarms are the product of the variability associated with the representation of the stimulus and the subject's criterion. Because the perceptual system is imperfect, such variability exists even on noise trials. As such, when the strength of evidence due to this noise exceeds the decision criterion, the subject mistakenly reports that a signal occurred. This important theoretical difference between HTT and SDT is empirically manifested in their different predictions about the form of the receiver operating characteristic (ROC). The ROC curve is a standard psychophysical graph showing the hit rate as a function of the false alarm rate over conditions in which the subject's response bias has been manipulated. The predicted shape of the ROC

according to the HTT model is a linear function:

$$H = p + (1 - p)FA, \quad (1)$$

where H represents the hit rate, FA represents the false-alarm rate, and p represents the proportion of signal trials in which the signal strength theoretically exceeded the threshold (see details in Swets et al., 1961). In contrast, the ROC function predicted from SDT is curvilinear. Although there is no exact algebraic expression for this function, a close approximation is given by

$$H = 1/\{e^{-ad'(1+r)}[(1/FA) - 1]^r + 1\}, \quad (2)$$

where a is a constant equal to 0.85 (Ogilvie & Creelman, 1968). This is a two-parameter model in which d' represents the distance between the mean of the signal distribution and the mean of the noise distribution, and r represents $1/\sigma$, where σ is the standard deviation of the signal distribution. For simplicity, both of these parameters are scaled in units equal to the standard deviation of the noise distribution, which is arbitrarily set to 1.

Wixted's (1993) and Dougherty and Wixted's (1996) experiments tested whether an HTT-default response or an SDT account provided a better description of the choice performance of pigeons in a matching task requiring them to remember the presence or absence of a sample stimulus.

Analyses of the pigeons' ROC functions revealed stronger support for a curvilinear signal detection analysis of their performance than the linear HTT–default response account. This outcome suggests that the pigeons made a memory-based decision on every trial instead of responding by default on trials that did not involve a memory trace.

Extension to Same–Different Texture Discriminations

With this as background, we can return to the question that motivated the current experiment: What are the mechanisms underlying same–different discrimination and concept formation by pigeons as found in recent texture experiments? Figure 3 shows some examples of the many Same and Different displays used in the experiment described below. In the current experiment, as well as in Cook, Cavoto, et al. (1995), the Different displays contained a contrasting target region on a uniform background. The Same displays consisted of a uniform set of elements having no target region. A hopper-choice procedure was used to test the pigeons' discrimination between these two classes of stimuli. Following the presentation of a to-be-discriminated texture stimulus in this procedure, Same and Different choice hoppers, located on the right and left sides of the testing chamber, are illuminated indicating that a choice can now be made. The pigeon then indicates its choice by entering its head into one of the two hoppers (LEDs just in front of the hoppers detect this behavior). If the entered hopper is the correct one for that stimulus, it is raised allowing access to its grain contents. If it is the incorrect hopper that is entered instead, the pigeon experiences a brief time-out in the dark.

Cook, Cavoto, et al. (1995) found that pigeons readily learned this type of discrimination, could discriminate large numbers of displays concurrently, and could transfer this discrimination to novel color and shape textures not experienced during training, suggesting the pigeons may have used a generalized same–different rule to solve the task.

Of particular relevance to the current discussion is a subsequent unpublished experiment (Cook, Blondeau, Cavoto, & Katz, 1995) that involved manipulating the texture display's presentation duration. This experiment revealed an asymmetry in the pigeons' processing of these types of Same and Different displays, an asymmetry similar to that observed when the delay length is varied in the food and no-food discriminations discussed previously. Specifically, the pigeons showed a strong and relatively constant tendency to correctly respond same on Same trials at both very short and long presentation durations, whereas performance with Different displays systematically improved with increasing duration.

Both HTT and SDT can provide reasonable accounts of both the original discrimination studied by Cook, Cavoto, et al. (1995) as well as this latter asymmetrical duration effect. According to HTT, the Different (i.e., target) displays produce a target code, whereas the Same (i.e., uniform) displays produce no comparable target code. When a target code is present, the Different hopper is chosen; whereas in its absence, the Same hopper is chosen by default. At short durations, the display is too brief to generate a target code even if a target is present in the display. Thus, on both short duration Same and Different trials, the Same hopper is chosen by default. At longer durations, the display will be
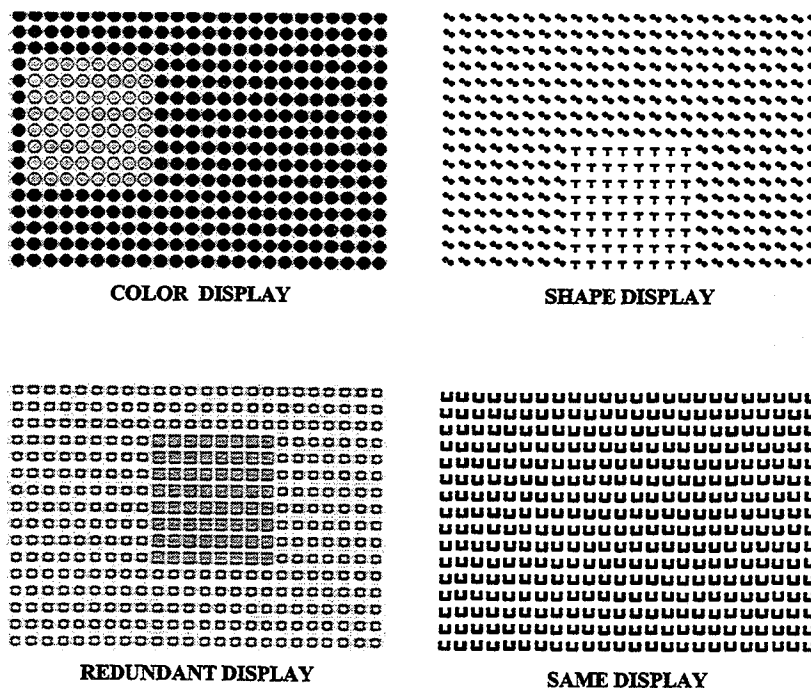

COLOR DISPLAY   SHAPE DISPLAY

REDUNDANT DISPLAY   SAME DISPLAY

Figure 3. Illustrative examples of the many color, shape, and redundant Different and Same displays used in these experiments.

presented long enough for the pigeon to detect the target if one is present. Thus, on Different trials, the tendency to choose the Different hopper should increase with presentation duration. On Same trials, by contrast, the tendency to choose the Same hopper should remain constant as no target code is created whether the duration of the display is short or long. This is precisely the pattern that was observed.

SDT offers an equally plausible explanation for the observed asymmetry. From this point of view, the Different displays can be treated as the equivalent of signal + noise trials, whereas Same displays function as noise-only trials. On the basis of sensory evidence collected during presentation, and in conjunction with a decision criterion, the pigeons then decide which choice alternative to select. When a display is presented for a short time, whether a target is present or not, there is little time to detect it and thus brief displays have a greater likelihood of being perceived as uniform than not. As a result, most responses will be directed toward the Same hopper at short presentation durations on both Same and Different trials. At longer presentation durations, evidence for the presence of a target increases on Different trials (increasing the tendency of the pigeons to choose the Different hopper), but not on Same trials. Thus, on Different trials the tendency to choose the Different hopper should increase with presentation duration, whereas on Same trials the tendency to choose the Same hopper should not change.

The experiment in this article was designed to test these two alternative conceptions of the pigeons' performance of the same–different texture discrimination task. This was done by testing the fits of the linear HTT and curvilinear SDT predictions to ROC functions collected from pigeons experienced in this same–different task (i.e., the same pigeons as used by Cook, Cavoto, et al., 1995). Individual ROC functions for 6 pigeons were generated by systematically varying the a priori probabilities of Same and Different displays across blocks of 18 sessions. Each block defined a different biasing condition in which (a) the majority of the trials within a daily session were Same trials (*Same bias* condition) (b) an equal number of Same and Different trials were presented (*Equal bias* condition), or (c) the majority of trials within a session were Different trials (*Different bias* condition). The effect of this response biasing manipulation was assessed by looking at the hit rate to Different displays and false-alarm rate to Same displays at the end of each of these biasing phases. These values were then used to generate functions that let us directly evaluate the contrasting ROC predictions of the HTT and SDT accounts of the pigeons' performance.

Avian Element and Compound Stimulus Processing

Besides deciding between these rival theories of choice behavior, we were also keenly interested in examining the issue of information integration in animals (cf. Cook et al., 1992; Heinemann & Chase, 1970; Kehoe & Gormezano, 1980; Riley, 1984). Following the completion of Cook, Cavoto, et al.'s (1995) observations, redundant displays involving targets that differed from the background along both color and shape dimensions were added to the daily sessions

of the pigeons. As might be expected, these redundant Different stimuli produced superior performance relative to the elemental color and shape Different displays. Our main focus here is on which of several plausible models of compound stimulus integration (cf. Macmillan & Creelman, 1991; Mulligan & Shaw, 1980; Shaw, 1982) best explains this redundant facilitation in the pigeons' performance. That is, how was the information from the color and shape dimensions combined when both redundantly defined the target?

Answering the question of how pigeons integrate color and shape information presupposes knowledge of their decision strategy. In the human literature, SDT has repeatedly been shown to adequately account for decision making in visual discrimination tasks. Within that framework, several models of information integration have been proposed (although the models could be worked out in terms of HTT as well). These standard models can be distinguished by the number of input channels they postulate (one vs. two) and the manner in which multidimensional information is jointly used to arrive at a decision (independent observations vs. additive integration).

The independent observations model is a two-channel model that postulates that color and shape information are processed in separate and independent channels. Given this assumption, the pigeon then may make two separate signal detection decisions on each trial—(a) is a shape difference detected? and (b) is a color difference detected?—and choose the Different hopper if either of these answers is yes. The integration model is a second type of two-channel model that postulates that the evidence for a difference from the shape channel is added to evidence for a difference from the color channel. The Different hopper is then chosen if this summed evidence exceeds the decision criterion. Closely related to the integration model is a unidimensional model, which by contrast assumes that only a single channel is used to evaluate both the color and shape information in a display (i.e., they are not separately processed). This model postulates that the pigeons process only one kind of information from the display on each trial, such as the degree of difference or target contrast present. If only one such channel of information is used, then on redundant trials the difference evidence provided by the color dimension would be added to difference evidence provided by the shape dimension. As will be detailed later, these three information-integration models make different predictions about (a) how much d' should improve on redundant trials given color and shape performance, and (b) the degree of variability calculated for the signal distribution associated with these redundant trials. Quite propitiously, the biasing manipulation designed to address the HTT versus SDT question also provides the information needed to distinguish between these information-integration models as well.

Method

Animals

Six male White Carneaux pigeons (*Columba livia*) were tested. No training was needed as these pigeons had participated in this

discrimination task for the prior 2 years (Cook, Cavoto, et al., 1995). They were maintained at 80% of their free-feeding weights in a colony room with a 12-hr light–dark cycle and during testing had free access to water and grit in their home cages.

Apparatus

Testing was conducted in a flat-black Plexiglas chamber (39 cm wide × 33 cm deep × 41 cm high). All stimuli were presented by a computer on a color monitor (NEC Multisync 2A; Wooddale, IL) visible through a 26 × 18 cm viewing window in the middle of the front panel. The viewing window's bottom edge was 20 cm above the chamber floor. A thin piece of glass mounted in this window protected the monitor. Pecks to the monitor screen were detected by an infrared touchscreen (resolution of 80 × 48; EMS Systems, Champaign, IL) mounted behind a 40-mm wide Plexiglas ledge that went around the inside edge of the viewing window. A 28-V houselight was located in the ceiling and illuminated at all times, except when an incorrect choice was made. Identical food hoppers (Coulbourn #E14-10, Lehigh Valley, PA) were located in the right and left walls of the chamber, each 3 cm from the front panel and flush with the floor. Infrared LEDs mounted 2.5 cm in front of each hopper detected the approach of a pigeon's head toward its opening.

All experimental events were controlled and recorded with a 386-class computer. A video card (VGA Wonder; ATI Technologies, Scarborough, Ontario, Canada) controlled the monitor in the VGA graphics mode (640 × 480 pixels). Computer-controlled relays (Metrabyte, Taunton, MA) operated the hoppers and houselight. Stimulus and event programming were done with QuickBasic (1989) with an attached graphics library (Xgraf, 1989). Programming of the video card's color palette was used to control stimulus onset and offset.

Procedure

Stimulus material. Texture stimuli were 18 × 12 cm in size and consisted of 384 elements arranged in a 24 × 16 matrix at 0.75-cm intervals. Individual elements were 3–6 mm in size depending on their shape. Figure 3 shows examples of each type of the display tested in the experiment (see also texture examples in Figure 1 of Cook, Katz, & Cavoto, 1997). Different displays consisted of an 8 × 8 target region of small colored elements embedded in a surrounding 24 × 16 region of distractor elements. On each trial, this target region was randomly located within the 24 × 16 matrix of distractors (153 total locations). Same displays were made by repeating the same element throughout the 24 × 16 display matrix. One hundred small elements were used to create the Different and Same displays. These elements were made from the pairwise combination of 10 colors (blue VGA #1, green #2, cyan #3, red #4, purple #5, brown #20, gray #7, orange #52, rose #37, yellow #54) and 10 shapes (U, T, triangle, square, chevron, dots, circle, vertical line, horizontal line, star).

These elements were combined to form three types of Different displays. In color Different displays the target region's repeated element was a different color from the distractor region's repeated element, but both elements had the same shape. There were 900 possible color displays (10 target colors × 10 distractor colors × 10 irrelevant shapes minus Same displays). In shape Different displays the target region's repeated element was a different shape from the distractor region's repeated element, but both were the same color. There were 900 possible shape displays (10 target shapes × 10 distractor shapes × 10 irrelevant colors minus Same displays). In redundant Different displays, the target region's repeated element differed in both color and shape from the distractor region's repeated element. There were 8100 possible redundant displays (100 target elements × 100 distractor shapes minus color Different and shape Different and Same displays). There were 100 possible Same displays. The specific stimuli tested in each session were randomly selected from these pools of Different and Same displays.

Discrimination procedure. Each texture discrimination trial began with a peck to a circular white ready signal (1.5 cm in diameter). This ready signal was randomly located within the upper half of the display area on each trial. This was followed by presentation of a Different or Same display for either a fixed number of pecks or fixed duration, after which the texture stimulus was turned off and the two choice hoppers illuminated. A pigeon indicated its choice by entering its head into one of the two hoppers. If the correct hopper was selected, it was raised for 2 s. If the incorrect hopper was selected, the hopper lights were turned off and the overhead houselight extinguished for 15 s. A 5-s intertrial interval followed either outcome.

Daily sessions consisted of 164 texture discrimination trials. Of these trials, 72 were used to test stimuli presented at four different stimulus durations (100, 500, 1000, 2000 ms). Three randomly sampled redundant, color, and shape Different stimuli, and nine Same stimuli, were tested at each duration. In addition, two no-stimulus *blank* trials were also conducted at each duration. These blank trials were conducted like the other duration trials, except that no texture stimuli appeared after the pigeons' pecked the ready signal, and either choice response was rewarded. These trials evaluated the pigeons' hopper preference in the absence of stimulus information. The number and composition of these duration trials (72 standard duration and 8 blank trials) did not change across the biasing manipulation (see *Response biasing procedures*). The remaining 84 trials of each session involved texture stimuli presented for a fixed number of responses (FR) prior to the illumination of the choice hoppers. Each pigeon had a different FR requirement, which had in the past produced comparable levels of choice accuracy (Kate = 1, Ethel = 2, Ellen = 2, Lucy = 5, Betty = 2, and Ginger = 2 pecks).

Response biasing procedures. The response biasing manipulation used to generate the ROC functions was implemented by changing the proportion of the Different and Same FR trials in a session. The number of FR trials was adjusted so that either 30%, 50%, or 70% of the total stimulus trials (FR and duration combined = 156 trials) consisted of Same displays, and correspondingly 70%, 50%, or 30% of the trials tested Different displays. Same biasing condition sessions consisted of 72 Same and 12 Different stimuli (4 color, 4 shape, and 4 redundant) FR trials plus the 72 duration trials. Different biasing condition sessions consisted of 12 Same and 72 Different stimuli (24 color, 24 shape, and 24 redundant) FR trials plus the 72 duration trials. Equal biasing condition sessions consisted of 42 Same and 42 Different stimuli (14 color, 14 shape, and 14 redundant) FR trials plus the duration trials. Each pigeon participated in a total of five 18-session blocks testing these three biasing conditions. Each pigeon was tested twice in the Same and Different biasing conditions and once in the Equal biasing condition. The testing order for each pigeon is presented in Table 1.

Results

The analysis of this experiment is divided into three sections. The first section summarizes basic performance across the three biasing conditions. The second section compares and evaluates the fit of the SDT and HTT models

Table 1
Testing Order of Biasing Conditions for Each Pigeon

Pigeon	Block				
	1	2	3	4	5
Kate	Equal	Same	Different	Same	Different
Ellen	Equal	Different	Same	Different	Same
Ethel	Different	Same	Equal	Different	Same
Lucy	Different	Equal	Same	Different	Same
Betty	Same	Equal	Different	Same	Different
Ginger	Same	Different	Equal	Same	Different

to these observations. The third section compares and evaluates the three different models of information integration. As detailed below, from these analyses we conclude that an unequal variance signal detection model better describes the pigeons' performance in this task than does the HTT–default response model. Furthermore, the final parameters derived from the best fitting SDT solution also best conform to those predicted by the unidimensional integration model—suggesting these pigeons were making their discriminations on the basis of the extraction of a single type of information from across the different types of textured stimuli.

Basic Performance and Effects of Response Biasing

As these pigeons were highly experienced at the task, performance was good from the beginning of the experiment. Over the experiment, the pigeons' mean accuracy was 82% on FR trials, with their performance with Same displays (87%) being slightly better than with Different displays (77%). Among the three types of Different displays, the percentage of correct choices was greater following redundant displays (91%) than with either color (69%) or shape (70%) displays. As evaluated by a repeated measures analysis of variance (ANOVA), the differences among these Different displays were significant, $F(2, 10) = 31.4$ (all statistical tests reported in this article used an alpha level set at $p \leq .05$).

Because the data from the duration trials were not of primary interest in the present context, they are not discussed below, except to note the asymmetry in the pigeons' responding to the Same and Different displays as a function of stimulus duration. Over the 90 experimental sessions, the pigeons generally responded quite accurately on Same trials regardless of their duration (100 ms = 65%, 500 ms = 87%, 1000 ms = 90%, 2000 ms = 91%), whereas performance on Different trials (color, shape, and redundant combined) systematically increased with increasing duration (100 ms = 46%, 500 ms = 64%, 1000 ms = 72%, 2000 ms = 77%). For reasons perhaps related to further experience, the size of the asymmetry was somewhat reduced, but in the same general direction as found in Cook, Blondeau, et al.'s (1995) results.

The biasing manipulation, as implemented by the variation in a priori probabilities of Same and Different FR trials across blocks, had its intended effects on response bias. That

is, the pigeons' inclination to choose the Same and Different choice alternatives was affected by the relative frequency of each type of display. For many of the following analyses, we particularly concentrated our attention on those data from the last the six sessions of each of the five test blocks, because it was in these sessions that we expected to see the greatest cumulative effect of the biasing manipulation. Over the last six sessions of the Different biasing, Equal biasing, and Same biasing conditions, the percentage of *same* responses on Same trials was 76%, 92%, and 96%, respectively, and the percentage of *different* responses on Different trials was 82%, 76%, and 72%. A repeated measures ANOVA confirmed that these systematic changes in response tendencies were reliable, as indicated by the presence of a significant Biasing Condition × Trial Type interaction in these responses, $F(2, 10) = 14.4$.

Examination of individual choice behavior confirmed this pattern was the same for all pigeons except one (Betty). This pigeon showed little sensitivity to the biasing manipulation, which is easily seen by the very close clustering of the data points in its ROC plots in Figures 4, 5, and 6. Why this pigeon was unaffected by the biasing manipulation is unclear. Because of this fundamental insensitivity, this pigeon's data were not useful in discriminating between the various models, and were excluded from the remaining analyses discussed below.

Evaluation of the SDT and HTT Models

We evaluated next whether the HTT or SDT account provided a better description of the choice behavior of the 5 pigeons that were affected by the biasing manipulation. Figures 4, 5, and 6 show the ROC functions for each pigeon with color, shape, and redundant displays, respectively. These ROC plots represent performance in terms of the pigeons' hit and false-alarm rates across the different blocks of the biasing manipulation. The hit rate is the proportion of Different trials in which the pigeons correctly reported the presence of a target. The false-alarm rate is the proportion of Same trials in which the pigeons incorrectly reported the presence of a target. The larger black data points in each plot are the mean hit and false-alarm rates from the last six sessions of each of the five test blocks. As mentioned earlier, these data best reflect the cumulative effects of biasing. These last six sessions of each block were used for the fits generated by the maximum likelihood estimation procedure

Color Trials

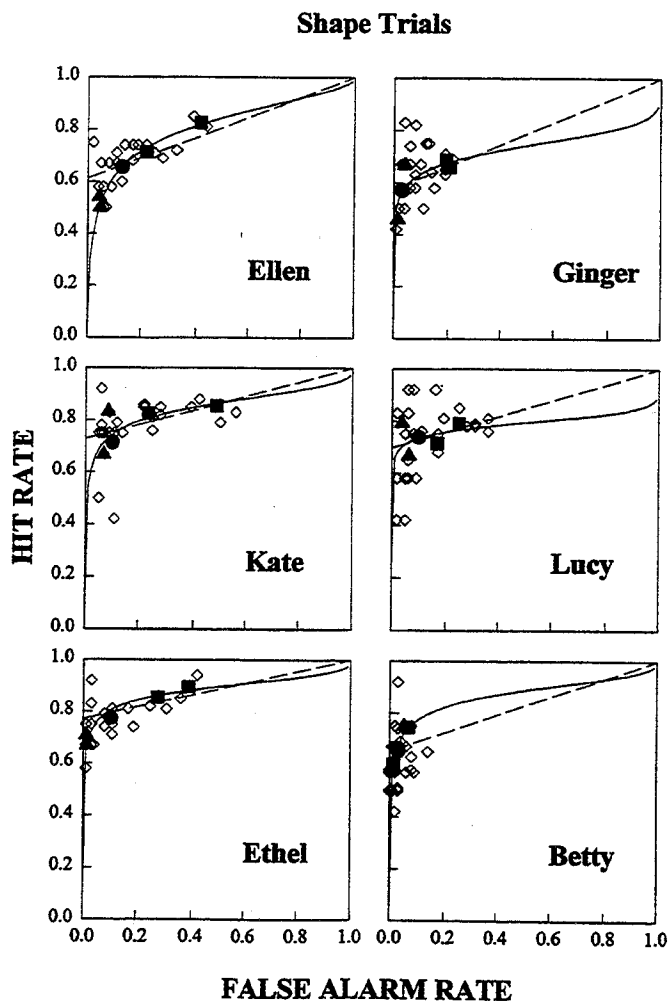**HIT RATE**
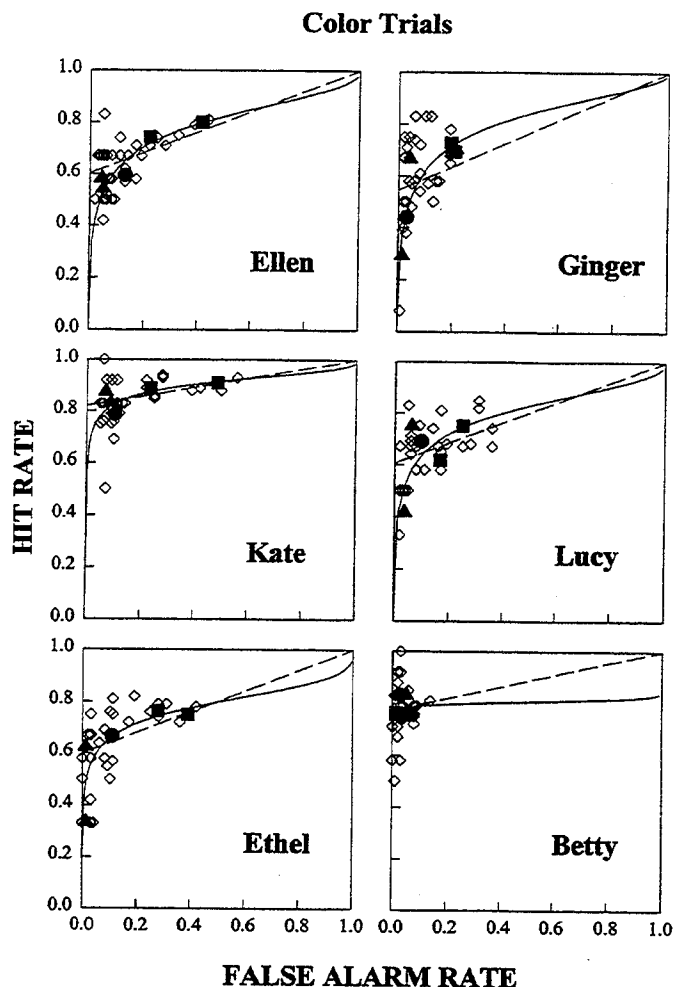
FALSE ALARM RATE

Figure 4. Individual receiver operating characteristic curves of color performance for each pigeon. Each plot shows the mean hit rate versus false-alarm rate for the last six sessions of the five different biasing phases of the experiment (black squares = Different biasing conditions; black circles = Equal biasing condition; black triangles = Same biasing conditions) and for successive three-session blocks across the entire experiment (open diamonds). The solid curves represent the best fitting functions based on signal detection theory. The dashed lines represents the best fitting function based on high-threshold theory.

described next. The smaller unfilled data points in each plot are the pigeons' hit and false-alarm rates across all sessions of the experiment as blocked by groups of three successive sessions (30 total points). These data points capture the intermediate changes in the pigeons' response bias as they reacted to the new proportions of trials within and between biasing conditions. It is these unfilled data points that were used in the least squares estimations described later.

The black square data points higher and to the right in each ROC plot represent performance from sessions collected under the Different bias condition. The black triangles that are generally lower and to the left represent performance from sessions collected under the Same bias condition. The black circles represent performance from sessions collected under the Equal bias condition. The curves drawn

in each plot represent the best fit of the SDT (solid line) and HTT (dashed line) models as based on these five data points in each plot. These fits were generated by using the maximum likelihood estimation (MLE) procedure (Ogilvie & Creelman, 1968). This method involves adjusting the free parameters of each model until the theoretical probability of obtaining the observed distribution of hits, misses, correct rejections, and false alarms reaches the highest possible value. The MLE method was used first to fit the data because it takes into account the error in measurement in both hit and false-alarms rates, unlike the more traditional least squares method (Ogilvie & Creelman, 1968). The SDT fits were actually performed two ways, once by using the equation for logistic distributions (Ogilvie & Creelman, 1968) and once

Shape Trials

HIT RATE

FALSE ALARM RATE

Figure 5. Individual receiver operating characteristic curves of shape performance for each pigeon. Each plot shows the mean hit rate versus false-alarm rate for the last six sessions of the five different biasing phases of the experiment (black squares = Different biasing conditions; black circles = Equal biasing condition; black triangles = Same biasing conditions) and for successive three-session blocks across the entire experiment (open diamonds). The solid curves represent the best fitting functions based on signal detection theory. The dashed lines represent the best fitting functions based on high-threshold theory.

Redundant Trials
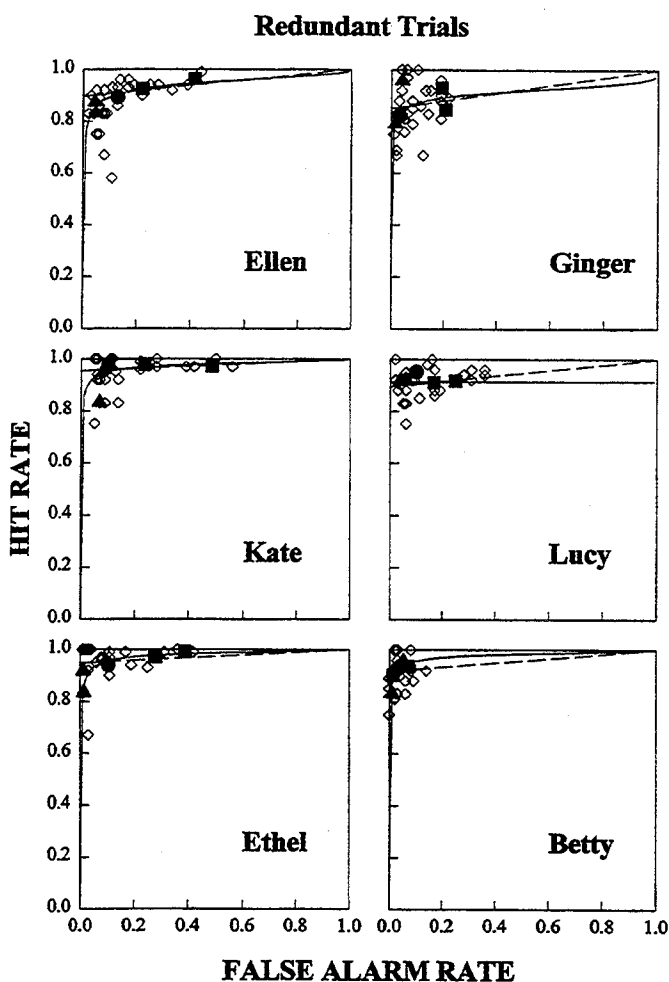
Figure 6. Individual receiver operating characteristic curves of redundant performance for each pigeon. Each plot shows the mean hit rate versus false-alarm rate for the last six sessions of the five different biasing phases of the experiment (black squares = Different biasing conditions; black circles = Equal biasing condition; black triangles = Same biasing conditions) and for successive three-session blocks across the entire experiment (open diamonds). The solid curves represent the best fitting functions based on signal detection theory. The dashed lines represent the best fitting functions based on high-threshold theory.

by using an approximation of the Gaussian distribution (Equation 26.2.19 in Abramowitz & Stegun, 1965). Because the results were virtually identical, only the estimates based on the logistic are reported. For the HTT fits, the equation for a straight line was used in place of the logistic (with one guessing parameter for each biasing condition and one overall discriminability parameter, p), but the MLE procedure was otherwise identical.

The final best fitting MLE parameter values for both the SDT and HTT models are listed in Tables 2 and 3. The MLE fit for each model was assessed by the chi-square statistic. A significant chi square indicates that the observed deviations from a model's predictions for that type of display were greater than would be expected by chance. The HTT model yielded mixed results overall, with significant deviations

Table 2
Maximum Likelihood Parameter Estimates (p) for the High-Threshold Model

Pigeon	Color	Shape	Redundant
Ellen	.60	.61	.90
Ethel	.60	.77	.95
Ginger	.54	.57	.85
Kate	.82	.73	.95
Lucy	.60	.70	.89
M	.63	.68	.91

from this model observed in three cases and marginally significant deviations observed in four cases: Ellen—color, $\chi^2(4) = 4.1$, shape, $\chi^2(4) = 4.4$, redundant, $\chi^2(4) = 3.4$; Ethel—color, $\chi^2(4) = 9.2, p < .10$, shape, $\chi^2(3) = 6.3, p < .10$, redundant, $\chi^2(3) = 12.9, p < .05$; Ginger—color, $\chi^2(4) = 18.2, p < .05$, shape, $\chi^2(4) = 2.6$, redundant, $\chi^2(4) = 8.4$, $p < .10$; Kate—color, $\chi^2(4) = 2.8$, shape, $\chi^2(4) = 3.4$, redundant, $\chi^2(4) = 10.5, p < .05$; Lucy—color, $\chi^2(4) = 9.1$, $p < .10$, shape, $\chi^2(4) = 2.1$, redundant, $\chi^2(3) = .1$. The SDT fits were better with significant deviations from this model observed in only one case and marginally significant deviations observed in three others: Ellen—color, $\chi^2(3) = 1.8$, shape, $\chi^2(3) = .03$, redundant, $\chi^2(3) = .8$; Ethel—color, $\chi^2(2) = 6.5, p < .05$, shape, $\chi^2(3) = 1.2$, redundant, $\chi^2(3) = 2.4$; Ginger—color, $\chi^2(3) = 3.5$, shape, $\chi^2(3) = 3.5$, redundant, $\chi^2(3) = 7.7, p < .10$; Kate—color, $\chi^2(3) = 2.0$, shape, $\chi^2(3) = 2.2$, redundant, $\chi^2(3) = 7.0, p < .10$; Lucy—color, $\chi^2(3) = 7.0, p < .10$, shape, $\chi^2(3) = 2.7$, redundant, $\chi^2(2) = .8$ ($N = 1,680$ for all chi-square tests; all chi-squares for SDT fits were evaluated by using one less degree of freedom due to the additional free parameter in this model).

At first inspection, both models appear to be adequate first approximations of the pigeons' choice behavior. For several reasons, we believe the curvilinear predictions of the SDT account provides for a better description of the pigeons' behavior. Closer inspection of these ROC plots, for example, reveals systematic deviations from the predictions of the HTT model—deviations that are more consistent with the

Table 3
Maximum Likelihood Parameter Estimates (σ and d') for the Signal Detection Model

Pigeon	Color		Shape		Redundant	
	σ	d'	σ	d'	σ	d'
Ellen	2.1	1.2	1.8	1.2	1.8	2.5
Ethel	3.0	1.2	2.9	1.9	1.7	3.2
Ginger	1.8	1.3	4.3	1.3	4.2	2.3
Kate	2.6	2.0	2.7	1.6	2.1	3.2
Lucy	2.1	1.3	6.5	1.4	—	2.8
M	2.3	1.4	3.6	1.5	2.4	2.8

Note. An estimate of σ for Lucy in the redundant condition could not be obtained.

predictions of SDT. This is especially true for the color and shape data for 3 of the pigeons (Ellen, Ginger, and Ethel) and more moderately so for the other 2 (Kate and Lucy).

Specifically, in most cases the left-most large black points fall below and the right-most points fall above the best fitting HTT prediction. These are exactly the types of deviation that would be expected if the data were actually following the curvilinear path predicted by SDT. In the color condition, 7 of the 10 left-most points (2 from each pigeon) fell beneath the dashed lined and 7 of the 10 right-most data points fell above the dashed line. The probability of obtaining 14 or more deviations in the predicted direction out of 20 opportunities is .058 by chance alone. In the shape condition, 17 of 20 deviations were in the direction predicted by SDT ($p = .003$), and in the redundant condition 16 of 20 were in the predicted direction ($p = .006$). Thus, although the high-threshold fits were not grossly inaccurate, the systematic nature of the observed deviations across pigeons suggests that the HTT model is fundamentally incorrect.

Evaluation of these two models by using the more traditional least squares methods further confirmed this last conclusion. Despite some technical limitations when applied to ROC analyses (Ogilvie & Creelman, 1968), this analysis permits the evaluation of performance by using more of the data than could be done with the superior MLE procedure. For these analyses we used the 30 points derived by grouping the results into three-session blocks (the unfilled diamonds in Figures 4, 5, and 6). The best fitting parameter values for each model obtained by using this method are shown in Tables 4 and 5. Overall, the results are virtually identical to those generated by the MLE procedure (see Tables 2 and 3).

Of most importance, the same pattern of systematic deviations described above were found for the HTT model but not for the SDT account. Linear regressions through the residuals between the obtained and predicted values for each pigeon's 30 points revealed a significant positive slope (mean $r = .25$) for the HTT fit, $t(4) = 3.7$. This positive slope to the HTT residuals again reflects its consistent overestimation of hit rate for the Same biasing conditions and underestimation of this value for the Different biasing condition. The same analysis for the SDT fit revealed no significant slope for its residual values (mean $r = -.03$), indicating no systematic deviations of this model's fit to the data.

The blank trials, where no stimulus was presented,

Table 4
Least Squares Parameter Estimates (p) for the High-Threshold Model

Pigeon	Color	Shape	Redundant
Ellen	.57	.61	.90
Ethel	.57	.75	.94
Ginger	.58	.58	.84
Kate	.80	.72	.93
Lucy	.60	.68	.89
M	.62	.67	.90

Table 5
Least Squares Parameter Estimates (σ and d') for the Signal Detection Model

Pigeon	Color		Shape		Redundant	
	σ	d'	σ	d'	σ	d'
Ellen	3.5	1.1	3.2	1.1	3.0	2.3
Ethel	2.7	1.2	4.6	1.7	3.8	3.3
Ginger	2.8	1.3	3.7	1.1	3.1	2.4
Kate	2.8	2.0	3.6	1.6	1.3	3.0
Lucy	3.1	1.2	3.2	1.5	7.6	2.7
M	3.0	1.4	3.7	1.4	3.8	2.7

included in this experiment provided yet another means of testing the adequacy of the HTT–default response model. For HTT to account for the asymmetrical duration results both described in the introduction and obtained during the experiment—where the pigeons showed a bias to choose Same across all durations—the default response of the pigeons would have had to have been directed towards the Same choice alternative. Examination of the results from the blank trials, however, collected during the last six sessions of each test block, showed that all pigeons, except Betty again, showed a bias to choose the Different hopper: Ellen, 61%; Ethel, 90%; Ginger, 72%; Kate, 69%; and Lucy, 67%. Only Betty showed a bias to choose the Same alternative, entering this hopper on 73% of the blank trials. This general bias towards the Different hopper in the absence of stimulus information is yet another line of converging evidence against any HTT account of the pigeons' choice behavior.

Collectively, the weight of the evidence (the overall better fit, the lack of systematic deviations in the residuals, the pattern of blank trial responding) all more strongly favor an SDT account of these results over an HTT–default response account. Given this, it makes sense to examine more closely the parameters obtained from the best fit signal detection model. These parameter estimates revealed an interesting pattern of three results that are novel and theoretically revealing about the structure of the pigeons' representation of the current task. First, the average value of d' in the compound redundant condition was essentially equal to the sum of the d' values from the elemental color and shape conditions. This was true for both the MLE ($1.4 + 1.5 \approx 2.8$, from Table 3) and least square fits ($1.4 + 1.4 \approx 2.7$, from Table 5) of the data. Paired t tests revealed no significant differences between the additive combination of these element d' values ($d'_{color} + d'_{shape}$) and the observed d' value for the compound redundant trials, $ts(4) < 1$. Secondly, the values estimated for σ for all the three Different conditions were consistently greater than 1 (the expected value given an equal variance SDT solution). The mean σ value across the three conditions was 2.92 based on the MLE fit and 3.46 for the least squares fit, values both significantly greater than the predicted value of 1: MLE, $t(4) = 4.95$; least squares, $t(4) = 7.3$. Thirdly, the standard deviations of the signal distributions estimated for the color, shape, and redundant displays were similar to one another. For the MLE fits these estimates

were: color = 2.3, shape = 3.6 (2.8, if one excludes Lucy's uncharacteristic value), redundant = 2.4. For the least square fits these estimates were: color = 3.0, shape = 3.7, redundant = 3.8 (2.8 without Lucy's value). Broadly speaking, it thus appeared that the estimated standard deviations of the signal distributions (i.e., Different displays) were generally about the same for all three Different conditions, and about three times greater than that of the noise distribution (i.e., the Same displays).[1] Collectively, these three outcomes motivated our next analyses.

Models of Information Integration

In the present experiment, the Different trials were defined by element differences along one of two dimensions (color or shape) or by a redundant compound of both of these dimensions (color and shape). One issue of long-standing theoretical interest is how such individual elements combine to govern performance on compound trials (for a review see Kehoe & Gormezano, 1980). With this goal, we next evaluated the fit of three different models of information integration to the present results (see Macmillan & Creelman, 1991; Mulligan & Shaw, 1980; Shaw, 1982). Two of these, the independent observations and integration models, assume that color and shape information are processed as separate independent channels, but use different combination rules for this information as the basis for a final decision. The third unidimensional model, in contrast, assumes only a single common channel of information is used in making judgments about all three types of Different displays.

As indicated earlier, the independent observations model assumes that a *yes* (Different) response is made if the strength of evidence for the presence of a target exceeds a decision criterion in either the color channel or the shape channel. As such, a false alarm on a Same trial will occur if noise in the color channel exceeds the color criterion or noise in the shape channel exceeds the shape criterion. If neither of these values exceeds its decision criterion, a correct *no* (Same) response is made. Similarly, a hit on a color trial will occur if signal strength from the color channel exceeds its criterion or if noise from the shape channel happens to exceed its criterion (and vice versa for a shape trial). If neither exceeds criterion, an incorrect *no* response will be made.

As described by Mulligan and Shaw (1980), the independent observations model predicts that the following relationship should hold:

$$(1 - p_{color})(1 - p_{shape}) = (1 - p_{same})(1 - p_{redundant}), \quad (3)$$

where p_{color}, p_{shape}, and $p_{redundant}$ represent the hit rate on color, shape, and redundant trials, respectively, and p_{same} represents the false-alarm rate on Same trials. In most previous tests of this kind, only one condition was run for each subject. Thus, each subject produced only one set of hit and false-alarm rates that could be used to test whether the above equality held. In our experiment, five biasing conditions were conducted, with each pigeon providing five separate sets of hit and false-alarm rates. Thus, the above equality could be tested a total of 25 times (5 times for each of 5 pigeons).

Even if the independent observations model were correct, one would not expect the left and right sides of Equation 3 to be exactly equal for all 25 tests. Nevertheless, by chance, one would expect the left side of Equation 3 to yield a slightly higher value than the right side as often as the right side yields a slightly higher value than the left. When each pigeon's hit and false-alarm rates were entered into the above equation, however, the left side consistently produced a larger value than the right side. Out of the 25 tests, the product on the left yielded the higher value 20 times (which according to the binomial would occur by chance with a .002 probability). As such, this independent observations model generates predictions that deviate systematically from the obtained data and can be safely rejected on those grounds.[2,3]

---

[1] This theoretically interesting result raised a technical issue. When the standard deviations of the signal and noise distributions differ, the unit in which d' is scaled is no longer equal to the standard deviation of the noise distribution, but is instead an intermediate value approximately (though not exactly) equal to the average of the standard deviation of the signal and noise distributions. Thus, if the standard deviations of the noise and signal distributions are 1 and 3 respectively, then the unit of measurement is about 2 (the mean of 1 and 3). A d' of 1 under these conditions implies that the means of the two distributions are separated by the average of the two standard deviations, and a d' of 2 implies that the means of the two distributions are separated by twice the average of the two standard deviations. The exact method of computing this measure, sometimes referred to as d'_e, is described by Macmillan and Creelman (1991, p. 71). They describe several other alternative d' measures that can be used when the signal and noise distribution variances differ. This is a technical point, however. Our conclusions remain unchanged regardless of which d' measure is used.

[2] A well-known prediction of the independent observations model is that $d'_{redundant}$ should equal the square root of $d'^2_{color} + d'^2_{shape}$. Because $d'^2_{color} = d'^2_{shape}$ it follows that $d'_{redundant}$ should equal the square root of $2 d'^2_{element}$ (i.e., $1.41 d'^2_{element}$) where $d'_{element} = d'_{color} = d'_{shape}$. That prediction only holds when the signal and noise distributions have equal variance, which is clearly not the case here.

[3] A further test of the independent observations model was performed by means of computer simulation. This test has the advantage of deriving predictions for both accuracy (as with the above test) and signal distribution variance on redundant trials. These simulations involved hypothetical color and shape distributions, and a hypothetical uniform distribution (the latter of which had a mean of 0 and a standard deviation of 1). All of the distributions were Gaussian in form. On a simulated uniform trial, a value representing noise in the color channel was randomly sampled from the uniform distribution, and a second value representing noise in the shape channel was sampled from the same distribution. If either exceeded its channel-specific criterion, a *different* response (i.e., a false alarm) was made, otherwise the response was *same*. On simulated color trials, one value was drawn from the color distribution (to simulate signal strength in the color channel) and another from the noise distribution (to simulate noise in the shape channel). Again, if either value exceeded its own

The additive integration model assumes that strength of evidence from the color channel is added to the strength of evidence from the shape channel to yield a summed variable. If this summed value exceeds a decision criterion, then a *different* response is made; otherwise a *same* response is made. For instance, a hit on a color trial will occur if signal strength from the color channel and the added noise from the shape channel exceeds the decision criterion. Likewise, a false alarm will occur on a Same trial, for example, if noise from the color channel when added to noise from the shape channel exceeds this decision criterion.

How does the integration model fare? The predictions of this model are more straightforward and can be derived analytically. On Same trials, noise from the color channel is theoretically added to noise from the shape channel. Because the mean of these two noise distributions is 0, the mean of the combined noise distribution will also be 0. On color trials, signal strength from the color channel is added to noise from the shape channel. That is, values from a distribution with a mean of d'_{color} are added to values from a distribution with a mean of 0 to produce a combined mean of d'_{color} (MLE and least squares estimates = 1.4). Similarly, on shape trials, the combined distribution is d'_{shape} (MLE = 1.5; least squares = 1.4). On redundant trials, signal strength from the color distribution (with a mean of d'_{color}) is added to signal strength from the shape distribution (with a mean of d'_{shape}) to theoretically yield a new distribution with a mean equal to the sum of the component distributions (i.e., d'_{color} + d'_{shape}). Thus, the predicted $d'_{redundant}$ from the MLE estimates would be 2.9 and from the least squares estimates 2.8. Both of these predicated values are indistinguishable from the observed d' values of 2.8 and 2.7 established with each of these methods.

---

criterion, a *different* response (i.e., a hit) was made. The process was repeated by using different criterion settings (to simulate different biasing conditions) and the resulting simulated ROC data were analyzed. Thus, for example, the criterion in both channels might be set at 0.5 standard deviations above the mean of the noise distribution for one simulation (which generates one point on the ROC), and at 1.0 for another simulation, and at 1.5 for still another. Note that the actual values used are unimportant so long as those values span a large enough range to generate points on the ROC that are reasonably spread out. The parameters of the color and shape distributions (mean and standard deviation) were adjusted manually until the simulated ROC data, when analyzed in the same way that the real data were analyzed, yielded estimated d' and σ values equal to the mean values reported in Table 2. Performance on redundant trials was then predicted by randomly sampling values from both the color distribution and the shape distribution. If either exceeded its decision criterion, a *different* response was made, otherwise a *same* response was made. Note that in these simulations, the equality shown in Equation 3 was always true (which must be the case for an independent observations model). When ROC data from the simulated redundant trials were analyzed, the predicted value of d' was approximately 4.9 (well above the observed value of 2.8), and the predicted value of σ was approximately 5 (also well above the observed value of 2.4). On the basis of these results, the independent observations model can be rejected.

The integration model also makes a clear prediction about the variance of the redundant signal distribution, a prediction that interestingly turns out to be incorrect. Consider that when values drawn from two distributions are summed, the resulting distribution has a mean equal to the sum of the means of the component distributions (as described above) and a variance equal to the sum of the variances of the component distributions. Generally speaking, therefore, if the integration model is correct the variance of the redundant distribution ($\sigma^2_{redundant}$) should be substantially greater than that of the individual color and shape distributions (σ^2_{color} and σ^2_{shape}).

To be more specific, it is important to first note that the variance of the combined noise distribution is equal to $\sigma^2_{N\text{-}color} + \sigma^2_{N\text{-}shape}$: that is, the variance of the noise distribution in the color channel plus the variance of the noise distribution in the shape channel. As such, on color trials the variance is theoretically equal to $\sigma^2_{color} + \sigma^2_{N\text{-}shape}$. It is this summed value divided by the summed values on noise trials that is theoretically measured in the ROC analysis of color trials. That is to say, the square root of ($\sigma^2_{color} + \sigma^2_{N\text{-}shape}$)/ ($\sigma^2_{N\text{-}color} + \sigma^2_{N\text{-}shape}$)—which is the variance on color trials relative to the variance on noise trials—is approximately equal to 3 on the basis of the estimates listed in Tables 3 and 5. The same logic applies on shape trials, where the absolute variance is theoretically equal to $\sigma^2_{shape} + \sigma^2_{N\text{-}color}$, and the relative standard deviation values shown in Tables 3 and 5 theoretically represent the square root of ($\sigma^2_{shape} + \sigma^2_{N\text{-}color}$)/ ($\sigma^2_{N\text{-}color} + \sigma^2_{N\text{-}shape}$). Thus, on redundant trials the absolute variance should be $\sigma^2_{color} + \sigma^2_{shape}$. If $\sigma^2_{N\text{-}color} + \sigma^2_{N\text{-}shape}$ (summed noise) is arbitrarily set to 1, then one can compute the predicted relative standard deviation estimates for the redundant trials for the integration model on the basis of the relative standard deviation estimates for the color and shape trials listed in Tables 3 and 5.

From the MLE estimates in Table 3, which tended to be slightly smaller than those for the least squares fits, the mean predicted relative standard deviation estimate for the redundant trials was 3.7. The observed value was 2.4. For all 4 pigeons for which a prediction could be derived (Lucy's fit did not converge on a solution in the redundant condition), the predicted values were greater than the observed values. The same phenomenon occured when predicted values were based on the least squares fits in Table 5. In this case, the mean predicted value was 4.6, whereas the observed value was 3.8. Again for all pigeons except Lucy, the predicted redundant values were greater than the observed values. Thus, although the integration model correctly predicts the d' relations between the color, shape, and redundant displays, it fails to account for the actual relations of σ observed between these different display types.

The last model we examined differs from the previous ones in that it involves a unidimensional assumption about the encoding of color and shape information. That is, the model does not envision separate contributions from the dimensions of color and shape used in combination to make judgments about each display. Instead, only a single and common type of information is being extracted from all

displays. This might be the degree of difference in the display or the strength of visual contrast formed at the boundary of the target–distractor regions. Regardless, the dimensional source (color, shape, or both) of this information is irrelevant, as only the strength of the perceived nonuniformity in each display is used to solve the task.

With respect to d' the predictions of the unidimensional model agree with the predictions of the integration model, because both assume that color and shape information simply summate on redundant trials. Thus, for exactly the same reasons, both models predict that $d'_{redundant}$ should equal the sum of d'_{color} and d'_{shape}. Because one model assumes two channels and the other assumes only one, however, their predictions with respect to the variance of the redundant distribution differ. More specifically, the unidimensional model does not necessarily predict that the variance will increase relative to color and shape trials. Because only one channel is assumed, adding color and shape information together is conceptually the same as increasing the strength of a color difference or increasing the strength of a shape difference. That is, it should not matter whether performance is enhanced by magnifying a color difference or by adding a color difference to a shape difference. If only one channel is involved, these are equivalent manipulations.

One might imagine that magnifying a color difference (e.g., switching from a low-contrast color to a high-contrast color) would increase the variance of the signal distribution. If it did, then the same result should have occurred on redundant trials. However, variations in strength of color, although having a large effect on d', do not affect the variance of the signal distribution. We were able to test this explicitly because some color–background combinations used in this experiment yielded poor performance (with d' values less than 1) whereas other color–background combinations yielded excellent performance (with d' close to 3). In both cases, ROC analyses suggested that the standard deviation of the signal distribution was about 3. The same result was obtained when strong and weak shape differences were examined. That being the case, adding the effects of color and shape differences together in a single channel should also have no effect on the variance estimate. These are essentially the results shown in Tables 3 and 5. The mean MLE variance estimate for the two elemental displays was 2.98, and the value for the compound redundant displays was 2.45. The corresponding least squares estimates were 3.32 and 3.76, respectively.

General Discussion

The results of this experiment merit several conclusions. First, the weight of evidence more strongly favors a signal detection account of the pigeons' choice behavior in this task than the HTT–default response account. Interpreted within this signal detection framework, we also established that the variance of the distribution of strength of evidence associated with the Different displays was substantially greater than that of the Same displays ($\sigma^2_{Same} < \sigma^2_{Different}$). Further, the variance estimates associated with color, shape, and redundant signal distributions were, within the limits of

measurement error, probably equivalent ($\sigma^2_{color} = \sigma^2_{redundant} \approx \sigma^2_{shape}$). Finally, the results suggest that performance with redundant displays was equal to the additive combination of discriminative performance on color and shape trials ($d'_{redundant} = d'_{color} + d'_{shape}$). Together, these observations best fit the predictions of a unidimensional decision model—one suggesting that these pigeons were making their discriminations on the basis of the extraction of a single type of information from the different types of textured stimuli. Figure 7 visually depicts these conclusions and their structural implications for the decision space used by the pigeons while performing this textured same–different task.

High-Threshold Versus Signal Detection Models of Discriminations

Historically, with human participants in a wide variety of tasks, the high-threshold model has failed to account for the shape of ROC plots. Almost invariably, the deviations with human participants are like those observed here with the pigeons, showing systematic differences in the directions predicted by a signal detection account. The present results add to a reasonably large literature suggesting that SDT nicely accounts for the discriminative performance of pigeons, and to the much smaller literature suggesting that it does a better job than its most frequently proposed alternative. In conjunction with Wixted's (1993) results based on a sample–no sample matching task, the present results involving target–no target stimuli in a same–different task offer little support for the conceptually appealing HTT–default response model as an adequate account of pigeon choice behavior. Instead for these two different types of discriminations, where a default response strategy was a distinct possibility because of the asymmetrical nature of the to-be-discriminated stimuli, a signal detection account offered a better description of the underlying decision processes.

Although little support for HTT was found in the present experiment, there remain several other situations in which asymmetrical stimuli have been used or asymmetrical responding has been observed where such an account remains plausible, including the memory of event duration by pigeons (Spetch & Wilkie, 1983) and the memory for sample modality by rats (Wallace, Steinert, Scobie, & Spear, 1980). It would be quite interesting to examine these types of discrimination by using the biasing protocol used here to see how well HTT and SDT would fare in these settings.

In fairness, it should be noted that the high-threshold model has only one free parameter (p) whereas the signal detection model tested here has two free parameters (d' and σ). Although the signal detection model was penalized by the loss of one additional degree of freedom in the chi-square goodness-of-fit tests, for example, the comparison of two models involving different numbers of parameters is not always a straightforward exercise. A more elaborate modification of the high-threshold account, such as Luce's low-threshold model (1963, which similarly has two free parameters) might have described these data accurately. We did not
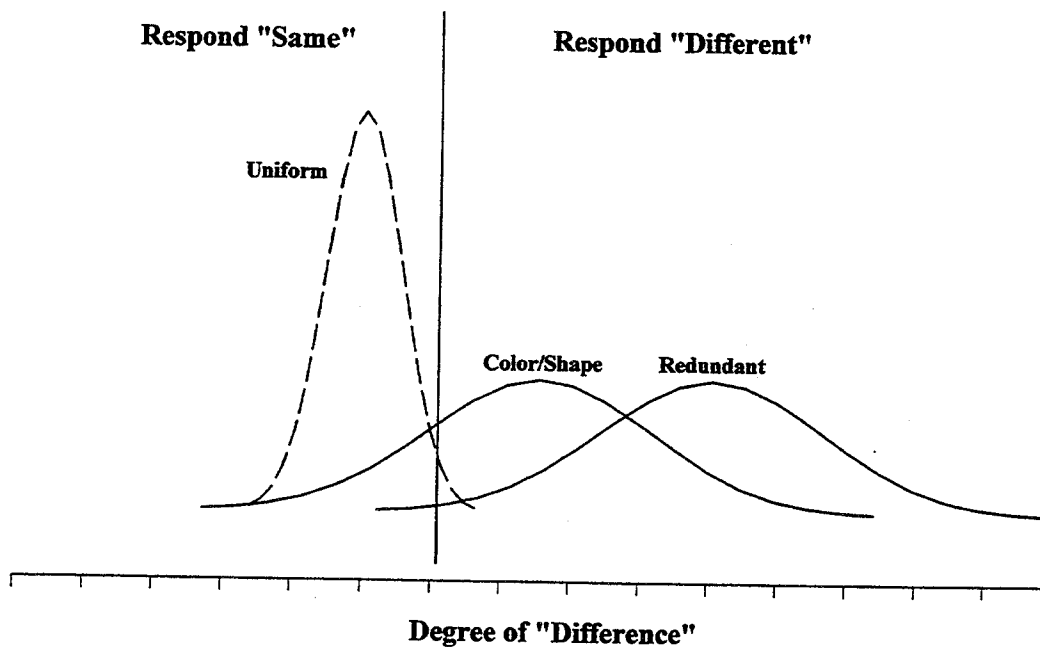
Figure 7. Illustration of the final parameters for the best fitting signal detection account of the present experiment.

pursue this latter model in greater detail for several reasons. First, we chose to concentrate on the one-parameter HTT because of its direct relation to currently proposed mechanisms of pigeon choice behavior. Second, it has been historically quite difficult to distinguish Luce's model from SDT on the basis of quantitative fits. Lastly, it should be noted that SDT assumes that pigeons always make their choice only on the basis of subjective information from the evidence dimension. Although this assumption is unlikely to be true in the extreme (Blough, 1996), as long as the pigeons' rates of guessing were constant across the different biasing conditions, this consideration would not change the conclusions reached in this article.

Difference in Signal and Noise Variance

Although an SDT account was found to better describe the results, it did require estimating signal distributions for color, shape, and redundant Different trials with standard deviations nearly a factor of three greater than that of the Same–noise distribution. This result was not anticipated, although given the asymmetric nature of the stimuli in terms of structure and number, it is perhaps not surprising. Wixted (1993) also found unequal variances in his sample–no sample matching task, although in that particular case it was the noise distribution that was wider than the signal distribution. Unlike a simple red versus green discrimination task, which involves stimuli that lie on a single physical continuum, both the present experiment and Wixted's may have involved discriminations among qualitatively distinct stimuli that may have been responsible for this difference in the variance of the signal and noise distributions. Both discriminations involved in some sense the discrimination of the

presence and absence of a particular type of stimulation. The wider variance of the noise distribution in Wixted's sample–no sample discrimination would seem to make good sense, as the absence of an event seems inherently more difficult to encode than its physical presence. The current experiment required the pigeons to discriminate a uniform display from a Different display that contained the additional presence of a highly variable target. As such, the greater width of the signal distribution in the latter case also seems to make intuitive sense.

But what factors associated with the Different displays were specifically responsible for this effect? We investigated several possibilities, all of which we eventually rejected as the source of this variability. The first possibility we considered is that the assortment of targets used in this experiment varied in their discriminability. Some targets (e.g., red dots in brown dots) were much less discriminable than others (e.g., yellow triangles in blue triangles). Perhaps the mixture of salient and nonsalient target stimuli produced more variability than the various types of uniform stimuli (such as all red dots or all yellow triangles). Although this intuitively appealing explanation can actually be rejected analytically, we nevertheless identified and classified the target–distractor combinations that yielded the best performance and separated them from those that yielded the worst performance (see also Cook, Katz, & Cavoto, in press). As indicated earlier, separate ROC analyses of these more homogeneous discriminability groupings of the Different trials yielded the same higher estimates of σ for all levels. That is, the ratio of the signal to noise σ was approximately 3 in all of these cases (although d' for the discriminable stimuli was, of course, much greater than d' for the less discriminable stimuli).

Next we tried classifying the Different trials according to whether the pigeon pecked the target area (which generally produced a higher d') from those in which they did not peck at the target (which yielded lower d' estimates). Separate analyses based on these differences in pecking behavior also produced no change in the estimates for signal distribution variance from those analyses described above.

Another possibility is that the relatively wide signal distribution variance may have resulted from the variable nature of the target's spatial position from trial to trial. Sometimes the target was on the left side of the screen and sometimes it was on the right. When the target appeared on the left, subjects may have been biased to choose the left hopper, and when it appeared on the right they may have been biased to choose the right hopper. To evaluate this possibility, we again reclassified the Different trials on the basis of target location in two ways (left vs. right and center vs. periphery). Again we fit the models to these more homogeneous location data sets, and in both cases the estimates of the signal distribution's variance remained stubbornly higher.

These analyses based on target discriminability, target pecking behavior, and target location all failed to reveal the reason for the wider estimate of the signal variance. Thus, perhaps our first intuition that the origin of the variance differences lie with the Different displays is incorrect. It is important to remember that these estimates are computed by fixing the value of the noise distribution at 1. As such, the problem can be conceptualized a little differently. Instead of the Different stimuli producing high variability, the Same stimuli may have been producing unexpectedly low variability. Such a perspective would suggest that uniformity may be a distinctive, psychologically less variable attribute, a possibility that has been considered in the human same–different literature to explain such phenomena as "fast same responses" (Krueger, 1978).

Although the psychological source of the unequal variance in the signal and noise distribution remains an intriguing mystery, the fact that they differ at all leads to an important consideration in judging choice accuracy among different conditions. Specifically, any model of how pigeons integrate color and shape information on redundant trials cannot rely solely on accuracy information as derived from a single test condition. Had only a single biasing condition been conducted, the information about the relative widths of the signal and noise distributions could not have been computed. Under such single test conditions, researchers adopt by necessity the simplifying assumption that the two distributions have the same standard deviations (e.g., Mulligan & Shaw, 1980). Had that strategy been used here, incorrect conclusions would have been reached. For example, we could have relied on hit and false alarm rates from the unbiased (i.e., equal presentation probability) condition only. By adopting the assumption that the variances of the signal and noise distribution were equal, we could easily obtain d' scores from published tables. Using that approach, we determined the mean d' scores from color, shape, and redundant trials in the unbiased condition to be 1.90, 1.91, and 2.91 respectively. The implication of these values is that

performance in the redundant condition is significantly less than the additive combination of performance in the color and shape conditions, a result that would have led us to reject incorrectly the unidimensional model.

In conclusion, the present research demonstrated that a signal detection analysis of ROC data provides an efficient way to differentially evaluate models of multidimensional perception and choice in pigeons. By using this approach, we rejected two multiple-channel models in favor of a simpler unidimensional account. Because the pigeon's job was primarily to detect the presence of a generalized target composed from a variety of values from different dimensions and their combination, perhaps this type of common encoding of the Different displays is not surprising. Such a general coding scheme for the display would permit the pigeons to both work with large numbers of stimuli and transfer to novel exemplars as found by Cook, Cavoto, et al. (1995; see also Wasserman et al., 1995). What is the possible psychological identity of this proposed dimension of discrimination? One possibility is that the pigeons learned a conceptual same–different rule, which was then used to judge all of the displays. For reasons outlined in those articles, both Cook, Cavoto, et al. (1995) and Wasserman et al. (1995) put forth this idea as the best explanation of their novel stimulus transfer results. Nevertheless, the present results are also consistent with another unidimensional possibility considered, but rejected, by these same authors. The idea here is that the pigeons are simply discriminating the training and transfer displays on the basis of a lower level, perceptually driven *contrast* detection strategy rather than using a higher level conceptual strategy. Recently, Young and Wasserman (1997; see also Young, Wasserman, & Garner, 1997) offered a third unidimensional possibility, reporting evidence for the idea that the critical dimension may involve the encoding of a display's *entropy*. Entropy is an information–theoretic measure of the relative variability created by the number and nature of different elements in multielement mixed displays. Although the present experiment cannot distinguish among these three alternative conceptions of the critical dimension (but see Cook, Katz, & Cavoto, 1997), its results do clearly support their pivotal idea that only a single kind of information is involved in the discrimination of these types of multielement same–different displays by pigeons.

References

Abramowitz, M., & Stegun, I. A. (1965). *Handbook of mathematical functions.* New York: Dover.

Blackwell, H. R. (1963). Neural theories of simple visual discriminations. *Journal of the Optical Society of America, 53,* 129–160.

Blough, D. S. (1996). Error factors in pigeon discrimination and delayed matching. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 118–131.

Colwill, R. M. (1984). Disruption of short-term memory for reinforcement by ambient illumination. *Quarterly Journal of Experimental Psychology, 36,* 235–258.

Colwill, R. M., & Dickinson, A. (1980). Short-term retention of "surprising" events by pigeons. *Quarterly Journal of Experimental Psychology, 32,* 539–556.

Cook, R. G. (1992). Dimensional organization and texture discrimi-

nation in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 18,* 354–363.

Cook, R. G., Blondeau, G., Cavoto, B. R., & Katz, J. (1995). [When and where in multidimensional texture discrimination by pigeons]. Unpublished data.

Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1995). Same–different texture discrimination and concept learning in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 253–260.

Cook, R. G., Cavoto, K. K., & Cavoto, B. R. (1996). Mechanisms of multidimensional grouping, fusion, and search in avian texture discrimination. *Animal Learning & Behavior, 24,* 150–167.

Cook, R. G., Katz, J. S., & Cavoto, B. R. (1997). Pigeon same–different concept learning with multiple stimulus classes. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 417–433.

Cook, R. G., Katz, J. S., & Cavoto, B. R. (in press). Processes of visual cognition in the pigeon. In S. A. Soraci & W. McIlvane (Eds.), *Perspectives in fundamental processes in intellectual functioning, Volume 1: A survey of research approaches.* Greenwich, CT: Ablex.

Cook, R. G., Riley, D. A., & Brown, M. F. (1992). Spatial and configural factors in compound stimulus processing by pigeons. *Animal Learning & Behavior, 20,* 41–55.

Dougherty, D. H., & Wixted, J. T. (1996). Detecting a nonevent: Delayed presence-vs.-absence discrimination in pigeons. *Journal of the Experimental Analysis of Behavior, 65,* 81–92.

Grant, D. S. (1991). Symmetrical and asymmetrical coding of food and no-food samples in delayed matching in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 186–193.

Heinemann, E. G., & Chase, S. (1970). Conditional stimulus control. *Journal of Experimental Psychology, 84,* 187–197.

Kehoe, E. J., & Gormezano, I. (1980). Configuration and combination laws in conditioning with compound stimuli. *Psychological Bulletin, 87,* 351–378.

Krueger, L. E. (1978). A theory of perceptual matching. *Psychological Review, 85,* 278–304.

Link, S. W. (1992). *The wave theory of difference and similarity.* Hillsdale, NJ: Erlbaum.

Luce, R. D. (1963). A threshold theory for simple detection experiments. *Psychological Review, 70,* 61–79.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Mulligan, R., & Shaw, M. L. (1980). Multimodal signal detection: Independent decisions vs. integration. *Perception & Psychophysics, 28,* 471–478.

Ogilvie, J. C., & Creelman, C. D. (1968). Maximum-likelihood estimation of receiver operating characteristic curve parameters. *Journal of Mathematical Psychology, 5,* 377–391.

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Institute of Radio Engineers Transaction, PGIT-4,* 171–212.

QuickBasic 7.0 [Computer programming language]. (1989). Redmond, WA: Microsoft Corporation.

Riley, D. A. (1984). Do pigeons decompose stimulus compounds? In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 333–352) Hillsdale, NJ: Erlbaum.

Riley, D. A., Cook, R. G., & Lamb, M. R. (1981). Classification and analysis of short term retention codes in pigeons. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory, Volume 15* (pp. 51–79) New York: Academic Press.

Shaw, M. L. (1982). Attending to multiple sources of information I. The integration of information in decision making. *Cognitive Psychology, 14,* 353–409.

Sherburne, L. M., & Zentall, T. R. (1993a). Asymmetrical coding of food and no-food events by pigeons: Sample pecking versus food as the basis of the sample code. *Learning and Motivation, 24,* 141–155.

Sherburne, L. M., & Zentall, T. R. (1993b). Coding of feature and no-feature events by pigeons performing a delayed conditional discrimination. *Animal Learning & Behavior, 21,* 92–100.

Spetch, M. L., & Wilkie, D. M. (1983). Subjective shortening: A model of pigeons' memory for event duration. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 14–30.

Swets, J. A., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review, 68,* 301–340.

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 51,* 401–409.

Wallace, J., Steinert, P. A., Scobie, S. R., & Spear, N. E. (1980). Stimulus modality and short-term memory in rats. *Animal Learning & Behavior, 8,* 10–16.

Wasserman, E. A., Hugart, J. A., & Kirkpatrick-Steger, K. (1995). Pigeons show same–different conceptualization after training with complex visual stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 248–252.

Wilson, B., & Boakes, R. A. (1985). A comparison of the short-term memory performance of pigeons and jackdaws. *Animal Learning & Behavior, 13,* 285–290.

Wixted, J. T. (1993). A signal detection analysis of memory for nonoccurrence in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 400–411.

Xgraf 5.0 [Computer software]. (1989). Pittsburgh, PA: Komputerwerk Inc.

Young, M. E., & Wasserman, E. A. (1997). Entropy detection by pigeons: Response to mixed visual displays after same–different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 157–170.

Young, M. E., Wasserman, E. A., & Garner, K. L. (1997). Effects of number of items on the pigeon's discrimination of same from different visual displays. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 491–501.