Cognitive theories as reinforcement history surrogates: The case of likelihood ratio models of human recognition memory

JOHN T. WIXTED and SANTINO C. GAITAN
University of California, San Diego, La Jolla, California

B. F. Skinner (1977) once argued that cognitive theories are essentially surrogates for the organism's (usually unknown) reinforcement history. In this article, we argue that this notion applies rather directly to a class of likelihood ratio models of human recognition memory. The point is not that such models are fundamentally flawed or that they are not useful and should be abandoned. Instead, the point is that the role of reinforcement history in shaping memory decisions could help to explain what otherwise must be explained by assuming that subjects are inexplicably endowed with the relevant distributional information and computational abilities. To the degree that a role for an organism's reinforcement history is appreciated, the importance of animal memory research in understanding human memory comes into clearer focus. As Skinner was also fond of pointing out, it is only in the animal laboratory that an organism's history of reinforcement can be precisely controlled and its effects on behavior clearly understood.

In a well-known article entitled "Why I Am Not a Cognitive Psychologist," B. F. Skinner (1977) made the following points:

> The variables of which human behavior is a function lie in the environment. (p. 1)

> Cognitive psychologists study these relations between organism and environment, but they seldom deal with them directly. Instead, they invent internal surrogates which become the subject matter of their science. (p. 1)

> The mental apparatus studied by cognitive psychology is simply a rather crude version of contingencies of reinforcement and their effects. (p. 9)

And in his well-known book *Science and Human Behavior* (1953), Skinner also asserted that

> the practice of looking inside the organism for an explanation of behavior has tended to obscure the variables which are immediately available for a scientific analysis. These variables lie outside the organism, in its immediate environment and in its environmental history. (p. 31)

Implicit in these quotations (and rather explicit in some of Skinner's other writings) is an indictment of cognitive theories and the research they generate. According to Skinner, the status of cognitive theories as reinforcement history surrogates and the dubious value of such theories

for the understanding of behavior are tightly connected notions. However, to us, the two notions seem to be quite distinct. The utility of cognitive theories seems well established, and the research generated by them has revealed interesting and nonobvious human capabilities (some of which we will review in this article). Nevertheless, it might still be true that cognitive models often serve as reinforcement history surrogates, thereby obscuring the environmental variables that are the more distal and potentially observable and controllable causes of behavior. In that sense, cognitive theories, useful though they may be, are open to criticism.

Skinner (1977) developed the cognitive-theories-as-surrogates idea by considering very general explanatory constructs that show up in one way or another in many different theories. Thus, for example, explanations appealing to an organism's "knowledge" or "representations," ideas that form the foundation of many cognitive models, were criticized as offering no real understanding and as drawing attention away from the true causes of behavior. Criticizing the very foundation of cognitive theories renders the argument widely applicable while, perhaps, simultaneously (and somewhat paradoxically) limiting its impact. Unless the argument is addressed to specific theories that have been advanced to explain specific findings in the literature, it may come across as philosophical speculation that is of interest mainly to philosophers. That the argument does apply to specific theories is what we will try to show in this article, and our case in point will be signal-detection–based likelihood ratio models of human recognition memory. Several comprehensive models of human recognition memory have been proposed in recent years that are based on the idea that the memory system com-

putes a likelihood ratio in order to determine whether or not an item was recently encountered on a list. These models assume that the understanding of human recognition memory is contingent on the understanding of the underlying computational machinery. Our central claim is that Skinner may have been right about this theoretical machinery: It is largely a stand-in for the organism's learning history. To the extent that this point is appreciated, the relevance of animal memory research to the understanding of human memory becomes much clearer than it otherwise might be. Indeed, it is only in the animal laboratory that the connection between an organism's prior reinforcement history and its current memory decisions is clearly evident (Alsop, 1998). To quote Skinner (1953) again: "We study the behavior of animals because it is simpler. Basic processes are revealed more easily and can be recorded over longer periods of time....We may arrange genetic histories to control certain variables and special life histories to control others" (p. 38). Illustrating the validity of this point is the main purpose of the present article.

The likelihood ratio models that we will consider in making our case represent one form of signal detection theory, so we will begin our inquiry into this issue with a brief overview of that basic framework.

Signal Detection Theory

Since the late 1960s, theories of human recognition memory have involved concepts drawn from signal detection theory. In a typical recognition memory task, subjects first study a list of items (e.g., words or pictures) and are then presented with a recognition test in which items are presented one at a time for an old/new recognition decision. A response of *old* means that the subject believes that the word appeared on the list; *new* means that the subject believes that it did not. Typically, half of the test items actually did appear on the list (targets), and half did not (lures). Imagine that subjects correctly respond *old* to 84% of the targets (i.e., the hit rate is .84) and *incorrectly* respond *old* to 16% of the lures (i.e., the false alarm rate is .16). How should we conceptualize a performance like this? Signal detection theory offers one appealing answer. This theory comes in two basic versions, one of which is fairly simple and the other of which is more complex.

In its simplest form, signal detection theory holds that a decision about whether or not an item was recently encountered on a list depends on its level of familiarity. If the item's level of familiarity exceeds a criterion value, it is judged to be old; otherwise, it is judged to be new. As is illustrated in the upper panel of Figure 1, familiarity values associated with the old and the new items (i.e., the targets and the lures, respectively) are usually assumed to be normally distributed, with the mean of the target distribution located at a higher point on the familiarity axis than the mean of the lure distribution. The decision criterion specifies the familiarity value above which an item is declared to be old, and it can be placed anywhere along the familiarity axis. In Figure 1, it is placed exactly halfway between the means of the target and the lure distributions. The shaded area in Figure 1 represents the proportion of
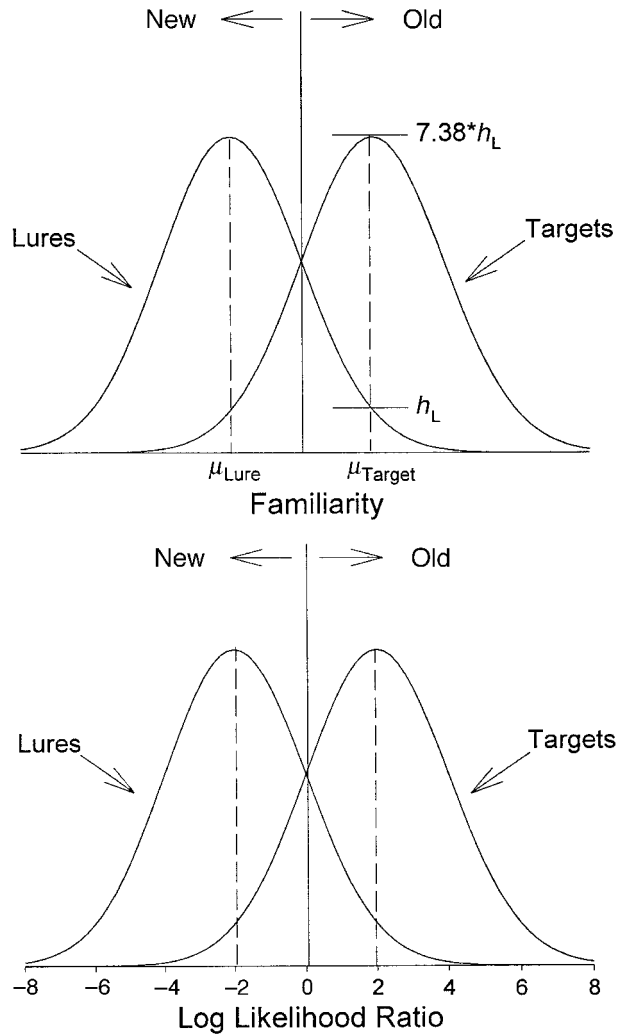

Figure 1. Upper panel: standard familiarity-based signal detection model of recognition memory. The decision axis represents a familiarity scale that ranges from low to high. Lower panel: likelihood ratio version of the signal detection account of recognition memory. The decision axis represents a log likelihood ratio scale that ranges from minus infinity to plus infinity.

lures that are incorrectly judged to be old because they are associated with familiarity values that fall above the criterion. That proportion is known as the false alarm rate. The false alarm rate for the situation depicted in Figure 1 would be about 16% (whereas the hit rate would be 84%).

An important assumption of the detection theory just described is that the decision axis represents a strength-of-evidence variable, such as familiarity. Although many signal detection models of recognition memory make precisely that assumption, another class of detection models does not. These models assume that the decision axis represents a log likelihood ratio scale. According to a likelihood ratio model, recognition decisions are based on a statistical computation: If the computed odds that the item appeared on the list are high enough (usually greater than even), the item is declared to be old; otherwise, it is de-

clared to be new. The lower panel of Figure 1 illustrates the likelihood ratio model, and it is immediately apparent that it looks a lot like the familiarity model shown in the upper panel of Figure 1. The only real difference is the decision axis, which now represents a log likelihood scale.

Likelihood ratio models assume that the operative psychological variable is an odds ratio associated with the test item, not its level of familiarity. The odds ratio is equal to the likelihood that the item was drawn from the target distribution divided by the likelihood that it was drawn from the lure distribution. Graphically, this value is given by the height of the target distribution divided by the height of the lure distribution at the point on the familiarity axis where the item falls. Consider, for example, an item that generates a level of familiarity equal to the mean familiarity of the targets. The height of the target distribution at its mean is 7.38 times the height of the lure distribution (h_L). As such, a test item that generates this level of familiarity (i.e., a familiarity of μ_{Target}) is 7.38 times as likely to have been drawn from the target distribution as from the lure distribution. Whenever the odds are greater than even, as they are in this case, an unbiased subject would declare the item to be old. Note that the natural log of 7.38 is 2.0, so the same item that generates a level of familiarity falling at the mean of the target distribution in the upper panel of Figure 1 generates a log likelihood ratio of 2.0 in the lower panel of Figure 1. Similarly, an item whose familiarity value falls midway between the means of the target and the lure distributions is associated with a likelihood ratio of 1.0 (i.e., the heights of the two distributions are equal at that point), which translates to a log likelihood ratio of 0 in the lower panel of Figure 1.

Whereas the strength version of signal detection theory requires only that the subject have some criterion familiarity value in mind in order to arrive at a recognition decision, the likelihood ratio version requires much more complete knowledge of the shapes and locations of the underlying familiarity distributions in order to make the relevant statistical computations. Although those added assumptions may seem implausible, it seems clear that subjects sometimes behave as if they do possess the relevant knowledge. The evidence bearing on this claim will be reviewed next, and the question we will later ask is this: Are humans biologically endowed with the requisite knowledge and computational ability, or have they been trained to behave that way as a function of prior (extralaboratory) consequences?

Explaining the Mirror Effect

The old/new decision criterion. What is the advantage of assuming a likelihood ratio model of memory? The familiarity and likelihood ratio accounts are equally able to explain, say, a hit rate of .84 and a false alarm rate of .16 (as is shown in Figure 1), but Figure 2 illustrates an important difference between them. This figure shows what the expected result would be if subjects maintained the same decision criterion as d' (the standardized distance between the means of the target and the lure distributions) changed from low to high. The upper panel shows the

familiarity-based model, and the lower panel shows the likelihood ratio model. We assume that d' was affected by a standard strength manipulation (e.g., in the strong condition, subjects had more time to study the items than in the weak condition) and that this manipulation affected the mean familiarity of the target items without affecting the mean familiarity of the lures (which, in both conditions, are simply items that did not appear on the list).

Note that the definition of the decision criterion differs between the two accounts. In the strength version, the criterion is a particular level of familiarity (so that items yielding higher familiarity levels than that are judged to be old). In the likelihood ratio version, the criterion is a particular odds ratio (so that items yielding higher odds than that are judged to be old).

In the familiarity-based model shown in the upper panel of Figure 2, the criterion is placed 0.75 standard deviations above the mean of the lure distribution whether the targets are weak or strong (i.e., the criterion is fixed at a particular point on the familiarity axis). As such, the hit rate increases considerably as a function of strength, but the false alarm rate remains constant. If the familiarity of the lures does not change as the targets are strengthened (which is the simplest assumption), the only way that the false alarm rate would change is if the criterion moved across conditions.

In the likelihood ratio model, the criterion is placed at the point where the odds are even that the item was drawn from the target or the lure distributions (i.e., where the odds ratio is 1.0). The point of even odds occurs where the heights of the two distributions are equal, and that occurs where the distributions intersect. In the familiarity model, the point of intersection occurs at one level of familiarity in the weak case and at a higher level of familiarity in the strong case. But these two familiarity values both translate to a value of 0 on the log likelihood ratio axis (i.e., $\ln[1.0] = 0$). Thus, from a likelihood ratio point of view, the criterion remains fixed at 0 across the two strength conditions.

Note the different patterns of hit and false alarm rates yielded by these two fixed-criterion models. Unlike the familiarity model, the likelihood ratio model naturally predicts a mirror effect (i.e., as the hit rate goes up, the false alarm rate goes down), and that pattern happens to be a nearly universal finding in the recognition memory literature (Glanzer, Adams, Iverson, & Kim, 1993). The mirror effect is typically observed not only for strength manipulations, but also for manipulations of word frequency, concreteness, and a variety of other variables (Glanzer & Adams, 1990). The ability of the likelihood ratio model to account for the mirror effect in a natural way is one of its most attractive features. Indeed, largely because of the mirror effect, recent theorizing about recognition memory appears to reflect a shift away from the idea that the decision axis represents a strength-of-evidence scale (like familiarity) and toward the idea that it represents a likelihood ratio scale. In fact, three of the most recent theories of human recognition memory are based fundamentally on the notion that recognition memory decisions are
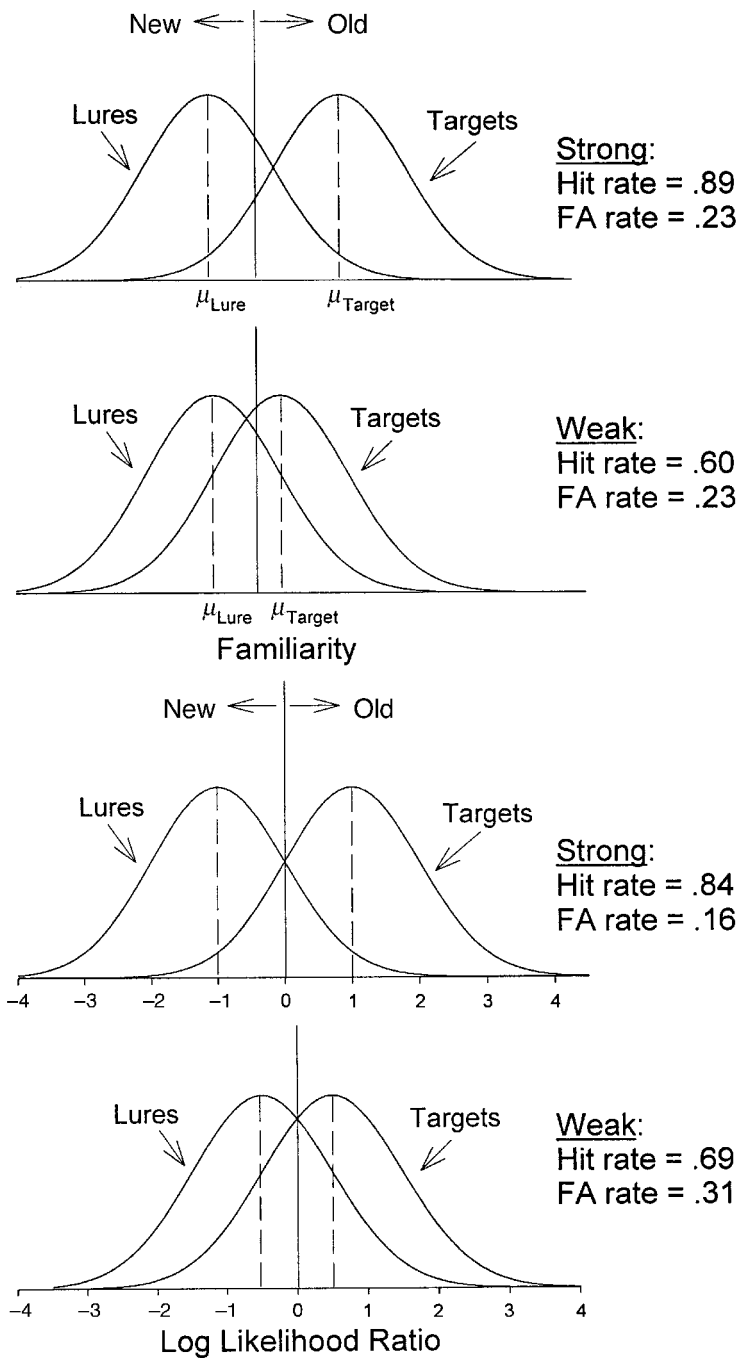
Figure 2. Upper panel: familiarity-based signal detection model of weak and strong conditions over which the decision criterion remains fixed on the decision axis. Lower panel: likelihood ratio signal detection model of weak and strong conditions over which the decision criterion remains fixed on the decision axis.

based on a likelihood ratio computation. These include theories known as attention-likelihood theory (Glanzer et al., 1993), retrieving effectively from memory (Shiffrin & Steyvers, 1997), and subjective-likelihood theory (McClelland & Chappell, 1998).

Although the very existence of the mirror effect supports the likelihood ratio view, that effect can still be eas-

ily represented (and explained) with the simpler version that assumes a familiarity axis. Thus, for example, imagine that, in one condition, the hit rate is .84 and the false alarm rate is .16 (so that $d' = 2$) and, in another condition, the hit rate is .69 and the false alarm rate is .31 (so that $d' = 1$). Further imagine that the mean familiarity of the lure distribution is the same in both conditions but

the mean familiarity of the targets is greater in the strong condition than in the weak condition. Figure 3 shows how to represent this situation, assuming a familiarity axis. Basically, when the target distribution weakens, the criterion shifts to the left in such a way as to remain midway between the means of the target and the lure distributions. Subjects might very well be inclined to do this because, after a strong list, they would presumably know that a test item, had it appeared on the list, would be very familiar, so a stringent criterion could be used without missing too many of the targets. After a weak list, by contrast, they would know that a target item might be somewhat unfamiliar even if it had appeared on the list, so a less stringent criterion might be used so as not to miss too many of the targets. If the criterion did shift in this way across conditions (which would not involve a likelihood ratio computation), a mirror effect would be observed. Nevertheless, a likelihood ratio theorist would say that the apparent criterion shift is an illusion: What looks like a criterion shift on the familiarity axis actually reflects the fact that the subject uses the same criterion odds ratio in both cases.

Confidence criteria. The argument just presented in connection with the movement of the old/new decision criterion as a function of strength can be generalized to the

situation in which subjects are asked not only for an old/new judgment, but also for a confidence rating. Thus, for example, when asked to decide whether or not a test item appeared on a study list, the subject might be offered six choice alternatives, labeled $---$, $--$, $-$, $+$, $++$, $+++$, where the first three represent *new* responses of varying confidence and the last three represent *old* responses of varying confidence. If the subject is absolutely sure that a particular test item appeared on the list, the "$+++$" button would be pressed. A less certain *old* response would be indicated by pressing the "$++$" button instead.

A straightforward extension of the basic detection model holds that confidence judgments are reached in much the same way that old/new decisions are reached (e.g., Macmillan & Creelman, 1991). That is, as is illustrated in Figure 4, a different criterion for each confidence rating is theoretically placed somewhere along the familiarity axis. Familiarity values that fall above O_H receive an *old* response with high confidence (i.e., $+++$), those that fall between O_M and O_H receive an *old* response with medium confidence ($++$), and those that fall between C and O_M receive an *old* response with low confidence ($+$).[1] On the other side of the criterion, familiarity values that fall below C but above N_M receive a *new* response with low confidence ($-$), those that fall below N_M but above N_H receive a *new* response with medium confidence ($--$), and those that fall below N_H receive a *new* response with high confidence ($---$).

As was indicated earlier, the likelihood ratio model holds that the mirror effect occurs because subjects use the same likelihood ratio to make decisions in both the strong and the weak conditions. A fixed likelihood ratio criterion shows up as a shifting criterion as a function of strength when the situation is represented by the familiarity version of detection theory (as in Figure 3). A similar story applies to the way in which confidence criteria change as a function of strength. If subjects maintain constant likelihood ratios for each of the confidence criteria, then, when represented in terms of the familiarity version of detection theory, the confidence criteria should move in a predictable way as d' decreases. More specifically, as will be described next, the criteria should diverge, or fan out, as d' approaches zero. Figure 5 shows three different ways the confidence criteria might shift as d' changes from high to low. In all three cases, the old/new decision criterion shifts to the left (thereby yielding a mirror effect), but the confidence criteria shift in different ways. The upper panel shows the confidence criteria shifting in lockstep with the decision criterion, the middle panel shows the confidence criteria converging toward the old/new decision criterion, and the lower panel shows the confidence criteria diverging (as the likelihood ratio account uniquely predicts).

According to the likelihood ratio account, the old/new decision criterion, C, is placed in such as way as to maintain a likelihood ratio of 1 for unbiased responding. Similarly, the O_H criterion is placed in such a way as to maintain a larger likelihood ratio, such as 10 to 1. That is, any item that is 10 or more times as likely to have come from
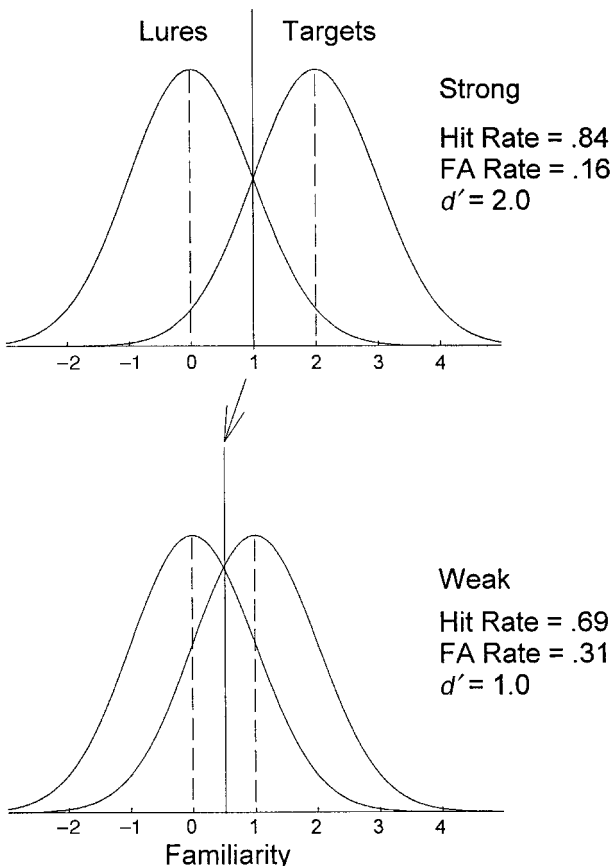

Figure 3. Familiarity-based signal detection model of weak and strong conditions over which the decision shifts on the decision axis.
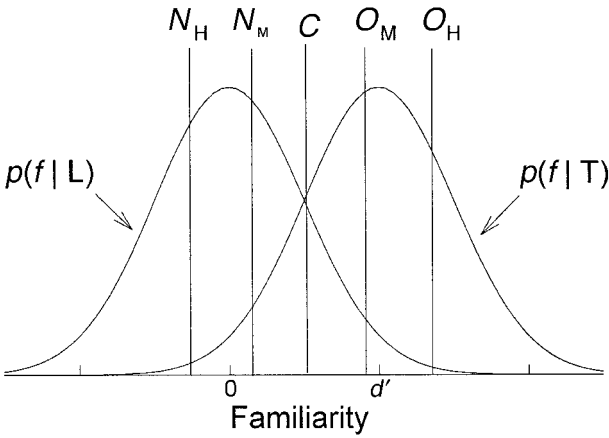
Figure 4. Illustration of the placement of confidence criteria in a signal detection framework. The old/new decision criterion is represented by C, and the remaining confidence criteria are represented by O_M, O_H, N_M, and N_H. Familiarity values falling above O_M or O_H receive medium- or high-confident *old* responses (respectively), whereas familiarity values that fall below N_M or N_H receive medium- or high-confident *new* responses (respectively).

the target distribution than the lure distribution receives an *old* response with high confidence. Different subjects might use different specific odds ratios before making a high-confident response, but we will assume that O_H falls where the likelihood ratio is 10 to 1 for the sake of simplicity. Similarly, the N_H criterion (i.e., the criterion for saying *new* with high confidence) might be placed in such a way as to maintain a small likelihood ratio, such as 1 to 10 (.1). That is, any item with only a 1 in 10 chance (or less) of having been drawn from the target distribution receives a *no* response with high confidence. As d' decreases, this model assumes that the criteria move in such a way as to maintain these confidence-specific likelihood ratios.

How could subjects possibly adjust the criteria in this way? The likelihood ratio model assumes that the subject not only has information about the location of the target and lure distributions, but also has information about their mathematical forms. Where the latter information might come from is usually underspecified (and is the main focus of this article). If the target and lure distributions are Gaussian, with means of d' and 0, respectively, and common standard deviation of σ, the probability of encountering a test item with a familiarity of f given that it is a target, $p(f \mid \text{target})$, is

$$p(f \mid \text{target}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-d')^2/\sigma^2} \,, \qquad (1)$$

and the probability of encountering a test item with a familiarity of f given that it is a lure, $p(f \mid \text{lure})$, is

$$p(f \mid \text{lure}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-0)^2/\sigma^2} \,, \qquad (2)$$

where f is a value along the familiarity axis. The mean of the lure distribution is arbitrarily set to 0 for the sake of simplicity. All other values (e.g., the mean of the target

distribution and the location of the decision criterion) are measured with respect to the mean of the lure distribution. The mean of the target distribution, d', represents the number of standard deviations the mean of that distribution falls above the mean of the lure distribution. Indeed, whether we are considering the mean of the target distribution or the location of the decision and confidence criteria, the values represent the number of standard deviation units above (or below) the mean of the lure distribution.

According to likelihood ratio models of human recognition memory, subjects use their knowledge of the mathematical forms of the target and lure distributions to place the old/new decision criterion, C, at the point on the familiarity axis where the height of the signal distribution equals that of the noise distribution—that is, where the likelihood ratio, $p(f \mid \text{target})/p(f \mid \text{lure})$, is equal to 1. This is the optimal location of the decision criterion, because this is where the odds that the test item was drawn from the target distribution exactly equals the odds that it was drawn from the lure distribution. For any familiarity value greater than that, the odds that the item was drawn from the target distribution (i.e., that it appeared on the list) are better than even, in which case an *old* response makes sense. As was indicated above, the remaining confidence criteria are placed in such a way as to maintain specific odds ratios that are greater than or less than 1. If O_H is associated with a likelihood ratio of 10 to 1, that confidence criterion is placed on the familiarity axis at the point where the height of the target distribution is 10 times the height of the lure distribution regardless of the value of d'.

For any given likelihood ratio, L, the criterion is placed on the familiarity axis at the point f that satisfies the following equation:

$$L(f) = \frac{p(f \mid \text{target})}{p(f \mid \text{lure})} \,. \qquad (3)$$

The right side of the equation is simply the ratio of the height of the target (signal plus noise) distribution to the height of the lure (noise) distribution at the point f on the familiarity axis. That is, using the identities expressed in Equations 1 and 2,

$$L(f) = \frac{\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-d')^2/\sigma^2}}{\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-0)^2/\sigma^2}} \,,$$

which, after setting σ to 1 for the sake of simplicity and rearranging terms, simplifies to

$$L(f) = e^{-.5(f-d')^2 + .5f^2} \,.$$

Solving this equation for f yields

$$f = \frac{d'}{2} + \frac{\ln[L(f)]}{d'} \,. \qquad (4)$$

To determine where a particular criterion should be placed on the familiarity axis to satisfy a particular odds ratio, one need only enter the value of L theoretically as-
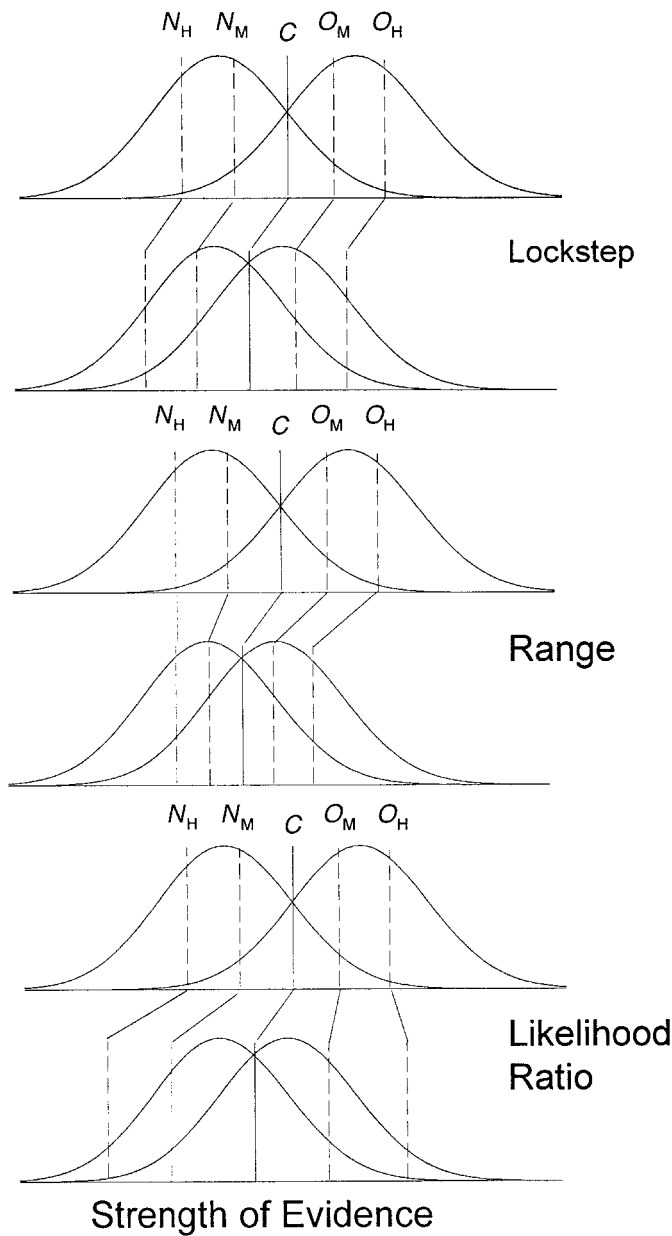
Figure 5. Illustration of the movement of the confidence criteria as a function of d' according to the lockstep, range, and likelihood ratio models (upper, middle, and lower panels, respectively). In each panel, the top figure represents a strong condition, and the bottom figure represents a weak condition.

sociated with that criterion. For example, theoretically, the old/new decision criterion, C, is placed at the point where the odds ratio is 1, which is to say $L(f) = 1$. Substituting 1 for $L(f)$ (and C for f) in Equation 2 reveals that $C = d'/2$. That is, according to this model, the decision criterion should be placed midway between the target and the lure distributions no matter what the value of d' is. If the criterion were always placed in that location in order to always maintain a likelihood ratio of 1, a mirror effect would always be observed (and it almost always is). If O_H is placed at the point on the familiarity axis where the odds

are 10 to 1 in favor of the item's having been drawn from the target distribution, Equation 2 indicates that

$$O_H = \frac{d'}{2} + \frac{\ln(10)}{d'}$$

$$= \frac{d'}{2} + \frac{2.30}{d'}.$$

In other words, O_H is placed at a point higher than $d'/2$ but by an amount that varies with d'. The more general version of this equation is $O_H = d'/2 + k_1/d'$, where k_1 is the constant log likelihood ratio associated with O_H (which is

10 in this example but could differ for different subjects). Similarly, if N_H is placed at the point where the odds are only 1 in 10 in favor of the item's having been drawn from the target distribution, Equation 2 indicates that

$$N_H = \frac{d'}{2} + \frac{\ln(0.10)}{d'}$$
$$= \frac{d'}{2} - \frac{2.30}{d'}.$$

The more general version of this equation is $N_H = d'/2 + k_2/d'$, where k_2 is the constant log likelihood ratio associated with N_H. This model predicts that the distance between O_H and N_H on the familiarity axis will be inversely related to d'. That distance (i.e., O_H minus N_H) is given by

$$O_H - N_H = \left(\frac{d'}{2} + \frac{k_1}{d'}\right) - \left(\frac{d'}{2} + \frac{k_2}{d'}\right)$$
$$= \frac{k_1 - k_2}{d'}.$$

Thus, as d' becomes very large, O_H and N_H should converge to the point that they coincide on the evidence axis (i.e., the distance between them should decrease to 0). By contrast, as d' approaches 0, those two criteria should move infinitely far apart.

The upshot of all of this is that, when the situation is represented with a familiarity scale, likelihood ratio models predict that the decision criterion (C) should move in such a way that the hit rate decreases and the false alarm rate increases as d' decreases (i.e., a mirror effect should be observed), because C will always be placed at $d'/2$ on the familiarity axis, since that is the point where the odds ratio is equal to 1. The more general version of this account is that all of the confidence criteria should diverge as d' decreases (as in the lower panel of Figure 5). Mirror effects are usually observed in real data, and, as will be described in more detail next, Stretch and Wixted (1998a) showed that the confidence criteria move in essentially the manner predicted by the likelihood ratio account as well.

How are the locations of the decision criterion and the various confidence criteria inferred from actual behavioral data? The process is not as mysterious as it might seem. Consider first how d' is computed. For a given hit and false alarm rate, only one arrangement of the target and lure distributions and the criterion is possible, assuming an equal-variance detection model. For example, symmetrical hit and false alarm rates of .84 and .16 yield a d' of 2.0, with a criterion placement of $d'/2$ (i.e., $C = 1.0$ in this example). Standard deviation units are used in both cases. That is, the only way to represent these hit and false alarm rates in an equal-variance signal detection model is for the means of the two distributions to be separated by two standard deviations and for the criterion to be placed one standard deviation above the mean of the lure distribution. No other placement of the criterion would correspond to a false alarm rate of .16 and, given that, no other placement of the mean of the target distribution would yield a hit rate

of .84. In practice, standard equations exist for making the relevant computations:

$$d' = z(\text{HIT}) - z(\text{FA})$$

and

$$C = -z(\text{FA}).$$

Thus, in this example, $C = -z(.16) = -(-1.0) = 1.0$, which is to say that the criterion is placed one standard deviation above the mean of the lure distribution. This method of computing the placement of C involves all *old* responses to lures with a confidence rating of *maybe old* or greater (which is to say, all *old* responses to lures, +, ++, or +++). A very similar method can be used to determine the locations of the various confidence criteria. To determine where the O_M confidence criterion is placed, for example, one can use only those responses to lures that received a confidence rating of *probably old* or greater (i.e., all false alarms committed by choosing the ++ or +++ alternatives). From these responses, a new false alarm rate can be computed, and the location of the O_M criterion can be estimated by converting that false alarm rate into a z score (exactly as one usually does for *old* responses that exceed the old/new decision criterion). In practice, a more comprehensive method is used (one that allows for unequal variance and that takes into account confidence-specific hit rates as well), but the method just described works about as well and is conceptually much simpler (Stretch & Wixted, 1998a). The locations of the various confidence criteria on the familiarity axis are simply the cumulative confidence-specific false alarm rates converted to a z score.

Stretch and Wixted (1998a) tested the predictions of the likelihood ratio account with a very simple experiment. In one condition, subjects studied lists of words that were presented quite slowly, after which they received a recognition test (this was the strong condition). In another, they studied lists of words that were presented more rapidly, followed by a recognition test (the weak condition). For both conditions, the locations of the various confidence criteria were computed in essentially the manner described above. As uniquely predicted by the likelihood ratio account, the confidence criteria fanned out on the familiarity axis as d' decreased. Figure 6 presents the relevant estimates of the locations of the confidence criteria in the strong and the weak conditions. Note that whereas the location of the decision criterion, C, decreased on the familiarity axis as conditions changed from weak to strong, the location of the high-confident *old* criterion, O_H, remained essentially constant, which means that even though the overall false alarm rate increased in the weak condition, the high-confident false alarm rate did not. This phenomenon is uniquely predicted by the likelihood ratio account.

Thus, whether we consider the basic mirror effect or the more complete picture provided by ratings of confidence, subjects behave as if they are able to maintain essentially (although not necessarily *exactly*) constant likelihood ratios across conditions. The fact that subjects are able to do
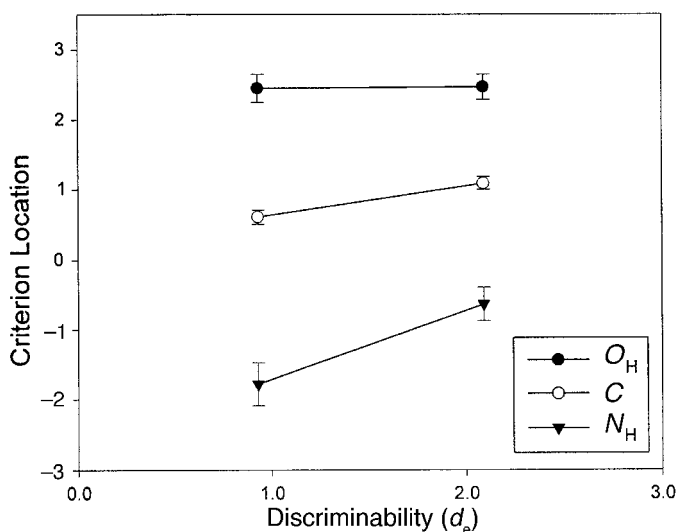
Figure 6. Estimates of the locations of the confidence criteria (O_H, C, and N_H) in the strong and weak conditions of a recognition memory experiment reported by Stretch and Wixted (1998a).

this is rather mysterious, at least on the surface, because it is not clear how they would have acquired the detailed information it requires (and likelihood ratio models typically have little to say about this key issue).

Within-list strength manipulations. It is worth noting that although subjects often behave in accordance with the predictions of likelihood ratio models, there are conditions under which they do not. All this really means is that it is a mistake to think of humans as being fully informed likelihood ratio computers. For the sake of completeness, we will turn now to a brief discussion of the conditions under which subjects do not behave in the manner predicted by likelihood ratio accounts, after which we will return to the central question of how it is that they often do.

Strength is usually manipulated between lists to produce a mirror effect, but Stretch and Wixted (1998b) also manipulated strength *within* list. In each list, half the words were colored red, and they were presented five times each. The other half were colored blue, and they appeared only once each. The red and the blue words were randomly intermixed, and the red word repetitions were randomly scattered throughout the list. On the subsequent recognition test, the red and the blue targets were randomly intermixed with red and blue lures. Obviously, the hit rate in the strong (red) condition was expected to exceed the hit rate in the weak (blue) condition, and it did. Of interest was whether or not the false alarm rate in the strong condition (i.e., to red lures) would be less than the false alarm rate in the weak condition (i.e., to blue lures). If so, the within-list strength manipulation would yield results similar to those produced by the between-list strength manipulation. However, the results reported by Stretch and Wixted (1998b) showed a different pattern. In two experiments in which strength was cued by color, the hit rate to

the strong items was much higher than the hit rate to the weak items (as was expected), but the false alarm rates were nearly identical. Similar results when other list materials were used were also recently reported by Morrell, Gaitan, and Wixted (2002). These results correspond to the situation depicted earlier in the upper panel of Figure 2, according to which the decision criterion remains fixed on the familiarity axis as a function of strength. A fully informed likelihood ratio computer would not respond in this manner but would, instead, adjust the criterion as a function of strength in order to maintain a constant likelihood ratio across conditions (as in the between-list strength manipulation).

The main difference between the within- and the between-list strength manipulations lies in how rapidly the conditions change during testing. In the between-list case, the conditions change slowly (e.g., all the target items are weak after one list, and all are strong after the other). In the within-list case, the conditions change at a rapid rate (e.g., a strong test item might be followed by a weak test item followed by a strong test item, and so on). Under such conditions, subjects tend to aggregate over those conditions, rather than respond in accordance with each condition separately. We shall see later that a similar issue arises in recognition memory experiments involving animal subjects, but the important point, for the time being, is that these data show that organisms do not always behave as if they were fully informed likelihood ratio computers.

Even though subjects do not always behave in exactly the manner one would expect according to a likelihood ratio model, they usually do, and the question of interest concerns how they acquire the information needed to do that. More specifically, what accounts for the fact that, when strength is manipulated between lists, subjects have

information available so that the confidence criteria fan out on the familiarity axis as d' decreases in just the manner predicted by the likelihood ratio account? That is a question left largely unanswered by likelihood ratio accounts, and suggesting an answer to it is the issue we will consider next.

Mirror Effects in Pigeons

What is often overlooked in the human cognitive literature is the possibility that subjects have experienced years of making recognition decisions and expressing various levels of confidence. Presumably, those expressions of confidence have been followed by social consequences that have served to shape their behavior. Conceivably, it is those consequences that shaped behavior to mimic what a history-less likelihood ratio computer might do. We begin our investigation into this possibility by making the point that pigeons show a mirror effect, too. In pigeon research, unlike in human research, recognition memory decisions are followed by experimentally arranged consequences, and the connection between where a pigeon places its decision criterion and the consequences it has received in the past is easier to appreciate. That explicitly arranged reinforcers could influence the location of the decision criterion was recognized from the earliest days of detection theory (e.g., Green & Swets, 1966), but research into that issue never really blossomed in the literature on humans. The most influential account along these lines in the literature on animals, where relevant research did blossom, was proposed some time later by Davison and Tustin (1978). Nevertheless, it is probably fair to say that the connection between their research (and the research their ideas generated) and what a likelihood ratio theorist might say about the movement of confidence criteria across conditions is not easy to see. Making that connection explicit is what we will try to accomplish in this section and the next.

Mirror effects in pigeons can be observed by using a simple procedure in which the birds are required to discriminate the prior presence or absence of a sample stimulus. In a typical experiment of this kind, sample and no-sample trials are randomly intermixed, and the pigeon's task is to report whether or not a sample appeared earlier in the trial. On sample trials, the end of the intertrial interval (ITI) is followed by the presentation of a sample (e.g., the brief illumination of a keylight or the houselight) and then by the retention interval. On no-sample trials, by contrast, the end of the ITI is followed immediately by the onset of the nominal retention interval (i.e., no sample is presented). Following the retention interval, two choice stimuli are presented (e.g., red vs. green). A response to one choice alternative is correct on sample trials, and a response to the other is correct on no-sample trials. Pigeons learn to perform this task with high levels of accuracy when the retention interval is short, and manipulating the retention interval across conditions yields a mirror effect.

In Wixted's (1993) Experiment 2, 4 pigeons were exposed to an ascending series of retention intervals (0.50,

2, 6, 12, and 24 sec), with each retention interval condition in effect for 15 sessions. The results are shown in Table 1. The hit rate represents the probability of a correct response on sample trials (i.e., a correct declaration that the sample occurred), and the false alarm rate represents the probability of an incorrect response (i.e., an incorrect declaration that the sample occurred) on no-sample trials. The data presented in boldface in Table 1 were averaged over the last 5 sessions of each condition and over the 4 subjects, and these data constitute the main results of the experiment. The data shown in standard typeface are from the first session of each new retention interval condition (save for the 0.50 sec condition, because the birds had not yet learned the task in the 1st session of that condition). The data from the 1st session of each condition are instructive and will be discussed in more detail after we consider the main findings. The main results show that, not surprisingly, d' decreased as the retention interval increased. The more important point is that as the hit rate decreased, the false alarm rate reliably increased. In other words, a mirror effect was observed, just as it almost always is in human recognition memory experiments.

What explains the existence of a mirror effect in this case? Figure 7 presents *sense-of-prior-occurrence* distributions for three retention interval conditions (short, medium, and long). These distributions are exactly analogous to the familiarity distributions previously discussed in relation to human recognition memory, and they represent the bird's subjective sense that a sample was presented. That subjective sense is assumed to vary from trial to trial on both sample and no-sample trials (according to Gaussian or Gaussian-like distributions) even for a constant retention interval. Some sense of prior occurrence is evident even on no-sample trials, perhaps owing to the lingering influence of samples presented on earlier trials or to inherent noise in the memory system.

When the retention interval is short (upper panel), the sample distribution falls far to the right because, despite some variability, the subjective sense of prior occurrence

Table 1
Hit and False Alarm (FA) Rates and d' Scores From a Sample/No-Sample Procedure in Which Retention Interval Was Manipulated Between Conditions

Retention Interval (sec)	Hit Rate	FA Rate	d'
0.5	**.94**	**.10**	**2.83**
2.0 (1st)	.84	.08	
2.0	**.88**	**.14**	**2.26**
6.0 (1st)	.60	.14	
6.0	**.88**	**.18**	**2.10**
12.0 (1st)	.67	.19	
12.0	**.74**	**.19**	**1.52**
24.0 (1st)	.46	.24	
24.0	**.65**	**.28**	**.97**

Note—The data were taken from Experiment 2 of Wixted (1993). The values in boldface were averaged over the last five sessions of each condition and over the 4 subjects. The values in standard typeface are from the first session of each new retention interval condition (except for the 0.50 sec condition).
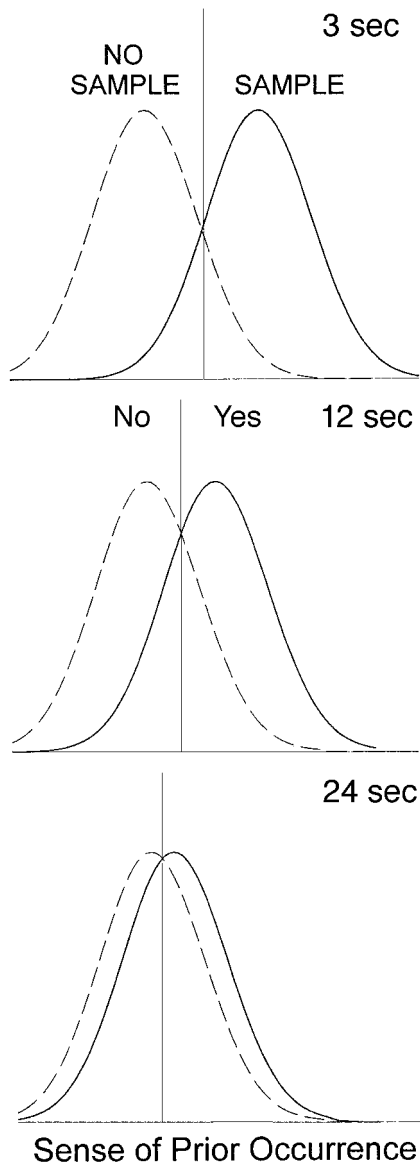
Figure 7. Hypothetical signal and noise distributions (corresponding to sample and no-sample trials, respectively) for three retention intervals (short, medium, and long) manipulated between conditions. The location of the decision criterion, *C*, changes with each condition.

will generally be quite high (and will fall well above the decision criterion). Thus, on the majority of sample trials, the subject would respond correctly (yes). When the retention interval increases (and the sense of prior occurrence weakens), the sample distribution gradually shifts toward the no-sample distribution (middle panel). The no-sample distribution does not change as a function of retention interval, because there is no memory trace established on those trials (so there is no trace that weakens with retention interval). When the retention interval is very

long, so that memory for the sample fades completely, the sample and no-sample distributions will overlap, so that performance on sample and no-sample trials will be indistinguishable (a situation that is approached in the lower panel).

Because each retention interval was in effect for 15 sessions, the birds had time to adjust the location of the decision criterion in light of the reinforcement consequences for doing so. For an equal-variance model, reinforcers will be maximized when the criterion is placed midway between the two distributions. Leaving a particular retention interval in effect for a prolonged period gives the birds time to realize that and to adjust the decision criterion accordingly. By adjusting the decision criterion in such a way as to maximize reinforcement, the performance of the birds exhibits a mirror effect.

That the birds actually do shift the criterion in light of the reinforcement consequences is perhaps seen most easily by considering how they behave when a retention interval is first increased from one condition to the next. The relevant data are presented in standard typeface in Table 1. Note that each time a retention interval was increased (e.g., from 0.5 to 2 sec, or from 2 to 6 sec), the hit rate dropped immediately (and precipitously), whereas the false alarm rate generally did not change. Over the next 15 sessions, both the hit rate and, to a lesser extent, the false alarm rate increased.

What explains that pattern? Presumably, on the first day that a new, longer retention interval was introduced, the effect on the mean of the sample distribution was immediate. That is, because of the longer retention interval, the memory trace of the sample faded to a greater extent, leading to a less intense sense of prior occurrence, on average. But that was the only immediate effect. The mean of the no-sample distribution did not change when the retention interval was increased, because there was no memory trace established on such trials to fade away. The criterion did not change, because the bird did not have enough time to learn where the optimal placement of the decision criterion might be in order to yield a higher number of reinforcers. As time passed, though, the bird learned to shift the criterion to the left on the sense-of-prior-occurrence axis (as is shown in Figure 7), thereby increasing both the hit rate and the false alarm rate.

The increase in the hit and the false alarm rates as a function of training at a particular retention interval is unmistakably asymmetric. That is, generally speaking, the hit rate increased a lot, whereas the false alarm rate increased only a little. Part of the explanation for this is that the appropriate detection model for this situation is not the equal-variance model we have used throughout this article but an unequal-variance model (so that the no-sample distribution is more variable than the sample distribution). A criterion sweeping across a narrow sample distribution will affect the hit rate to a much greater extent than a criterion sweeping across a wide no-sample distribution will affect the false alarm rate. In fact, in a separate experi-

Table 2
Hit and False Alarm (FA) Rates and *d'* Scores From
a Sample/No-Sample Procedure in Which Retention Interval
Was Manipulated Within Sessions

Retention Interval (sec)	Hit Rate	FA Rate	d'
0.5	.94	.14	2.63
2.0	.93	.20	2.32
6.0	.73	.17	1.56
12.0	.63	.16	1.32

Note—The data were taken from Experiment 1 of Wixted (1993).

ment, Wixted (1993) reported an analysis of the receiver operating characteristic that provided clear support for the unequal-variance model. As was discussed in detail by Stretch and Wixted (1998), the unequal-variance detection model complicates likelihood ratio models that assume underlying Gaussian distributions (e.g., under such conditions, there are two points on the sense-of-prior-occurrence axis where the likelihood ratio equals 1.0). The most straightforward way out of this problem is to assume that the distributions are Gaussian-like (e.g., Glanzer et al., 1993, assume a binomial distribution). The essence of the likelihood ratio account does not change by assuming different forms, so we will continue to use the equal-variance Gaussian model for the sake of simplicity.

It might still be tempting to argue that the increasing false alarm rate as a function of retention interval evident in Table 1 actually reflects the increasing strength of the no-sample distribution on the sense-of-prior-occurrence axis as the retention interval increased, rather than a criterion shift. An upward movement of the no-sample distribution would certainly increase the false alarm rate even if the location of the decision criterion remained fixed (although why it would increase in strength is not clear). However, evidence against this fixed-criterion interpretation is provided by studies that manipulate retention interval within sessions. Just as human subjects appear to be reluctant to shift the decision criterion item by item during the course of a recognition test, pigeons appear to be reluctant to shift the decision criterion trial by trial within a session. Thus, when retention interval is manipulated within sessions on a sample/no-sample task, the typical result is that the false alarm rate remains constant but the hit rate decreases rapidly as a function of retention interval (Dougherty & Wixted, 1996; Gaitan & Wixted, 2000; Wixted, 1993). Some relevant data from Wixted are presented in Table 2, and the theoretical representation of this situation in terms of signal detection theory is shown in Figure 8. The fact that the false alarm rate remains constant suggests that variations in the size of the retention interval do not affect the properties of the no-sample distribution. Instead, the simplest interpretation is that the no-sample distribution and the decision criterion are unaffected by the within-session retention interval manipulation. When retention interval is manipulated between conditions, there is no reason to assume that the properties of the no-sample distribution would now be affected by

that manipulation, although such an explanation is technically possible. Instead, it is much simpler (and much more theoretically sensible) to assume that the location of the decision criterion changes as a function of d'.

As an aside, it should be noted that pigeons are not always reluctant to shift the criterion from trial to trial. In a standard delayed matching-to-sample procedure, for example, pigeons appear to be quite willing to do just that. White and Cooney (1996) provided the most compelling evidence in this regard by showing that birds could learn to associate different reinforcement histories with different retention intervals (manipulated within session) when
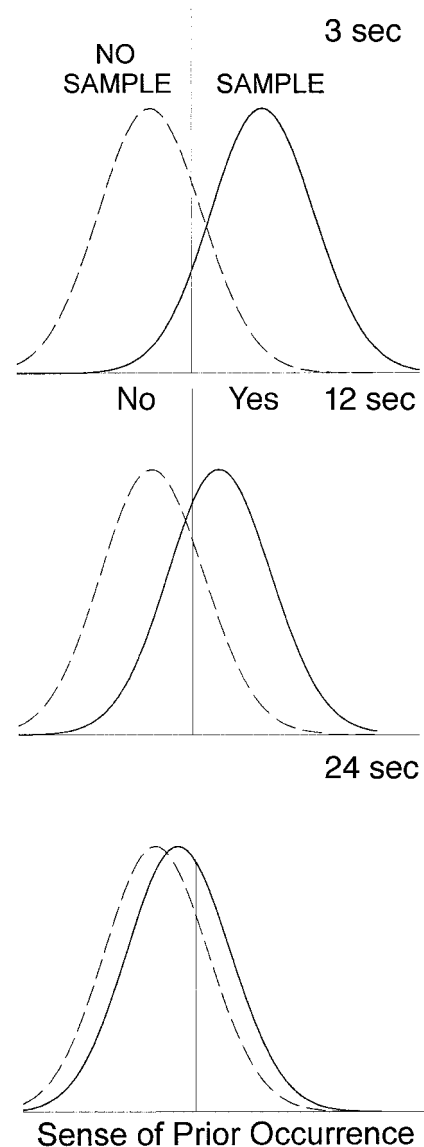

Figure 8. Hypothetical signal and noise distributions (corresponding to sample and no-sample trials, respectively) for three retention intervals (short, medium, and long) manipulated within sessions. The location of the decision criterion, *C*, is fixed.

the scheduled relative reinforcer probabilities for the two choice alternatives differed as a function of retention interval. Thus, although subjects are apparently reluctant to shift the criterion trial by trial on a sample/no-sample task (in the case of pigeons) or item by item in a yes/no recognition task (in the case of humans), it is clear that this is not an inviolable principle. Although this issue is clearly important, it is orthogonal to the main issue at hand, which is where the information comes from that allows subjects to behave like likelihood ratio computers whenever they do shift the criterion.

White and Wixted's (1999) Reinforcement History Model

It is important to emphasize that the decision criterion is a device used by theoreticians to represent one aspect of a bird's (or a human's) decision-making process. White and Wixted (1999) proposed that pigeons do not actually place a decision criterion on a subjective sense-of-prior-occurrence axis, even though it can be convenient to represent the situation as such. Instead, they suggested that, given a particular sense of prior occurrence on a given trial, the bird's behavior is governed by its history of reinforcement for making one choice or the other under similar conditions in the past. Consider, for example, a particular sense of prior occurrence equal to f in Figure 9. In the past, this value has sometimes been generated on no-sample trials and sometimes on sample trials. Usually, though, that value of f was generated on sample trials, so the probability of reinforcement for choosing the *sample* choice alternative given f is considerably higher than the probability of reinforcement for choosing the *no-sample* choice alternative given f. According to White and Wixted's model, birds learn the relative probabilities of reinforcement for each value of f along the sense-of-prior-occurrence axis.[2] If, for a given value of f, the probability of reinforcement for choosing the *sample* option is four times that for choosing the *no-sample* option, then, in accordance with the matching law, the bird will be four times as likely to choose *sample* as *no sample*. Thus, instead of *computing* the odds that the current value of f was produced by a sample trial

$p(f \mid NS)$ $p(f \mid S)$

f

Sense of Prior Occurrence

Figure 9. Hypothetical sample and no-sample distributions illustrating their relative heights at a particular point, f, on the sense-of-prior-occurrence axis.

(as opposed to a no-sample trial), the pigeon *learns* the odds that a reinforcer is arranged on the sample choice alternative (as opposed to the no-sample choice alternative) for the current value of f.

Note how similar this model is to the likelihood ratio model described earlier. Just as in the likelihood ratio account, the ratio of the height of the signal distribution to the ratio of the noise distribution for a particular value of f is of critical importance. Where the relevant distributional information comes from, though, is not well specified in contemporary likelihood ratio accounts. White and Wixted (1999) assumed that the organism had learned the relevant reinforcer ratios on the basis of prior experience and responded accordingly. As will be described next, a model like this winds up making predictions that are a lot like a likelihood ratio account.

Earlier, we noted that if the underlying sense-of-prior-occurrence distributions are Gaussian, the sample (or target) distribution with mean d' and standard deviation σ is given by

$$p(f \mid S) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-d')^2/\sigma^2}, \qquad (5)$$

and the no-sample (or lure) distribution with mean 0 and standard deviation σ is given by

$$p(f \mid NS) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-0)^2/\sigma^2}, \qquad (6)$$

where f now represents the *sense of prior occurrence* (analogous to *familiarity* in human recognition memory tasks) on a particular trial and $p(f \mid S)$ and $p(f \mid NS)$ represent the probability that a particular sense of prior occurrence will occur given a sample (S) or no-sample (NS) trial, respectively. If birds were likelihood ratio computers, we would assume that they know the forms of these distributions and could compute the ratio of $p(f \mid S) / p(f \mid NS)$ on a given trial. What the pigeons instead learn, according to White and Wixted's (1999) account, is the probability that a reinforcer is arranged on the sample choice alternative given f, which will be denoted as $p(R_S \mid f)$, and the probability that a reinforcer is arranged on the no-sample choice alternative given f, which will be denoted as $p(R_{NS} \mid f)$. Assuming that a sample trial is as likely to occur as a no-sample trial, the probability that a reinforcer is arranged on the sample choice alternative given f is

$$p(R_S \mid f) = \frac{p(f \mid S)}{p(f \mid S) + p(f \mid NS)} r_S, \qquad (7)$$

where r_S is the experimenter-arranged probability that a correct choice of the sample alternative will be reinforced (r_S is usually 1.0). Similarly, the probability that a reinforcer is arranged on the no-sample choice alternative given f is

$$p(R_{NS} \mid f) = \frac{p(f \mid NS)}{p(f \mid S) + p(f \mid NS)} r_{NS}, \qquad (8)$$

where r_{NS} is the experimenter-arranged probability that a correct choice of the sample alternative will be reinforced (r_{NS} is also usually 1.0).
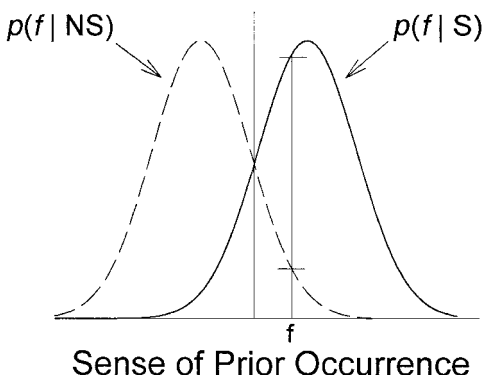
White and Wixted's (1999) model assumes that pigeons learn $p(R_S \mid f)$ and $p(R_{NS} \mid f)$ during the course of training, and in light of this information, they respond to the choice alternatives in accordance with the matching law. If $p(B_S \mid f)$ represents the probability of choosing the sample choice alternative given f and $p(B_{NS} \mid f)$ represents the probability of choosing the no-sample choice alternative given f, then

$$\frac{p(B_S \mid f)}{p(B_{NS} \mid f)} = \frac{p(R_S \mid f)}{p(R_{NS} \mid f)}, \qquad (9)$$

which, after replacing $p(R_S \mid f)$ and $p(R_{NS} \mid f)$ with their equivalent values from Equations 7 and 8, can be rewritten as

$$\frac{p(B_S \mid f)}{p(B_{NS} \mid f)} = \left(\frac{r_S}{r_{NS}}\right)\left[\frac{p(f \mid S)}{p(f \mid NS)}\right]. \qquad (10)$$

These equations state that the behavior ratio given f is equal to the reinforcer ratio given f, which is essentially the matching law for each value of f along the sense-of-prior-occurrence axis.

If $r_S = r_{NS} = 1.0$ (i.e., if the probability of reinforcement for a correct response on sample and no-sample trials, respectively, is 1.0), as is typically the case, Equation 10 reduces to

$$\frac{p(B_S \mid f)}{p(B_{NS} \mid f)} = \frac{p(f \mid S)}{p(f \mid NS)}.$$

Note that the term on the right side of this equation is the likelihood ratio. That is, after replacing $p(f \mid S)$ and $p(f \mid NS)$ with their equivalent values from Equations 5 and 6, this equation becomes

$$\frac{p(B_S \mid f)}{p(B_{NS} \mid f)} = \frac{\left[\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-d')^2/\sigma^2}\right]}{\left[\dfrac{1}{\sqrt{2\pi\sigma^2}} e^{-.5(f-0)/\sigma^2}\right]}.$$

Although the choice behavior ratio (left side of the equation) happens to equal the likelihood ratio (right side of the equation) when $r_S = r_{NS}$, a likelihood ratio computation is not assumed to be governing the bird's behavior. What drives the bird's choice, according to White and Wixted's (1999) model, are the learned odds that a reinforcer is arranged on one alternative or the other, not the computed odds that the current value of f was generated by a sample, as opposed to no sample. In a typical concurrent VI VI experiment, pigeons are generally assumed to learn the probabilities of reinforcement associated with the two choice alternatives and to respond in accordance with the matching law. White and Wixted's account makes exactly the same assumption, except that it assumes that the birds learn different reinforcement probabilities associated with the two alternatives, depending on what f happens to be. What f happens to be is determined, in part, by the form of the underlying sense-of-prior-occurrence distributions, although the birds are not assumed to have any direct knowledge of that.

There is some point on the sense-of-prior-occurrence axis (i.e., some value of f) that leads the bird to be indifferent between the two choice alternatives because, at that point, the experienced probability of reinforcement happens to be the same for both the sample and the no-sample choice alternatives. At the indifference point, $p(B_S \mid f) = p(B_{NS} \mid f)$. If $r_S = r_{NS} = 1.0$, it is easily shown that the value of f that yields indifference is $d'/2$. If $p(B_S \mid f) = p(B_{NS} \mid f)$, which is to say that $p(B_S \mid f) / p(B_{NS} \mid f) = 1$, then, according to Equation 10,

$$1 = \left(\frac{r_S}{r_{NS}}\right)\left[\frac{p(f \mid S)}{p(f \mid NS)}\right].$$

Solving this equation for f provides the point on the sense-of-prior-occurrence axis that yields indifference. Substituting the appropriate Gaussian density functions for $p(f \mid S)$ and $p(f \mid NS)$ from Equations 5 and 6 and canceling out the $1/\sqrt{2\pi\sigma^2}$ terms that appear in the numerator and denominator yields

$$1 = \left(\frac{r_S}{r_{NS}}\right)\left[\frac{e^{-.5(f-d')^2}}{e^{-.5(f-0)^2}}\right]$$

(assuming that $\sigma = 1$ for simplicity). When solved for f, this equation yields

$$f = \frac{d'}{2} + \frac{\ln\left(\dfrac{r_{NS}}{r_S}\right)}{d'}. \qquad (11)$$

Note the similarity between Equation 11 and Equation 4. According to Equation 11, if $r_S = r_{NS}$, so that $\ln(r_{NS}/r_S) = 0$, the indifference point occurs at $d'/2$. This familiarity value corresponds to the indifference point and could be represented as $f_{1/1} = d'/2$, where the subscript $1/1$ represents the ratio of programmed reinforcer probabilities (r_{NS}/r_S). This is exactly analogous to a central prediction of the likelihood ratio model, according to which the criterion will be located midway between the target and the lure distributions even when d' changes (thereby yielding a mirror effect). The difference is that White and Wixted's (1999) model assumes that no actual criterion is placed on a decision axis. Instead, the organism is responding in accordance with the matching law on the basis of its prior history of reinforcement for choosing the sample and no-sample alternatives for particular values of f. In this case, the pigeon has learned that when $f = d'/2$, the probability of reinforcement is the same for the sample and the no-sample choice alternatives. For any value greater than $d'/2$, the odds favor a reinforcer's being available for choosing the sample alternative (because that is the way it has been in the past), so the bird will no longer be indifferent. For any value less than $d'/2$, the odds favor a reinforcer's being available for choosing the no-sample alternative, so the bird will tend to prefer that alternative. Although reinforcement history and likelihood ratio accounts differ fundamentally, the predictions of the two accounts are the same—namely, that a mirror effect should be observed.

Again, although we assume that the indifference point represents the learned value of f associated with equal re-

inforcement probabilities, this is not to say that the situation cannot be represented with a criterion placed on a decision axis, as in Figure 7. What the criterion represents, though, is not a value that the organism has calculated on the basis of its inherent knowledge of the situation or a value that the bird has decided to use as a criterion value against which to evaluate all trial-by-trial sense-of-prior-occurrence values. Instead, the criterion simply depicts the point on the sense-of-prior-occurrence axis where, according to the subject's learning history, the probability of reinforcement for choosing the sample alternative is the same as the probability of reinforcement for choosing the no-sample alternative.

The similarity between the standard likelihood ratio account and the reinforcement history account proposed by White and Wixted (1999) extends beyond the mirror effect. In fact, if the probabilities of reinforcement for correct responses on sample and no-sample trials differ (e.g., if $r_{NS} > r_S$), the effect of that manipulation on the indifference point along the sense-of-prior-occurrence axis is exactly analogous to the effect of changes in d' on the location of the confidence criteria discussed earlier in connection with studies of human recognition memory. Earlier, we considered an example in which the high-confident *old* criterion was placed on the decision axis at the point where the odds that the test item was drawn from the target distribution was 10 to 1 in favor (and the high-confident *new* criterion was placed where the odds were 10 to 1 against). What we showed is that the distance between these two extreme criteria should increase (i.e., the confidence criteria should fan out, as in the lower panel of Figure 5) as d' decreases, and we also showed that the data qualitatively correspond to this prediction (as is shown in Figure 6). The same story emerges from the reinforcement history account proposed by White and Wixted. Consider, for example, where the choice indifference point would fall on the sense-of-prior-occurrence axis if the probability of reinforcement for a correct no-sample response was 10 times that of a correct sample response (e.g., $r_{NS} = 1.0$ and $r_S = .1$). Because reinforcers are so likely to be arranged on the no-sample choice alternative, the pigeon should demand an especially high sense of prior occurrence that the sample occurred before actually choosing the sample choice alternative. In other words, according to one way of thinking, the bird should have high confidence that the sample really did occur before choosing that alternative. How high should this indifference point be on the sense-of-prior-occurrence axis when $r_{NS} = 10r_S$? Substituting 10 for r_{NS}/r_S in Equation 11 and solving for f yields

$$f_{10/1} = \frac{d'}{2} + \frac{\ln(10)}{d'}$$
$$= \frac{d'}{2} + \frac{2.30}{d'}.$$

In other words, $f_{10/1}$ (the indifference point when $r_{NS} = 10r_S$) occurs at a point higher than $d'/2$, but by an amount that varies with d'. The more general version of this equa-

tion is $f_{10/1} = d'/2 + k_1/d'$, where k_1 is the constant r_{NS}/r_S. Similarly, if $r_S = 10r_{NS}$, then

$$f_{1/10} = \frac{d'}{2} + \frac{\ln(0.10)}{d'}$$
$$= \frac{d'}{2} - \frac{2.30}{d'}.$$

The more general version of this equation is $f_{1/10} = d'/2 + k_2/d'$, where k_2 is the constant r_{NS}/r_S. This model predicts that the distance between $f_{10/1}$ and $f_{1/10}$ on the sense-of-prior-occurrence axis will be inversely related to d'. That distance (i.e., $f_{10/1}$ minus $f_{1/10}$) is given by

$$f_{10/1} - f_{1/10} = \left(\frac{d'}{2} + \frac{k_1}{d'}\right) - \left(\frac{d'}{2} + \frac{k_2}{d'}\right)$$
$$= \frac{k_1 - k_2}{d'}.$$

Thus, as d' becomes very large, $f_{10/1}$ and $f_{1/10}$ should converge to the point that they coincide on the evidence axis (i.e., the distance between them should decrease to 0). By contrast, as d' approaches 0, those two indifference points should move infinitely far apart.

All of this is just another way of saying that the pigeon's sensitivity to variations in relative reinforcer probabilities should be low when d' is large and high when d' is small. This was actually the prediction tested by the experiments conducted by White and Wixted (1999). When d' was high (owing to a short retention interval), variations in relative reinforcer probabilities (i.e., variations in the ratio of r_{NS} to r_S) had little effect on behavior. When d' was low (owing to a long retention interval), identical variations in relative reinforcer probabilities had a large effect on behavior. In that sense, the birds were behaving like likelihood ratio computers. Because White and Wixted knew the reinforcer probabilities, they were able to plot relative choice probabilities as a function of relative reinforcer probabilities to show that sensitivity changed in the predicted manner as d' decreased. However, they could have instead computed indifference points for each reinforcer ratio condition, and the data would have shown that these points fan out on the sense-of-prior-occurrence axis as d' decreased (because that is just another way of saying that sensitivity to reinforcement increased as d' decreased). The pattern is the same as that observed for humans making confidence judgments, but the connection to the organism's reinforcement history is transparent only for the pigeons whose reinforcement history we had control over.

Why Do Humans Behave Like Likelihood Ratio Computers?

These considerations offer a new way to understand why humans tend to behave like likelihood ratio computers. As we considered earlier, humans tend to exhibit a mirror effect in recognition memory tasks, and more generally, they adjust their confidence criteria across conditions in a manner predicted by likelihood ratio models. Let us consider

again why a subject might provide a high-confident *old* response when an item is so familiar that the odds are 10 to 1 in favor of its having been drawn from the target distribution. From the likelihood ratio point of view, this occurs because the subject computes the relative heights of the target and lure distributions on the basis of knowledge of the shapes of the underlying distributions. If

$$\frac{p(f \mid \text{target})}{p(f \mid \text{lure})} > 10,$$

a high-confident *old* response is given. According to the reinforcement history account, by contrast, a high-confident *old* response is given when, for the current value of f, the odds that an *old* response would be correct *on the basis of prior feedback* is 10 to 1 in favor:

$$\frac{p(R_{\text{Old}} \mid f)}{p(R_{\text{New}} \mid f)} > 10,$$

where $p(R_{\text{Old}} \mid f)$ represents the past probability that, given this value of f, socially reinforcing feedback for an *old* response occurred and $p(R_{\text{New}} \mid f)$ represents the past probability that, given this value of f, socially reinforcing feedback for a *new* response occurred. If humans obey the matching law, then

$$\frac{p(B_{\text{Old}} \mid f)}{p(B_{\text{New}} \mid f)} = \frac{p(R_{\text{Old}} \mid f)}{p(R_{\text{New}} \mid f)}, \qquad (12)$$

which is exactly analogous to Equation 9. According to this equation, a high-confident *old* response would not always occur even if the reinforcer ratio on the right side of the equation were equal to 10, because the choice rule involves matching, not maximizing. Still, such a response would be 10 times as likely as a (low-confident) *new* response. If humans maximize rather than match, they would always respond *old* when the right side of the equation exceeds 1 (and would always respond *old* with high confidence when it exceeds, say, 10). The essence of our model does not change whether we assume that humans match or maximize.

By using exactly the same reasoning that converted Equation 9 into Equation 10, Equation 12 becomes

$$\frac{p(B_{\text{Old}} \mid f)}{p(B_{\text{New}} \mid f)} = \left(\frac{r_{\text{Old}}}{r_{\text{New}}}\right)\left[\frac{p(f \mid \text{target})}{p(f \mid \text{lure})}\right]. \qquad (13)$$

Thus, a high-confident *old* response is likely to occur when, for example,

$$\left(\frac{r_{\text{Old}}}{r_{\text{New}}}\right)\left[\frac{p(f \mid \text{target})}{p(f \mid \text{lure})}\right] > 10,$$

where r_{Old} and r_{New} represent the probabilities that reinforcing feedback occurred for correct *old* and *new* responses, respectively. These values are exactly analogous to r_{S} and r_{NS}, the experimenter-arranged probabilities of reinforcement for correct *sample* and *no-sample* choices. The values of r_{S} and r_{NS} in a pigeon experiment are controlled by the experimenter and are usually set to 1.0. By contrast, we have no way of knowing if, in life, $r_{\text{Old}} =$

r_{New}. Assuming, for the sake of simplicity, that they are equal, the mathematics of the reinforcement history account becomes equivalent to the mathematics of the likelihood ratio account. Conceptually, though, the two accounts are quite different. The learning model assumes that our human subjects, like our pigeons, have learned the relevant reinforcer ratios associated with each value of f along the familiarity axis. Whereas our birds would be 10 times as likely to choose, say, the sample alternative over the no-sample alternative under these conditions, our humans would be 10 times as likely to say *old* as *new*, and, if they did say *old*, they would do so with high confidence (knowing full well that the probability of reinforcing social feedback would be high, because that is the way it has been in the past). In essence, the math of the reinforcement history account reflects what experience has taught the subject about reinforcement outcomes for each value of f.

For pigeons, we needed to adjust the scheduled reinforcement probabilities in order to move the indifference point on the sense-of-prior-occurrence axis to the relatively high value at which the experienced probabilities of reinforcement for the two choice alternatives were equal. That is, because we cannot easily ask the pigeon for a confidence rating, we instead arrange the reinforcement outcome probabilities in such a way that the bird would be reluctant to choose the sample alternative unless it were highly confident that the sample was actually presented on the current trial. Thus, for example, if $r_{\text{NS}} = 10r_{\text{S}}$, the bird would generally not be inclined to choose the sample alternative, because reinforcers are so much more likely to be available on the no-sample choice alternative. Indeed, the indifference point on the sense-of-prior-occurrence axis under these conditions occurs not at $d'/2$, but at a higher point—namely, $d'/2 + 2.30/d'$. For humans, we usually do not alter the relative probability of reinforcement, but their confidence ratings can be used to tell us (theoretically) where, on the familiarity axis, the odds of prior reinforcement for an *old* response were high. In this sense, a typical human recognition memory experiment is essentially a *probe* test, revealing the effects of prior learning. A subject expresses high confidence in an *old* decision when the familiarity of a test item falls at a point where, in the past, the probability of reinforcement for an *old* response was especially high (e.g., 10 to 1 in favor).

The reinforcement learning model becomes identical to the likelihood ratio model if we assume that organisms do not learn specific $f \rightarrow$ reinforcement outcome relations but, instead, that they learn the mathematical forms of the underlying distributions and make the relevant familiarity-specific reinforcer ratio computations (such as a trained neural network might do). Even if this assumption were made, however, the interesting part of the explanation of why humans behave like likelihood ratio computers lies in the training history. To take one analogy, if a dog is trained to catch a Frisbee while running on its hind legs and winking one eye, we could say that the dog's neural network is performing miraculous feats. But the neural network is just doing what it was trained to do (and it would be doing

something else if it had been trained in a different way). The real explanation for why it is that the dog behaves in such an interesting way lies in the specifics of the dog's training history. Our argument is that the same considerations apply to humans expressing recognition memory confidence judgments, and the best way to gain insight into this process may be to study the effects of reinforcement on the memory behavior of nonhuman subjects.

Conclusion

Both humans and pigeons can be shown to behave in accordance with likelihood ratio models that appear to require inherent knowledge of the underlying familiarity distributions. Prevailing explanations for such behavior tend to focus on the computational abilities of the organism's neural network without detailing the origins of the relevant distributional information. A neural network may indeed perform such computations in just the manner suggested by likelihood ratio models, but our point is that the network does so because of the way it was trained by life or (in the case of pigeons) by laboratory experience. The neural network would presumably perform in a different way had the subject's reinforcement history been such that, for example, incorrect high-confident *old* responses were reinforced with higher probability than correct ones. In reality, we assume that socially reinforcing feedback is much more likely for correct than for incorrect high-confident *old* responses.

Abundant research in recent years has suggested that the behavior of humans working on signal detection tasks is governed in lawful ways by reinforcement outcomes (e.g., Erev, 1998; Johnstone & Alsop, 1999, 2000). It would be surprising indeed if only laboratory-based reinforcers affected their behavior in this way. Indeed, our central assumption is that humans are exquisitely sensitive to reinforcers that are arranged by life experiences. From this point of view, the interesting part of the story, the part that explains why subjects (including pigeons) appear to behave like likelihood ratio computers, lies not in the inherent properties of the subject's neural network, but in the subject's history of reinforcement. The behavior of the neural network (and the behavior of the subject) reflects that history, and in that sense, Skinner (1977) was probably right when he asserted that "the mental apparatus studied by cognitive psychology is simply a rather crude version of contingencies of reinforcement and their effects" (p. 9). If so, much of what is interesting about human memory will be illuminated only by studying animals whose reinforcement history can be studied in a much more direct way.

REFERENCES

ALSOP, B. (1998). Receiver operating characteristics from nonhuman animals: Some implications and directions for research with humans. *Psychonomic Bulletin & Review*, **5**, 239-252.

DAVISON, M. C., & TUSTIN, R. D. (1978). The relation between the generalized matching law and signal detection theory. *Journal of the Experimental Analysis of Behavior*, **29**, 331-336.

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.

EREV, I. (1998). Signal detection by human observers: A cutoff reinforcement learning model of categorization decisions under uncertainty. *Psychological Review*, **105**, 280-298.

GAITAN, S. C., & WIXTED, J. T. (2000). The role of "nothing" in memory for event duration in pigeons. *Animal Learning & Behavior*, **28**, 147-161.

GLANZER, M., & ADAMS, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 5-16.

GLANZER, M., ADAMS, J. K., IVERSON, G. J., & KIM, K. (1993). The regularities of recognition memory. *Psychological Review*, **100**, 546-567.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

JOHNSTONE, V., & ALSOP, B. (1999). Stimulus presentation ratios and the outcomes for correct responses in signal-detection procedures. *Journal of the Experimental Analysis of Behavior*, **72**, 1-20.

JOHNSTONE, V., & ALSOP, B. (2000). Reinforcer control and human signal-detection performance. *Journal of the Experimental Analysis of Behavior*, **73**, 275-290.

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

MCCLELLAND, J. L., & CHAPPELL, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, **105**, 734-760.

MORRELL, H. E. R., GAITAN, C., & WIXTED, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1095-1110.

SHIFFRIN, R. M., & STEYVERS, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, **4**, 145-166.

SKINNER, B. F. (1953). *Science and human behavior*. New York: Macmillan.

SKINNER, B. F. (1977). Why I am not a cognitive psychologist. *Behaviorism*, **5**, 1-10.

STRETCH, V., & WIXTED, J. T. (1998a). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1397-1410.

STRETCH, V., & WIXTED, J. T. (1998b). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1379-1396.

WHITE, K. G., & COONEY, E. B. (1996). Consequences of remembering: Independence of performance at different retention levels. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 51-59.

WHITE, K. G., & WIXTED, J. T. (1999). Psychophysics of remembering. *Journal of the Experimental Analysis of Behavior*, **71**, 91-113.

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

NOTES

1. A potentially confusing issue is that the symbol C is sometimes used as a measure of response bias and its value is equal to the criterion's distance from the point of intersection between the signal and the noise distributions. By contrast, we use C to denote the location of the criterion relative to the mean of the noise distribution.

2. Actually, the probability that a continuous random variable assumes a particular value, f, is zero. Thus, in what follows, we will actually assume that the value in question is some small interval on the familiarity axis that does have a certain probability of occurrence.