Attention, Perception, & Psychophysics
2009, 71 (6), 1426-1433
doi:10.3758/APP.71.6.1426

Likelihood ratio, optimal decision rules, and relationship between proportion correct and *d′* in the dual-pair AB-versus-BA identification paradigm

Christophe Micheyl
University of Minnesota, Minneapolis, Minnesota

and

Huanping Dai
University of Arizona, Tucson, Arizona

The equal-variance Gaussian signal detection theory (SDT) decision model for the dual-pair (4IAX) change-detection paradigm has been described in earlier publications. In this research article, we consider the equal-variance Gaussian SDT model for the related 4IAX AB-versus-BA identification paradigm. The likelihood ratios, optimal decision rules, receiver-operating characteristics (ROCs), and relationships between *d′* and proportion correct (PC) are analyzed for two special cases: that of statistically independent observations, which typically applies in constant-stimuli experiments, and that of highly correlated observations, which typically applies in experiments where stimuli are roved widely across trials or pairs. A surprising outcome of this analysis is that, although these two situations lead to different optimal decision rules, the predicted ROCs and PC responses for these two cases are not substantially different and are either identical to or similar to those observed in the basic yes–no paradigm. Supplemental materials for this study can be downloaded from app.psychonomic-journals.org/content/supplemental.

The pioneers of psychophysical signal detection theory (SDT; e.g., Green & Swets, 1966) have made the two-interval, two-alternative forced choice (2I-2AFC) paradigm popular among psychophysicists by outlining its lower susceptibility to response biases, as compared with the yes–no and *same–different* paradigms. Consequently, numerous psychophysical studies over the past 50 years have used the 2I-2AFC paradigm to measure sensory "discrimination" thresholds. However, consistently correct performance in the 2I-2AFC paradigm usually requires more than just an ability to detect that two stimuli are different; in addition, the order of presentation of the two stimuli (e.g., low–high vs. high–low), or the "direction" of the stimulus change (e.g., upward vs. downward), must be correctly identified. These two abilities are not always closely related. For instance, recent results in the psychoacoustic literature reveal that some listeners are unable to correctly identify the direction of changes in pitch, even though they can very easily detect these changes (Semal & Demany, 2006).

One approach that can be used to measure a participant's ability to detect changes (or differences) between two stimuli without requiring him or her to be able to identify the direction of the change involves presenting two pairs of stimuli on each trial. One of the two pairs contains identical stimuli (AA or BB), whereas the other contains different stimuli (AB or BA). The two pairs are presented in random order, and the participant's task is to indicate which pair contains different stimuli. This "dual-pair" design is sometimes referred to in the literature as the *four-interval same–different* or *4IAX* paradigm (Creelman & Macmillan, 1979; Macmillan, Kaplan, & Creelman, 1977; Noreen, 1981; Rousseau & Ennis, 2001). Its main advantage is that it retains the relative independence from bias of the 2I-2AFC paradigm but does not require the participant to identify the direction of stimulus change along some dimension (e.g., intensity or frequency).

Several recent studies of auditory perception have used a 4IAX design to measure thresholds for the detection of differences in basic sound attributes, such as intensity, frequency, or modulation rate, and to compare these to thresholds for the identification of the directions of these changes (Demany, Carlyon, & Semal, 2009; Mathias & Bailey, 2008; Micheyl, Kaernbach, & Demany, 2008; Semal & Demany, 2006). It should be noted that the two types of thresholds were measured using the same stimuli; only the instructions given to the participants differed. In one case, participants were asked to indicate which of the

C. Micheyl, cmicheyl@umn.edu

two stimulus pairs contained a change, regardless of the direction (BA or AB) of that change. In the other case, they were asked to report the direction of the change in the pair containing different stimuli, regardless of whether that pair occurred first or second. This illustrated a unique feature of the 4IAX design, which is that it allows thresholds for the detection of stimulus changes and thresholds for the identification of the direction of these changes, to be measured adaptively under rigorously identical stimulus conditions; only the instructions given to the participants differed. This could not be achieved as easily with two-interval designs, such as *same–different* and 2I-2AFC, for at least two reasons. First, to accurately assess the false alarm rate in the *same–different* paradigm, the experimenter must present a sufficiently large number of trials on which the two stimuli are physically identical. Most existing adaptive threshold-tracking procedures make no provision for the introduction of such identical-stimulus trials. In addition, being asked to identify the direction of the change on such trials may confuse the participant. In contrast, in the 4IAX design, one of the two stimulus pairs always contains a physical change. Second, as mentioned earlier, the two-interval *same–different* task is prone to response biases. The 4IAX change-detection paradigm alleviates this problem by embedding the two-interval *same–different* task into a 2I-2AFC design. The 4IAX change-direction (AB vs. BA) identification task is no more likely to be affected by response biases than is the popular 2I-2AFC task. Thus, it can be used in conjunction with adaptive procedures that track a predefined proportion of correct responses (PC) on the psychometric function.

The 4IAX change-direction and change-detection identification paradigms are superficially very similar; they both involve two pairs of stimuli, only one of which contains different stimuli. However, this apparent similarity is misleading. The optimal decision rules in these two tasks are actually very different. Thus, one cannot directly compare PC thresholds across these two tasks. First, the measured PCs must be transformed into a common metric, such as d'. Appropriate formulas for transforming PCs measured in the 4IAX change-detection paradigm into d' have been worked out in previous publications after detailed analyses of the underlying equal-variance Gaussian SDT decision model (Creelman & Macmillan, 1979; Macmillan et al., 1977; Micheyl & Dai, 2008; Micheyl & Messing, 2006; Noreen, 1981; Rousseau & Ennis, 2001). In contrast, the equal-variance Gaussian SDT model for the 4IAX change-direction (AB vs. BA) identification task has not been analyzed in any great detail. To our knowledge, the only previous publication in which this model has been considered is a recent article by Micheyl et al. (2008). By considering the geometry of the decision space in the 4IAX change-direction identification task, these authors came to the conclusion that the optimal decision rule and expected relationship between PC and d' are quite different for this task than for its 4IAX change-detection counterpart.

The purpose of the present article is to provide a synthetic and more thorough analysis of the equal-variance Gaussian SDT model for the 4IAX direction-identification task. The article is organized into three sections. In the first section, the equal-variance Gaussian SDT model of the sensory and decision processes in the 4IAX change-direction identification paradigm is briefly described, and expressions for the likelihood ratio are provided for two common experimental situations: that of no stimulus *roving* (random variation over a wide range) and that of wide stimulus roving. It is found that the likelihood ratios for these two cases are quite different. In the second section, optimal decision rules based on the likelihood ratios derived in the first section for the no-roving and wide-roving cases are explained. In the third and final section, receiver-operating characteristic (ROC) curves and the relationship between d' and the PC achieved by an unbiased optimal observer (PC_{max}) are analyzed. One surprising outcome of this analysis is that, although the 4IAX AB-versus-BA identification task is superficially quite different from the basic yes–no task, the predicted relationship between d' and PC_{max} in these two paradigms is exactly the same when wide roving is used. Another unexpected outcome is that, although the optimal decision rules in the no-roving and wide-roving cases are different, the predicted relationships between d' and PC_{max} in these two cases are indistinguishable, as long as d' is less than about 1.5.

Likelihood Ratio

In modeling the sensory and decision processes underlying performance in the 4IAX AB-versus-BA identification paradigm, we used the framework of psychophysical SDT (Green & Swets, 1966). The sensory process is conceived as a probabilistic mapping from a physical stimulus parameter space onto a sensory observation space. Accordingly, each observation is modeled as a random variable with a probability distribution that is dependent on which stimulus is presented. In the 4IAX AB-versus-BA identification paradigm, a trial consists of four stimuli—s_1, s_2, s_3, and s_4—in which the subscript refers to the observation interval: Intervals 1 and 2 correspond to the first and second stimuli in the first pair, and Intervals 3 and 4 correspond to the first and second stimuli in the second pair. Assuming that each stimulus yields one sensory observation, a trial results in four observations—denoted here as y_1^k, y_2^k, y_3^k, and y_4^k—where the superscript k denotes the trial and the subscript denotes the observation interval. It is worth stressing that each y_i^k represents one realization of a random variable Y, the value of which fluctuates across stimulus intervals because of changes in stimulus parameter(s) and/or changes in internal noise. For simplicity, and without any loss of generality, the variance of the internal noise is set to 1.

The decision process is modeled as a deterministic mapping from the sensory-observation space to a response space, the elements of which are the different response alternatives. In the paradigm considered here, there are two response alternatives used for indicating whether the pair that contains the different stimuli is AB or BA. The mapping from the observations to the response space is determined by a decision rule. The optimal observer uses a decision rule that maximizes PC. It can be shown that

$$\lambda = \frac{f_{y|ABAA}(y \mid ABAA) + f_{y|ABBB}(y \mid ABBB) + f_{y|AAAB}(y \mid AAAB) + f_{y|BBAB}(y \mid BBAB)}{f_{y|BAAA}(y \mid BAAA) + f_{y|BABB}(y \mid BABB) + f_{y|AABA}(y \mid AABA) + f_{y|BBBA}(y \mid BBBA)} \quad (1)$$

this is achieved by always choosing the most likely alternative, given the observations, which amounts to basing the decision on the likelihood ratio (Green & Swets, 1966). Here, the likelihood ratio is given by Equation 1, above. In this equation, $f_{y|S}(y \mid S)$ is used to denote the conditional probability of the vector of observations, $y = [y_1, y_2, y_3, y_4]^T$, where the superscript T indicates vector transposition, given that the presented sequence of stimuli is S, with $S \in \{$ABAA, ABBB, AAAB, BBAB, BAAA, BABB, AABA, BBBA$\}$. In this and in the following equations, the superscript k is implicit.

Following a long tradition in the SDT literature, we assume that the conditional distribution functions in Equation 1 are multivariate Gaussian probability density functions:

$$f_{y|S}\left(y \mid S\right) = \frac{1}{\sqrt{2\pi \mid \Sigma^2 \mid}} e^{-\frac{1}{2}\left(y-\mu_S\right)^T \Sigma^{-2}\left(y-\mu_S\right)}. \quad (2)$$

In this equation, Σ^2 denotes the variance–covariance matrix of the observations, $|\Sigma^2|$ denotes the determinant of that matrix, and Σ^{-2} denotes its inverse. μ_S is a four-element vector that contains the conditional expected value of Y, given that stimulus sequence S was presented. As an example, for the AABA sequence, $\mu_S = \mu_{AABA} = [\mu_A, \mu_A, \mu_B, \mu_A]^T$, where μ_A is the expected value of the random variable Y_i conditioned on interval i containing stimulus A—that is, $\mu_A = \mathrm{E}[Y_i | S_i = \mathrm{A}]$—and μ_B is the expected value of Y_i conditioned on interval i containing stimulus B—that is, $\mu_B = \mathrm{E}[Y_i | S_i = \mathrm{B}]$. Introducing $d = \mu_B - \mu_A$, we have $\mu_{AABA} = [\mu - d/2, \mu - d/2, \mu + d/2, \mu - d/2]^T$.

Following Dai, Versfeld, and Green (1996), Versfeld, Dai, and Green (1996), and Micheyl and Messing (2006), we consider two extreme cases: that of statistically independent observations and that of highly correlated observations. The independent-observations assumption is commonly made when one analyzes the results of studies in which stimulus parameters are not roved across trials, and the only source of variability in the sensory observations stems from internal noise. To the extent that the internal-noise samples that contaminate the sensory observations are statistically independent, the sensory observations are also statistically independent from each other, conditional on a given stimulus sequence. If the internal-noise samples that are added to the observations are not statistically independent—due, for example, to slow fluctuations in internal noise, or if stimulus roving is used—the observations are no longer statistically independent from each other. Typically, across-trial stimulus roving is used by experimenters to discourage listeners from comparing individual stimuli on the current trial with memory traces of individual stimuli presented on previous trials. Mathematically, such across-trial roving is modeled by adding one realization (r^k) of a random variable (R) to all four sensory observations on a trial. The addition of a common source of variation to all four observations on a trial introduces some correlation between these observations. Typically, the magnitude of the random variation applied to the stimuli is large in comparison with the differences between the stimuli that the participant must discriminate; otherwise, the roving is unlikely to be effective (Green, 1988). Therefore, the extra noise introduced by roving is large, as compared with the internal noise. This results in high correlation between the sensory observations on a trial. This is referred to as the *highly correlated observations* case.

In the presence of across-trial roving or of some other source of correlation among all four observations on a trial, the variance–covariance matrix of the observations has the following form:

$$\Sigma^2 = \begin{bmatrix} v+1 & v & v & v \\ v & v+1 & v & v \\ v & v & v+1 & v \\ v & v & v & v+1 \end{bmatrix}, \quad (3)$$

where v denotes the common variance in the observations. This variance is related to the (Pearson product–moment) correlation coefficient ρ between the observations by $v = \rho/(1 - \rho)$. When ρ is 0 (no correlation), v is also 0. As ρ approaches $+1$ or -1, v tends toward $+\infty$ or $-\infty$, respectively.

The likelihood ratio for the case of independent observations (i.e., no roving) is obtained by setting v equal to 0 in Σ^2 (which then reduces to a four-by-four identity matrix), plugging Equation 3 into Equation 2, and plugging Equation 2 into Equation 1. The algebraic manipulations required for arriving at a simplified solution are fairly straightforward but tedious; they were performed using a computer algebra system, Mathematica 5.2 (Wolfram Research). The code is provided as a supplementary material for this article ("Mathematica Code for Deriving Likelihood Ratios in the Dual-Pair AB Versus BA Identification Paradigm With Independent and Highly Correlated Gaussian Observations"). The solution is

$$\lambda_{io} = \frac{e^{d(2\mu+y_2)} + e^{d(2\mu+y_4)} + e^{d(y_1+y_2+y_4)} + e^{d(y_2+y_3+y_4)}}{e^{d(2\mu+y_1)} + e^{d(2\mu+y_3)} + e^{d(y_1+y_2+y_3)} + e^{d(y_1+y_3+y_4)}}. \quad (4)$$

The case of highly correlated observations that are due to wide across-trial roving is obtained by letting $v \to +\infty$ in Σ^2, plugging Equation 3 into Equation 2, and then plugging Equation 2 into Equation 1. The result, which was also obtained using Mathematica (the code for which is available as a supplementary material, "Mathematica Code for Deriving Likelihood Ratios in the Dual-Pair AB Versus BA Identification Paradigm With Independent and Highly Correlated Gaussian Observations"), is a somewhat simpler expression:

$$\lambda_{hco} = e^{d(y_2 - y_1 + y_4 - y_3)}. \quad (5)$$

$$\lambda = \frac{f\left(y_{p_1} \mid AB\right)\left[f\left(y_{p_2} \mid AA\right)+f\left(y_{p_2} \mid BB\right)\right]+f\left(y_{p_2} \mid AB\right)\left[f\left(y_{p_1} \mid AA\right)+f\left(y_{p_1} \mid BB\right)\right]}{f\left(y_{p_1} \mid BA\right)\left[f\left(y_{p_2} \mid AA\right)+f\left(y_{p_2} \mid BB\right)\right]+f\left(y_{p_2} \mid BA\right)\left[f\left(y_{p_1} \mid AA\right)+f\left(y_{p_1} \mid BB\right)\right]} \tag{6}$$

Optimal Decision Rule

The likelihood ratios described in the preceding section can be used to determine optimal decision rules for the independent-observations (no-roving) case and the highly correlated (wide-roving) case. As mentioned above, an optimal decision rule—that is, a rule that maximizes PC—consists of selecting the response alternative that is the most likely a posteriori, given the observations. If, as is usually the case, the different stimulus sequences are all equally likely a priori, this is tantamount to responding "AB" if the likelihood ratio given by Equation 4 (for the no-roving case) or by Equation 5 (for the wide-roving case) is larger than 1 and responding "BA" otherwise. Equivalently, the decision may be based on any monotonic transformation of the likelihood ratio.

Taking the logarithm of the right-hand side of Equation 5, it is readily found that, for the case of highly correlated observations, an optimal decision rule for the unbiased observer is to respond "AB" if $y_2 - y_1 > y_3 - y_4$ and to respond "BA" otherwise. In words, this decision rule can be described as follows. First, subtract the observation evoked by the first stimulus in the first pair from that evoked by the second stimulus in the same pair. Second, subtract the observation evoked by the second stimulus in the second pair from that evoked by the first stimulus in that pair. Compare the two resulting differences and respond "AB" if the first difference is larger than the second; otherwise, respond "BA." An equivalent and perhaps more intuitive decision rule is to select the pair in which the difference between the two observations is larger in magnitude and respond "AB" if this difference is negative; otherwise, respond "BA." Decision rules based on differences between the observations on a trial are traditionally referred to in the SDT literature as *differencing strategies*.

It is worth noting that μ, which was introduced above to denote the expected value of the observations, does not appear in the likelihood ratio for the highly correlated observations case (Equation 5). This indicates that, in this case, having a fixed reference corresponding to the center of gravity of the observations is useless for task performance. This is consistent with the idea that roving the stimuli over a wide range across trials prevents listeners from benefiting from the use of a fixed (absolute) reference point along the observation axis. Note also that taking the difference between the observations in each pair removes any positive correlation that might exist between the two observations within a pair. In some experimental applications of the 4IAX paradigm, roving is applied not only across trials but also across the two pairs of observations on a trial. This form of roving is produced by adding different realizations of the roving variable R to the first and second pairs of stimuli. Such across-pair roving is intended to prevent listeners from comparing individual stimuli across the two pairs and to force them, instead, to compare the two stimuli in each pair. For the optimal observer, however, it does not matter whether roving is applied across trials or across pairs, because taking the difference between the two observations within each pair eliminates both forms of correlation between the observations.[1]

The optimal decision rule for the independent-observations case is less straightforward. Although Equation 4 provides one of the simplest expressions of the likelihood ratio for this situation, it does not provide a very intuitive interpretation of the optimal decision strategy. A more intuitive interpretation of the optimal decision rule is actually obtained by rearranging the terms in Equation 1 as shown in Equation 6, above. In this equation, y_{p_1} and y_{p_2} denote vectors containing the two observations from the first and second pairs—that is, $y_{p_1} = [y_1, y_2]^T$, and $y_{p_2} = [y_3, y_4]^T$, respectively. In this new expression, the numerator is the sum of two probabilities: the probability that the first pair is AB and the second pair is AA or BB, and the probability that the second pair is AB and the first pair is AA or BB. This is the same as the likelihood that the first or second pair is AB. Likewise, the denominator is the sum of the probability that the first pair is BA (and the second pair is AA or BB) and the probability that the second pair is BA (and the first pair AA or BB). This is the same as the likelihood that the first or second pair is BA. The optimal decision rule for the case of independent observations involves a comparison between these two likelihoods, as follows:

$$\lambda_{AB} \underset{\text{“}BA\text{”}}{\overset{\text{“}AB\text{”}}{\gtrless}} \lambda_{AB}, \tag{7}$$

with

$$\lambda_{AB} = f(y_{p_1}|AB)[f(y_{p_2}|AA) + f(y_{p_2}|BB)] + f(y_{p_2}|AB)[f(y_{p_1}|AA) + f(y_{p_1}|BB)]$$

and

$$\lambda_{BA} = f(y_{p_1}|BA)[f(y_{p_2}|AA) + f(y_{p_2}|BB)] + f(y_{p_2}|BA)[f(y_{p_1}|AA) + f(y_{p_1}|BB)].$$

In the supplementary materials for this article, in "The Optimal Decision Rules in the Dual-Pair AB Versus BA Identification Paradigm Explained in Terms of Template Matching," we describe a third view of the optimal decision rules for the independent- and highly correlated observations cases, in terms of template matching.

ROCs and the Relationship Between d' and PC_{max}

Since we have found that the optimal (likelihood ratio) decision rule in the 4IAX AB-versus-BA identification paradigm differs depending on whether stimulus roving is used, a question of practical interest for experiment-

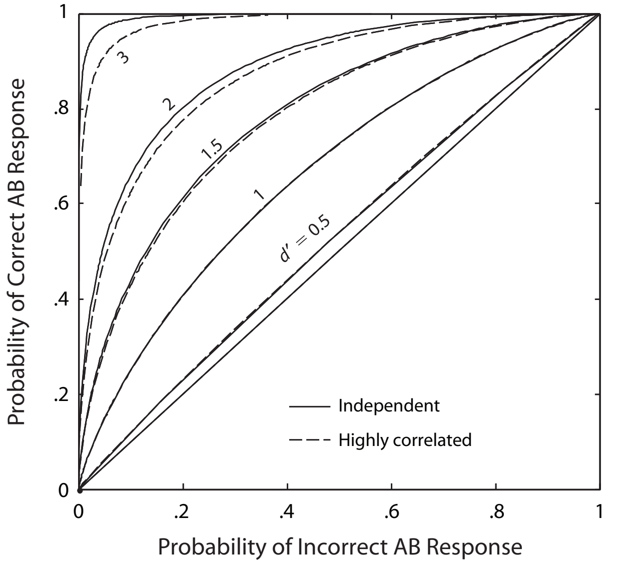

Figure 1. Receiver-operating characteristics (ROCs) in the dual-pair AB-versus-BA identification paradigm. The solid lines correspond to the independent-observation (no-roving) case. The dashed lines correspond to the highly correlated observations (wide-roving) case. The latter are identical to yes–no ROCs. Each curve corresponds to a different *d′* value: 0 (straight diagonal line), 0.5, 1, 1.5, 2, and 3, as indicated on the plot.

ers is how this may affect the measured PCs. A first step toward answering this question involves looking at ROCs for the version of the 4IAX paradigm considered here. Examples of such ROCs are shown in Figure 1 for both the independent-observations case (solid lines) and the highly correlated observations case (dashed lines). Each curve corresponds to a given d' value equal to 0 (straight diagonal line), 0.5, 1, 1.5, 2, or 3. Details of how these ROCs were computed can be found in the supplementary materials for this article as "Monte Carlo Simulations for Computing ROCs and PC_{max}."

Two observations of practical importance can be made regarding these ROCs. First, for d' values lower than about 1.5, the ROCs for the independent- and highly correlated observations cases are practically indistinguishable. For d' values larger than 1.5, ROCs for the independent-observations case are systematically closer to the upper left-hand corner than are the corresponding ROCs for the highly correlated observations case. Thus, over this range of sensitivity, the area under the ROC is larger for the independent-observations case than for the highly correlated observations case. Since the area under the ROC provides an unbiased measure of performance (Egan, 1975; Green & Swets, 1966), this indicates that, under these circumstances, an optimal observer should achieve a higher level of performance if the observations are uncorrelated (i.e., if the stimuli are not roved) than if they are correlated (i.e., if the stimuli are roved across trials or pairs). However, the difference is quite small: $\Delta A_{ROC} = 0.012$ for $d' = 1.5$; $\Delta A_{ROC} = 0.018$ for $d' = 2$; and $\Delta A_{ROC} = 0.011$ for $d' = 3$. Given the usually limited accuracy and inherent variability of empirical ROC data, such small differences are unlikely to matter in practice.

The analysis of 4IAX AB-versus-BA identification data is further simplified by a second observation regarding the ROCs in Figure 1. For the case of highly correlated observations, the ROCs (dashed lines) are identical to yes–no ROCs. This outcome may seem surprising at first sight, since the 4IAX paradigm considered here and the yes–no paradigm are superficially very different; the former involves the presentation of four stimuli on each trial, as opposed to a single stimulus in the latter. However, the deep connection between the two paradigms can be understood by noting that the likelihood ratio defined by Equation 5 is a monotonic function of $Z = y_1 - y_2 + y_3 - y_4$. Thus, the optimal decision rule can use Z instead of the likelihood ratio. Since the y_is are Gaussian random variables and Z is a sum of y_is, Z is itself a Gaussian random variable. Conditioned on the occurrence of an AB pair, the expected value and variance of Z are $E(Z \mid AB) = d$ and $Var(Z \mid AB) = 4\sigma^2$, respectively. Conditioned on the occurrence of a BA pair, the expected value and variance of Z are $E(Z \mid BA) = -d$ and $Var(Z \mid BA) = 4\sigma^2$, respectively. Thus,

$$d' = \frac{E(Z \mid AB) - E(Z \mid BA)}{\sqrt{Var(Z)}} = \frac{2d}{\sqrt{4\sigma^2}} = d. \qquad (8)$$

Above, d was defined as the difference between the expected values of the distributions of sensory observations evoked by the two stimuli, A and B. Therefore, Equation 8 reveals that, for the case of highly correlated observations, d' in the 4IAX AB-versus-BA identification paradigm is equal to the difference between the means of the distributions of observations corresponding to the two stimuli, just as in the basic yes–no paradigm. Additional insight into this unexpectedly simple connection between the yes–no paradigm and the 4IAX AB-versus-BA identification task with highly correlated observations can be obtained by noting that the decision spaces for these two tasks are equivalent (Micheyl et al., 2008).

Yes–no ROCs are discussed at length in SDT textbooks (Egan, 1975; Green & Swets, 1966; Macmillan & Creelman, 2005; Wickens, 2001). One of their most significant features is that, when plotted on z-coordinates, these predicted ROCs are linear, with an intercept equal to d'. Accordingly, experimenters can calculate d' in the 4IAX AB-versus-BA identification paradigm as $d' = \Phi^{-1}(P_H) - \Phi^{-1}(P_F)$ (Green & Swets, 1966; Macmillan & Creelman, 2005), where $\Phi^{-1}(.)$ denotes the inverse cumulative normal function, P_H is the measured proportion of hits (defined as an "AB" response to an AAAB, BBAB, ABAA, or ABBB stimulus), and P_F is the measured proportion of false alarms (defined as an "AB" response to an AABA, BBBA, BAAA, or BABB stimulus). The same equation can also be used to obtain approximate d' values for the case of independent observations, provided that d' is less than about 1.4. For higher values, the approximation becomes inaccurate to the first decimal. This is shown in Table 1, which displays $\Phi^{-1}(P_H) - \Phi^{-1}(P_F)$ values corresponding to d' values between 0.1 and 4, in steps of 0.1.

Table 1
$\Phi^{-1}(P_H) - \Phi^{-1}(P_F)$ Values Corresponding to d' Values Between 0.1 and 4.0 (in Steps of 0.1) in the Dual-Pair AB-Versus-BA Identification Paradigm Under the Independent Observations Model

d'	$\Phi^{-1}(P_H) - \Phi^{-1}(P_F)$	d'	$\Phi^{-1}(P_H) - \Phi^{-1}(P_F)$
0.1	0.102	2.1	2.324
0.2	0.202	2.2	2.449
0.3	0.300	2.3	2.587
0.4	0.399	2.4	2.723
0.5	0.500	2.5	2.857
0.6	0.602	2.6	2.995
0.7	0.703	2.7	3.132
0.8	0.807	2.8	3.274
0.9	0.912	2.9	3.415
1.0	1.017	3.0	3.552
1.1	1.125	3.1	3.700
1.2	1.239	3.2	3.846
1.3	1.349	3.3	3.988
1.4	1.460	3.4	4.133
1.5	1.577	3.5	4.281
1.6	1.694	3.6	4.429
1.7	1.817	3.7	4.576
1.8	1.939	3.8	4.716
1.9	2.064	3.9	4.866
2.0	2.192	4.0	5.013

Details on how these values were computed can be found in the supplementary materials for this article as "Monte Carlo Simulations for Computing ROCs and PC_{max}."

Assuming equal a priori probabilities of occurrence for the various stimulus sequences, one can determine PC_{max}, the PC that should be achieved by an unbiased optimal observer. It corresponds to the probability that the likelihood ratio in Equation 4 (for the independent-observations strategy) or Equation 5 (for the differencing strategy) exceeds 1, given that the stimulus sequence contains an AB pair. One approach to computing this probability involves integrating the joint probability density function of the observations y_1, y_2, y_3, and y_4, conditioned on the occurrence of an ABAA, ABBB, AAAB, or BBAB stimulus sequence. For the case of highly correlated observations, carrying out the integration is not necessary because, as we demonstrated above (Equation 8), d' simply equals the difference between the means of the distributions of observations corresponding to the two stimuli, as in the basic yes–no paradigm. It follows that, for this case, PC_{max} in the 4IAX AB-versus-BA identification paradigm is related to d' in exactly the same way as in the basic yes–no task: $PC_{max} = \Phi(d'/2)$ (Green & Swets, 1966; Macmillan & Creelman, 2005).

For the independent-observations case, the relationship between PC_{max} and d' is more complicated. The reason for this is that, as revealed by Equation 4, the likelihood ratio for this case involves a ratio of sums of exponentials, which makes the integration route less appealing. Fortunately, the solution can also be obtained using Monte Carlo simulations, as described in the supplementary materials for this article, "Monte Carlo Simulations for Computing ROCs and PC_{max}."

Figure 2 shows how PC_{max} varies as a function of d' in the 4IAX AB-versus-BA identification task. For comparison, the relationship is also shown for other popular paradigms. Consistent with our earlier observation—that when d' exceeds about 1.4, the ROC area in Figure 1 was larger for independent observations (no roving) than for

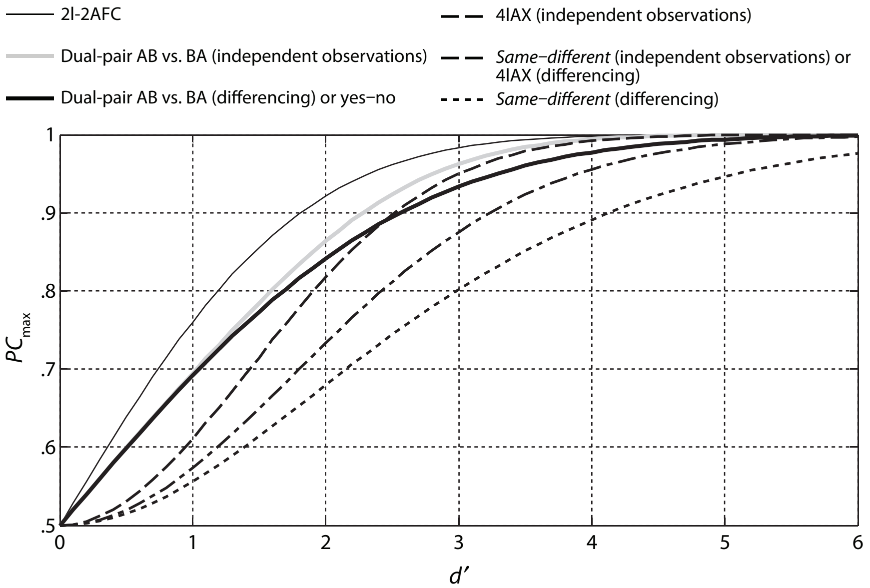

Figure 2. Relationship between d' and PC_{max} in the dual-pair (4IAX) AB-versus-BA identification paradigm and other psychophysical paradigms. The curves corresponding to the 4IAX AB-versus-BA identification paradigm are shown as thick lines. The dark thick line corresponds to the independent-observation (no-roving) case. The lighter (gray) thick line corresponds to the differencing strategy (wide-stimulus-roving, highly correlated observations). The function relating PC_{max} to d' for other psychophysical paradigms, including 2I-2AFC, yes–no, *same–different*, and 4IAX, are also shown here for comparison.

Table 2
***d′* Values Corresponding to PC_{max} Values Between .50 and 1.0 (in Steps of 0.01) for the Dual-Pair AB-Versus-BA Identification Paradigm Under the Independent-Observations Model**

PC_{max}	d'	PC_{max}	d'	PC_{max}	d'
.50	0.00	.67	0.87	.84	1.84
.51	0.05	.68	0.92	.85	1.91
.52	0.10	.69	0.98	.86	1.98
.53	0.15	.70	1.03	.87	2.05
.54	0.20	.71	1.08	.88	2.12
.55	0.25	.72	1.14	.89	2.20
.56	0.30	.73	1.19	.90	2.28
.57	0.35	.74	1.25	.91	2.37
.58	0.40	.75	1.30	.92	2.47
.59	0.45	.76	1.36	.93	2.57
.60	0.51	.77	1.42	.94	2.68
.61	0.56	.78	1.47	.95	2.81
.62	0.61	.79	1.53	.96	2.96
.63	0.66	.80	1.59	.97	3.14
.64	0.71	.81	1.65	.98	3.38
.65	0.77	.82	1.71	.99	3.76
.66	0.82	.83	1.78	1.00	$+\infty$

highly correlated observations (wide roving)—the PC_{max} curve corresponding to independent observations in Figure 2 (thick dark line) increases above that corresponding to highly correlated observations (thick gray line) as d' increases above about 1.4. However, the deviation between these two curves is not as large as that observed in other paradigms, in which the differencing strategy also differs from the optimal strategy—namely, the two-interval *same–different* paradigm and the 4IAX paradigm (Dai et al., 1996; Micheyl & Messing, 2006); the PC_{max} curves corresponding to these paradigms are shown as interrupted lines. In fact, for d' values lower than about 1 (which corresponds to a PC of about .7), the curves for uncorrelated and highly correlated observations in the 4IAX AB-versus-BA identification paradigm overlap almost perfectly. One practical implication of this observation is that, in experiments where d' values lower than 1 or PCs lower than about .7 are targeted, no significant difference in PC_{max} should be expected, depending on whether participants use the optimal strategy or the differencing strategy.

Table 2 lists d' values corresponding to PC_{max} values between .50 and 1 in steps of .01. Note that this table should be used only when the assumption of independent observations is warranted. For the case of highly correlated observations, readers can either use $PC_{max} = \Phi(d'/2)$ or refer to a yes–no paradigm table, which can be found in most psychophysical SDT textbooks (e.g., Macmillan & Creelman, 2005). It is important to note that the use of PC_{max} as a summary measure of performance is justified only when participants have no a priori bias toward the AB or BA response alternative and behave at least approximately as the above-described observer, operating on the basis of equal-variance Gaussian sensory observations.

Conclusions

In this article, an SDT model of the sensory and decision processes in the 4IAX AB-versus-BA identification paradigm was analyzed. Under the traditional assumption of equal-variance Gaussian observations, (likelihood ratio) optimal decision rules were derived for two cases: the case of statistically independent observations, which is likely to apply in experiments in which stimuli are not roved, and the case of highly correlated observations, which is likely to apply in experiments in which stimuli are roved widely across trials or pairs. Finally, the relationship between d' and PC_{max} was analyzed for these two cases.

Perhaps the most important conclusion of this work is that, although the optimal decision strategy depends on the degree of correlation between the sensory observations (and therefore, on whether roving is used), the relationships between d' and PC_{max} for the two limiting cases of independent and highly correlated observations do not differ appreciably. At most, the difference in PC_{max} is about 3% for a d' of about 2.7. This contrasts with other paradigms, such as the *same–different* or 4IAX paradigms, where using the differencing strategy instead of the independent-observations strategy can, according to the equal-variance Gaussian SDT model, result in substantially lower (>10%) PCs (Dai et al., 1996; Micheyl & Messing, 2006).

Another conclusion of practical importance is that, when the correlation between the sensory observations on a trial approaches 1, the relationship between d' and PC_{max} and the predicted ROC curves becomes the same as in the basic yes–no paradigm. Thus, when analyzing data from a 4IAX AB-versus-BA identification experiment where the stimuli are roved over a wide range, experimenters can use the tables and equations for the yes–no paradigm that can be found in most SDT textbooks. When one analyzes data from experiments using no stimulus roving, the formulas and tables for the yes–no paradigm can still be used for obtaining reasonably accurate approximations if d' is less than 1. If d' is larger than 1, or if more precise estimates are required, use of the tables provided in the present article is recommended.

AUTHOR NOTE

This work was supported by National Institutes of Health—National Institute on Deafness and Other Communication Disorders Grant R01 DC 05216, as well as by the University of Arizona. The authors are grateful to Douglas Creelman, Joshua Solomon, and an anonymous reviewer for many helpful suggestions on the manuscript, as well as to Laurent Demany and Samuel Mathias for stimulating discussions, which helped inspire the analyses described in this article. Correspondence concerning this article should be addressed to C. Micheyl, Department of Psychology, University of Minnesota, 75 East River Rd., Minneapolis, MN 55455-0344 (e-mail: cmicheyl@umn.edu).

REFERENCES

Creelman, C. D., & Macmillan, N. A. (1979). Auditory phase and frequency discrimination: A comparison of nine procedures. *Journal of Experimental Psychology: Human Perception & Performance*, **5**, 146-156.

Dai, H., Versfeld, N. J., & Green, D. M. (1996). The optimum decision rules in the *same–different* paradigm. *Perception & Psychophysics*, **58**, 1-9.

Demany, L., Carlyon, R. P., & Semal, C. (2009). Continuous versus discrete frequency changes: Different detection mechanisms? *Journal of the Acoustical Society of America*, **125**, 1082-1090.

Egan, J. P. (1975). *Signal detection theory and ROC analysis.* New York: Academic Press.

Green, D. M. (1988). *Profile analysis: Auditory intensity discrimination.* New York: Oxford University Press.
Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.
Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
Macmillan, N. A., Kaplan, H. L., & Creelman, C. D. (1977). The psychophysics of categorical perception. *Psychological Review*, **84**, 452-471.
Mathias, S. R., & Bailey, P. J. (2008, September). *Insensitivity to the direction of pitch change: The role of stimulus uncertainty and correlations with self-reported hearing*. Poster presented at the British Society of Audiology Meeting on Experimental Studies of Hearing and Deafness, University of York, York, England.
Micheyl, C., & Dai, H. (2008). A general area theorem for the *same–different* (2IAX) paradigm. *Perception & Psychophysics*, **70**, 761-764.
Micheyl, C., Kaernbach, C., & Demany, L. (2008). An evaluation of psychophysical models of auditory change perception. *Psychological Review*, **115**, 1069-1083.
Micheyl, C., & Messing, D. P. (2006). Likelihood ratio, optimal decision rules, and correct response probabilities in a signal detection theoretic, equal-variance Gaussian model of the observer in the 4IAX paradigm. *Perception & Psychophysics*, **68**, 725-735.
Noreen, D. L. (1981). Optimal decision rules for some common psychophysical paradigms. In S. Grossberg (Ed.), *Mathematical psychology and psychophysiology* (Vol. 13, pp. 237-279). Providence, RI: American Mathematical Society.
Rousseau, B., & Ennis, D. M. (2001). A Thurstonian model for the dual pair (4IAX) discrimination method. *Perception & Psychophysics*, **63**, 1083-1090.
Semal, C., & Demany, L. (2006). Individual differences in the sensitivity to pitch direction. *Journal of the Acoustical Society of America*, **120**, 3907-3915.
Versfeld, N. J., Dai, H., & Green, D. M. (1996). The optimum decision rules for the oddity task. *Perception & Psychophysics*, **58**, 10-21.
Wickens, T. D. (2001). *Elementary signal detection theory*. Oxford: Oxford University Press.

NOTE

1. It is important to note that this conclusion concerning the respective influence of across-trial and across-pair roving is based entirely on an ideal-observer analysis, which may not correctly describe the behavior of actual listeners. SDT is a normative theory, and ideal-observer analyses, such as the one described here, are only meant to provide an upper limit of performance. Empirical results can be compared with it in order to determine whether participants in an experiment behave like the ideal observer.

SUPPLEMENTAL MATERIALS

Mathematica code for deriving likelihood ratios in the dual-pair AB-versus-BA identification paradigm, further explanation of the optimal decision rules, and Monte Carlo simulations for calculating ROCs and PC_{max} for our data may be downloaded from app.psychonomic-journals.org/content/supplemental.

(Manuscript received December 29, 2008;
revision accepted for publication April 3, 2009.)