A computational technique for improving estimates of discriminability and bias across multiple dimensions of choice

GLENN S. BROWN AND K. GEOFFREY WHITE
University of Otago, Dunedin, New Zealand

Brown and White (2009) proposed measures of discriminability and bias that accommodate additional dimensions of choice—and hence, bias—in conditional discriminations such as matching-to-sample and the yes–no signal detection task. Their proposed measures increase the statistical independence of discriminability and bias estimates, thus improving their accuracy. Because Brown and White's (2009) equations partition response data more than do standard equations, however, their measures have a slightly lower ceiling. Consequently, measurements can be less accurate when there are few trials and discriminability and bias are extreme. We introduce a computational estimation technique that overcomes this limitation. It estimates Brown and White's (2009) discriminability and bias measurements from an array of related measures that have a higher ceiling. Simulations show that resulting estimates of discriminability and bias are either comparable to or more accurate than measurements calculated from traditional equations or Brown and White's (2009) direct measures, even with few trials. A worked example of our technique may be downloaded from brm.psychonomic-journals.org/content/supplemental.

Choice in conditional discriminations such as delayed matching-to-sample (Blough, 1959; White, Ruske, & Colombo, 1996) and the yes–no signal detection task (Macmillan & Creelman, 1991) is traditionally conceptualized as being one-dimensional. For example, the experimentally defined dimension for choice might be *red* versus *green*, or *yes* versus *no*. Brown and White (2009) argued that choice in such paradigms may be conceptualized as multidimensional. Depending on the procedure, other dimensions, such as choice side (*left* vs. *right*) or match to a priming stimulus (*congruent* vs. *incongruent*), could influence responding, in addition to an overall bias to choose *red* versus *green* or *yes* versus *no*. Brown and White (2009) showed that, if they do, standard measures of discriminability and bias will underestimate true performance and will not be statistically independent. They proposed alternative equations for such situations. Simulations and reanalyses of previous data confirmed that Brown and White's (2009) equations produced more accurate estimates of discriminability than did traditional measures, and to an extent that could alter the conclusions of an experiment. As we will show below, however, the disadvantage of their equations is that they have a lower measurement ceiling and range than do standard measures. This results from a finer partitioning of response data and, consequently, smaller cell frequencies. As a result, Brown and White's (2009) measures can be more susceptible to inaccuracies that arise from a small number of trials and extreme discriminability or bias (Brown & White, 2005; Hautus, 1995). In the present article, we introduce a computational estimation technique that overcomes this disadvantage.

Standard Choice Measurements in Conditional Discrimination Procedures

Choice in conditional discrimination tasks such as matching-to-sample is typically characterized along one dimension, B_1 versus B_2 (e.g., choice of *red* vs. *green*) following the presentation of a sample stimulus, S_1 or S_2 (red or green). *Discriminability* measures the tendency to choose the "correct" responses:[1] B_1 given S_1, and B_2 given S_2. Discriminability is commonly measured by calculating the ratio of correct to incorrect responses in each trial type and then taking the log of their geometric mean. Examples include choice theory's $\ln \alpha$ (Luce, 1963; McNicol, 1972), Nevin's (1981) $\log 1/\eta$, and Davison and Tustin's (1978) $\log d$. All are equivalent and directly proportional to each other (Brown & White, 2005). We focus on the commonly used $\log d$, but all conclusions apply equally to $\ln \alpha$ or $\log 1/\eta$. The equation for $\log d$ is

$$\log d = \frac{1}{2} \cdot \log_{10}\left(\frac{B_1 \mid S_1}{B_2 \mid S_1} \cdot \frac{B_2 \mid S_2}{B_1 \mid S_2} \right). \qquad (1)$$

A similar measure is signal detection theory's d' (Green & Swets, 1966). It has an almost linear relationship with

G. S. Brown, glenn@psy.otago.ac.nz; K. G. White, geoff.white@otago.ac.nz

log d and ln α (Brown & White, 2005; Macmillan & Creelman, 1991), but it is based on an inverse-normal transformation, rather than on the almost indistinguishable logit transformation. Although we focus on log d, the methods described in the present article may, in principle, also be applied to d' (it is, however, more computationally intensive; see the Appendix).

Choice on the experimenter-defined B_1–B_2 dimension is also measured by *response bias*, the tendency to choose B_1 over B_2 (or vice versa), regardless of whether it is correct. Response bias toward B_1 will increase accuracy on S_1 trials and will decrease accuracy on S_2 trials. Response bias is often measured in a way analogous to that for discriminability: Measure the tendency toward B_1 (i.e., the ratio of B_1 to B_2 responses) separately in each trial type and then take the log of the geometric mean:

$$\log b = \frac{1}{2} \cdot \log_{10}\left(\frac{B_1 \mid S_1}{B_2 \mid S_1} \cdot \frac{B_1 \mid S_2}{B_2 \mid S_2} \right). \quad (2)$$

This equation (Davison & Tustin, 1978) is equivalent to the bias equations for choice theory (Luce, 1963; McNicol, 1972) and Nevin's (1981) behavioral detection theory.

Multidimensional Choice and Performance Measurements

Brown and White (2009) argued that the choice between B_1 and B_2 may vary on more than one dimension. A common example is that the experimenter may define the relevant choice as one between red and green colors but may randomly vary the position of those choices from trial to trial between left and right side keys. Even though the left–right side of the choice should be irrelevant to the arranged task, it can nevertheless have a strong influence on choice: Subjects can exhibit *side bias* (a bias to left vs. right; see, e.g., Jones, 2003; Katz, 1989; Nevin & Grosch, 1990). When side bias—or any other unaccounted-for bias—is evident, it tends to deflate estimates of discriminability and bias and causes the two measures to become interdependent (Brown & White, 2009). This effect undermines a key goal of these analyses: to obtain statistically independent measurements of discriminability and bias (Macmillan & Creelman, 1991).

Brown and White (2009) proposed an equation for discriminability that accommodates the additional dimension of choice side or any equivalent dimension. With an additional dimension of choice such as side, there are four trial types rather than two: S_1 with B_1 on the left, S_1 with B_1 on the right, S_2 with B_2 on the left, and S_2 with B_2 on the right. Although, for clarity, we will focus on one concrete example, side bias, the same notion applies to other types of bias. In such cases, left versus right might be replaced by dimensions such as *same as* versus *different from* the prior trial's sample (proactive interference) or congruent versus incongruent with the response (priming). Using the side bias example, Brown and White (2009) used the same principles as in the standard discriminability equation: They took the ratio of correct to incorrect responses for each trial type, and found the log of the geometric mean:

$$\log d_{bs} = \frac{1}{4} \cdot \log_{10}\left(\frac{B_{1,\text{Left}} \mid S_1}{B_{2,\text{Right}} \mid S_1} \cdot \frac{B_{1,\text{Right}} \mid S_1}{B_{2,\text{Left}} \mid S_1} \right.$$
$$\left. \cdot \frac{B_{2,\text{Left}} \mid S_2}{B_{1,\text{Right}} \mid S_2} \cdot \frac{B_{2,\text{Right}} \mid S_2}{B_{1,\text{Left}} \mid S_2} \right), \quad (3)$$

where log d_{bs} is log d calculated to be independent of response bias (b) and side bias (s). Using this nomenclature, log d in Equation 1 may be written log d_b.

Brown and White's (2009) corresponding response bias equation that accounts for side and discriminability (d) dimensions is

$$\log b_{ds} = \frac{1}{4} \cdot \log\left(\frac{B_{1,\text{Left}} \mid S_1}{B_{2,\text{Right}} \mid S_1} \cdot \frac{B_{1,\text{Right}} \mid S_1}{B_{2,\text{Left}} \mid S_1} \right.$$
$$\left. \cdot \frac{B_{1,\text{Right}} \mid S_2}{B_{2,\text{Left}} \mid S_2} \cdot \frac{B_{1,\text{Left}} \mid S_2}{B_{2,\text{Right}} \mid S_2} \right), \quad (4)$$

and their equation for side bias that accounts for response bias and discriminability dimensions is

$$\log s_{bd} = \frac{1}{4} \cdot \log\left(\frac{B_{1,\text{Left}} \mid S_1}{B_{2,\text{Right}} \mid S_1} \cdot \frac{B_{2,\text{Left}} \mid S_1}{B_{1,\text{Right}} \mid S_1} \right.$$
$$\left. \cdot \frac{B_{2,\text{Left}} \mid S_2}{B_{1,\text{Right}} \mid S_2} \cdot \frac{B_{1,\text{Left}} \mid S_2}{B_{2,\text{Right}} \mid S_2} \right). \quad (5)$$

All of Brown and White's (2009) equations work in fundamentally the same way. They are based on the same principles as the standard discriminability and bias equations, but they accommodate an additional dimension of choice (and hence, bias). Although Brown and White (2009) pointed out that many dimensions of choice can be considered together, for clarity we will focus here on one additional choice dimension only.

Brown and White (2009) showed that when choice is significantly influenced by the additional choice dimension (e.g., side), the standard discriminability equation, log d_b (Equation 1), tends to underestimate discriminability overall, whereas their proposed discriminability equation, log d_{bs} (Equation 3), tends not to. Furthermore, when performance is side biased, log d_b measurements are dependent on both response bias and side bias, whereas log d_{bs} measurements are not (assuming no biases on other, unaccounted-for dimensions). The effects were illustrated in simulations, and reanalysis of previous data showed that they could be of sufficient magnitude to alter the conclusions of experiments (Brown & White, 2009). The Brown and White (2009) bias equations (Equations 4 and 5) are equally advantageous, being of the same mathematical form as the discriminability equation (Equation 3).

Disadvantages of the Multidimensional Approach

The potential disadvantage of Brown and White's (2009) equations is that they reduce the ceiling or range of obtainable measurements. This is a result of partitioning the data into a greater number of cells and, hence,

smaller cell frequencies. With the standard equations (1 and 2), there are four response counts: B_1 in S_1, B_2 in S_1, B_1 in S_2, and B_2 in S_2. To accommodate an additional choice dimension such as side, however, each response count must be divided into left and right B_1 and B_2 positions, thus producing eight response counts in total. This halves their frequency overall and thus lowers the maximum discriminability or bias measurement that can be obtained (Brown & White, 2005), even when a suitable correction technique is used to avoid indeterminate measurements (Brown & White, 2005; see also Hautus, 1995; Kadlec, 1999; Miller, 1996). As a consequence, the multidimensional discriminability and bias measurements are more likely to underestimate their true values when discriminability and bias are extreme and when the number of trials is low.

The outcome of the low-trial-number problem is illustrated in Figure 1. We simulated data for different underlying parameter values of discriminability, response bias, and side bias, following Brown and White (2009), but for a very small number of trials (40 in total; 10 of each type). For each different parameter–value combination, 1,000 sets of response counts were randomly selected from the full sampling distribution of possible response counts (Miller, 1996), weighted by their probability. For each data set, log d_b and log d_{bs} were calculated. To avoid indeterminate measurements, the commonly used correction factor of 0.5 was added to all response counts in the calculation, regardless of their value (Brown & White, 2005; after Goodman, 1970). The absolute difference between each discriminability value and the known, programmed discriminability parameter was found. These *misestimations* were averaged over the 1,000 data sets and are plotted with their standard deviations in Figure 1.

In Brown and White's (2009) main simulation, a large number of trials was used, and log d_{bs} consistently estimated discriminability as well as or better than did log d_b. In Figure 1, where a small number of trials was used, this was no longer true. The reason is that using a small number of trials lowered the ceiling value and range of log d_{bs} or log d_b well below many of the programmed discriminability and bias parameter values. For example, with 100 trials of each type, the maximum log d is about 2.3, whereas with 20 trials each, it is about 1.6 (using Goodman's, 1970, correction). That is, the measurement ceiling is lower with fewer trials. Since log d_{bs} has a lower ceiling range than does log d_b, its underestimates tended to be larger when discriminability and response bias were extreme. (Log d_{bs} was still more likely to outperform log d_b when side bias was strong, however.)

This effect does not occur if there are a sufficient number of trials. So, an obvious way of avoiding the problems associated with using Brown and White's (2009) multidimensional equations is to run more trials. Here, however, we present a mathematical procedure that renders this time- and resource-consuming solution unnecessary. Although it might seem involved at first, it is in fact easy to use, once it is incorporated into an electronic spreadsheet or similar tool.

THE COLLAPSED MATRIX INFERENCE TECHNIQUE

Overview

The eight response counts required to calculate log d_{bs} (Equation 3) come from a $2 \times 2 \times 2$ matrix (S_1 vs. $S_2 \times B_1$ vs. $B_2 \times$ left vs. right). The estimation technique described below avoids the problems of a smaller range of obtainable measurements from these eight-cell matrices by basing its calculations on three 2×2 matrices instead. The three 2×2 matrices are constructed by collapsing the $2 \times 2 \times 2$ matrix in three different ways. All data contribute to each matrix. Because each matrix contains only four cells, response counts are higher, and the range of obtainable measurements is greater. If in-cell response counts of zero and, hence, ceiling ratios do occur, their effects are diluted, because the analysis is based on three matrices rather than one. Unfortunately, log d_{bs} (Equation 3), log b_{ds} (Equation 4), and log s_{bd} (Equation 5) cannot be calculated directly from these three matrices. It is possible, however, to infer their values by comparing response measurements from these empirical matrices with those from similar, but programmed, matrices in which the underlying discriminability, response bias, and side bias parameters are known. If the empirical matrices produce the same discriminability and bias measurements as the programmed matrices, we infer that the empirical and programmed matrices are based on the same underlying parameter values. Step-by-step details of this procedure, which we refer to as the *collapsed matrix inference technique*, will be shown below. Furthermore, a Microsoft Excel workbook that provides an automated example of the procedure may be downloaded in the online supplement.

Step 1: Input Data Into an Extended Signal Detection Matrix

A standard, 2×2 signal detection matrix consists of two stimuli (S_1 or S_2) $\times$ two responses (B_1 or B_2). When an additional dimension of choice, such as side, is included, the matrix becomes a $2 \times 2 \times 2$ matrix. It still consists of two stimuli (S_1 or S_2) and two corresponding "correct" choices (B_1 and B_2) but also consists of two response sides (left and right). It is easier and clearer to display this $2 \times 2 \times 2$ matrix as a 2×4 matrix, however. Such an extended matrix is shown in Table 1. In Table 1, as in the subsequent examples, S_1 and S_2 are red and green stimuli, and B_1 and B_2 are choices of red and green, respectively. The same type of matrix applies to other stimuli and choices, but we will use these concrete examples for clarity.

In the extended signal detection matrix shown in Table 1, there are eight cells. Observed response counts should be entered into the appropriate cells. The eight cells represent

Table 1
Extended Signal Detection Matrix

Sample	Response			
	Red, Left	Red, Right	Green, Left	Green, Right
Red	A	B	C	D
Green	E	F	G	H
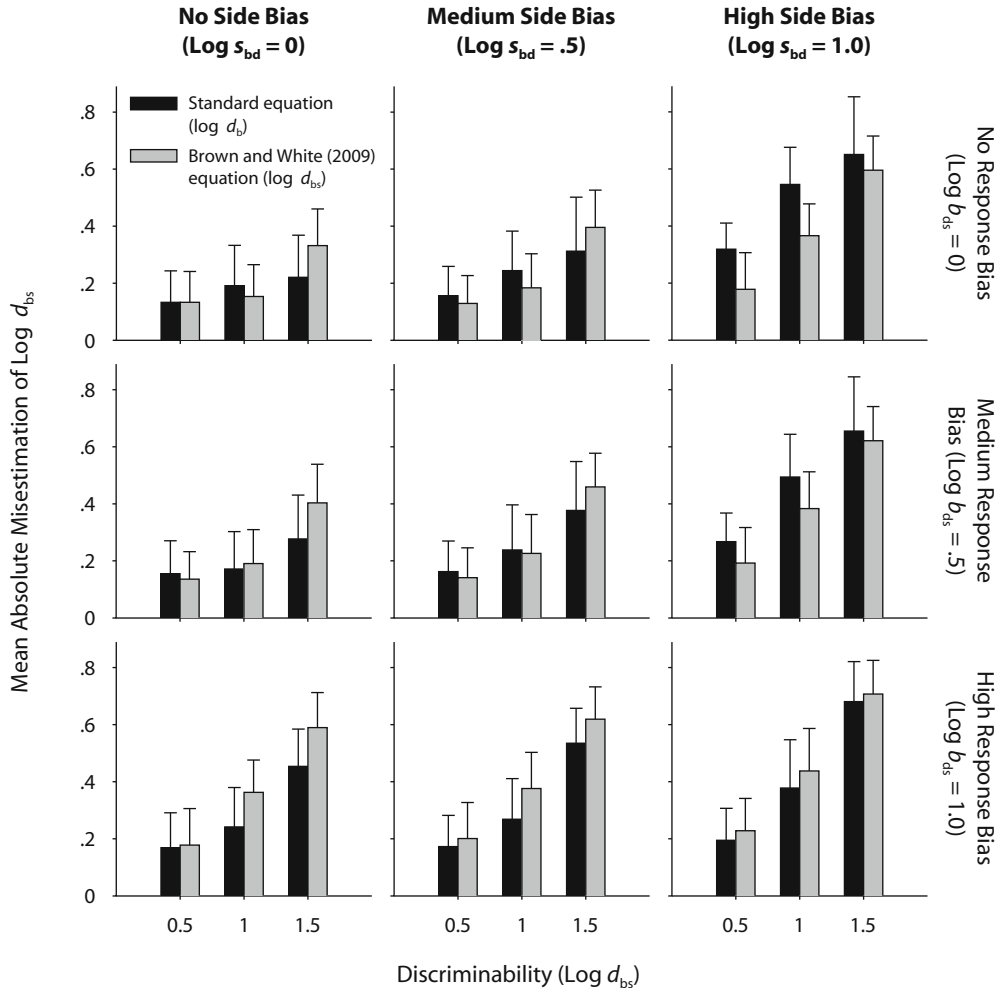
Figure 1. Simulated misestimations of discriminability for a very small number of trials for the standard discriminability measure, log d_b (black bars), and Brown and White's (2009) multidimensional discriminability measure, log d_{bs} (gray bars). Each condition constitutes 1,000 simulations of response counts, each consisting of 10 trials of each of the four types. The left column of the graphs shows data for simulations in which there was no inherent side bias (log $s_{bd} = 0$), the middle column for medium side bias (log $s_{bd} = .5$), and the right column for high side bias (log $s_{bd} = 1$). Similarly, the top row of the graphs is for no response bias (log $b_{ds} = 0$), the middle row for medium response bias (log $b_{ds} = .5$), and the bottom row for high response bias (log $b_{ds} = 1$). For each of the nine graphs, programmed discriminability was low (log $d_{bs} = .5$), medium (log $d_{bs} = 1$), or high (log $d_{bs} = 1.5$). Error bars show one standard deviation.

four sets of concurrently available options. For example, "correct" red responses on the left are available at the same time as "incorrect" green responses on the right. These cells therefore complement each other. Cells A and D, B and C, E and H, and F and G in Table 1 are all pairs of complementary cells. Since each pair contains all of the responses for a particular trial type, they are the basis of the four ratios used when estimating discriminability and bias parameters.

Step 2: Construct a Programmed Extended Signal Detection Matrix

The key component of the present procedure is to compare response measurements from empirical data with corresponding measurements generated from programmed parameters. Accordingly, every matrix and equation used with actual data must have a programmed counterpart. Since Step 1 constructed an extended matrix for observed data,

Step 2 must construct the corresponding matrix generated from programmed parameter values. Table 2 shows how log d_{bs}, log b_{ds}, and log s_{bd} parameters can be programmed into such a matrix in order to generate hypothetical data. It assumes that a positive color bias is toward red and a positive side bias is toward left. We generated this matrix by working back from the equation for log d_{bs} (Equation 3), its

Table 2
Extended Signal Detection Matrix
With Programmed Parameters

	Response			
Sample	Red, Left	Red, Right	Green, Left	Green, Right
Red	$d_{bs} \times b_{ds} \times s_{bd}$	$d_{bs} \times b_{ds}$	s_{bd}	1
Green	$b_{ds} \times s_{bd}$	b_{ds}	$d_{bs} \times s_{bd}$	d_{bs}

Table 3
Normalized Extended Signal Detection Matrix

	Response			
Sample	Red, Left	Red, Right	Green, Left	Green, Right
Red	$\dfrac{A}{(A+D)}$	$\dfrac{B}{(B+C)}$	$\dfrac{C}{(B+C)}$	$\dfrac{D}{(A+D)}$
Green	$\dfrac{E}{(E+H)}$	$\dfrac{F}{(F+G)}$	$\dfrac{G}{(F+G)}$	$\dfrac{H}{(E+H)}$

Table 4
Normalized Programmed Extended Signal Detection Matrix

	Response			
Sample	Red, Left	Red, Right	Green, Left	Green, Right
Red	$\dfrac{d_{bs} \times b_{ds} \times s_{bd}}{\left(d_{bs} \times b_{ds} \times s_{bd}\right)+1}$	$\dfrac{d_{bs} \times b_{ds}}{\left(d_{bs} \times b_{ds}\right)+s_{bd}}$	$\dfrac{s_{bd}}{\left(d_{bs} \times b_{ds}\right)+s_{bd}}$	$\dfrac{1}{\left(d_{bs} \times b_{ds} \times s_{bd}\right)+1}$
Green	$\dfrac{b_{ds} \times s_{bd}}{\left(b_{ds} \times s_{bd}\right)+d_{bs}}$	$\dfrac{b_{ds}}{\left(d_{bs} \times s_{bd}\right)+b_{ds}}$	$\dfrac{d_{bs} \times s_{bd}}{\left(d_{bs} \times s_{bd}\right)+b_{ds}}$	$\dfrac{1}{\left(b_{ds} \times s_{bd}\right)+d_{bs}}$

analogous equation for response bias (log b_{ds}; Equation 4), and the formula for side bias (log s_{bd}; Equation 5).

Table 2 contains the antilog of the discriminability and bias parameters. In any "correct" cell, d_{bs} (the antilog of the discriminability parameter, log d_{bs}) appears. In any red response cell, b_{ds} appears, and in any left response cell, s_{bd} appears. If a cell contains more than one term, those terms are multiplied together. The only cell containing no terms simply has a value of one. (A slightly different method is required when measurements are based on signal detection theory's sensitivity and response criteria parameters; see the Appendix.)

Step 3: Normalize the Extended Signal Detection Matrices

Cells in the extended matrix must be combined as part of the present analysis. Because of this, it may be necessary to normalize the contents of the extended matrix in order to ensure that the subsequent analysis is weighted evenly over trial types. Empirical matrices need to be normalized only if there are unequal numbers of trials across trial type. Programmed matrices, on the other hand, must always be normalized. Normalization converts the contents of a cell into a proportion of itself and its complementary cell (the other cell of the same trial type). This is shown explicitly in Table 3. (Note that although later we will refer to the cell labels in Table 1, the equivalent cells in Table 3 should be used if normalization is required.) For the programmed extended matrix, Table 4 specifies the appropriate calculations.

Step 4: Generate 2 × 2 Matrices From the Extended Matrices

An important part of the present procedure is to use three 2 × 2 matrices, rather than one extended matrix. Doing this increases the range or ceiling of obtainable measurement values. These three matrices are shown in Tables 5, 6, and 7. Table 5 is the standard signal detection matrix in which responses are collapsed over left and right sides. Table 6 is a similar matrix, except that it is collapsed

over red and green responses, rather than over left and right responses. Table 7 follows the same principles, except that, quite unusually, it is collapsed over sample stimulus. Each table uses the cell labels defined in Table 1. As for all other steps, it is necessary to generate both empirical matrices and their programmed counterparts. Both are calculated from their parent extended matrix in exactly the same way. (Unless the number of trials is the same for each trial type, normalized matrices should be used in both cases.)

Step 5: Calculate Bias Values From the 2 × 2 Matrices

In the present procedure, empirical and programmed 2 × 2 matrices are compared by using the response measures associated with each matrix. For example, "mea-

Table 5
2 × 2 Signal Detection Matrix Collapsed Over Side of Response

Correct Choice	Response	
	Red	Green
Red	A + B	C + D
Green	E + F	G + H

Table 6
2 × 2 Signal Detection Matrix Collapsed Over Response Color

Correct Choice	Response	
	Left	Right
Left	A + G	D + F
Right	C + E	B + H

Table 7
2 × 2 Signal Detection Matrix Collapsed Over Stimulus Color

Choice Side	Response Color	
	Red	Green
Left	A + E	C + G
Right	B + F	D + H

sured" values of log d_b and log b_d (see Equations 6 and 7) are calculated from the empirical and the programmed 2×2 matrix that is collapsed over response side. The response measurements from the empirical and programmed versions of Tables 5, 6, and 7 are compared with each other. For example, log d_b from the empirical matrices is compared with log d_b from the programmed matrices.

It might seem odd to examine measurements from collapsed matrices, since they may be contaminated by bias or discriminability (Brown & White, 2009). The first reason that this is not a limitation here is that three matrices—each collapsed in a different way—are examined together rather than individually. The second reason is that response measures based on collapsed matrices are only part of an intermediate step. They are used to check whether there is a match between empirical measurements and corresponding values based on known, programmed parameter values. As long as response measures from all three matrices are used, no dimension of choice or bias is ignored in the comparison.

From Table 5, the standard signal detection matrix in which responses are collapsed over response side, the discriminability equation is

$$\log d_b = \frac{1}{2} \cdot \log\left(\frac{A+B}{C+D} \cdot \frac{G+H}{E+F}\right), \qquad (6)$$

where log d_b measures discriminability, taking response bias into account. The terms A through H represent response counts from Table 1. From Table 5, the response bias equation is

$$\log b_d = \frac{1}{2} \cdot \log\left(\frac{A+B}{C+D} \cdot \frac{E+F}{G+H}\right), \qquad (7)$$

where log b_d measures response bias, taking discriminability into account. The equivalent measures from Table 6, in which responses are collapsed over color, are

$$\log d_s = \frac{1}{2} \cdot \log\left(\frac{A+G}{D+F} \cdot \frac{B+H}{C+E}\right), \qquad (8)$$

where log d_s measures discriminability, taking side bias into account, and

$$\log s_d = \frac{1}{2} \cdot \log\left(\frac{A+G}{D+F} \cdot \frac{C+E}{B+H}\right), \qquad (9)$$

where log s_d measures side bias, taking discriminability into account.

In Table 7, responses are collapsed over stimulus type. This simply means that discriminability—which represents the behavioral tendency to choose the correct response—must be ignored. Two bias equations can still be constructed, though: response (color) bias and side bias. To calculate response bias in this case,

$$\log b_s = \frac{1}{2} \cdot \log\left(\frac{A+E}{D+H} \cdot \frac{B+F}{C+G}\right), \qquad (10)$$

where log b_s is an estimate of response bias, taking side bias into account. Similarly,

$$\log s_b = \frac{1}{2} \cdot \log\left(\frac{A+E}{D+H} \cdot \frac{C+G}{B+F}\right), \qquad (11)$$

where log s_b represents side bias, taking response bias into account. Equivalent calculations for signal detection theory's sensitivity and response criteria parameters are shown in the Appendix.

## Step 6: Compare Data-Based Measurements With Hypothetical Measurements

At this stage, there are six response measures based on empirical data and six based on the programmed matrices. If the empirical measurements are all very similar to their programmed counterparts, it suggests that the empirical measurements are based on the same underlying parameters as the measurements derived from the programmed matrices. In other words, they arise from the same log d_{bs}, log b_{ds}, and log s_{bd} parameters that are known and were entered into the programmed extended matrix (Tables 2 and 4).

If the empirical measurements differ somewhat from their programmed counterparts, the programmed log d_{bs}, log b_{ds}, and log s_{bd} parameters should be adjusted until they do not. Our method of doing this is to use electronic spreadsheet software. Using this software, we set out all of the matrices and equations described in Steps 1–5 and set the programmed, underlying parameter values of d_{bs}, b_{ds}, and s_{bd} to zero. We then use Microsoft Excel's Solver to adjust those parameter values (and consequently, the response measures based on the programmed 2×2 matrices) until the sum of squares of the differences between the 2×2 empirical response measurements and their programmed counterparts is minimized. When this happens, estimates of log d_{bs}, log b_{ds}, and log s_{bd} values that underpin the empirical data can be inferred from the known parameters in the programmed matrix.

EVALUATION

We used simulations to evaluate how well the collapsed matrix inference technique estimates log d_{bs} when there is a small number of trials. The same method and parameters were used as in the simulations above (see Brown & White, 2009, for methodological details), except that the simulations were based on 64 trials, a minimum of what might be expected in experiments. In addition, log d_{bs} was calculated both from Equation 3 and from the collapsed matrix inference technique. The average misestimations of each method and of the standard log d_b equation are shown in Figure 2.

Figure 2 shows that, even with no side bias, the collapsed matrix inference technique produced discriminability estimates that were comparable to, or more accurate than, those generated using the standard discriminability measure, log d_b. When side bias was high, the technique consistently produced much more accurate estimates of discriminability than did the standard measure. Under conditions of low trial number and medium-to-strong discriminability and bias, the technique also tended to produce more accurate estimates of discriminability than did Brown and White's (2009) log d_{bs} equation. An important conclusion is that measuring discriminability with log d_{bs} from the collapsed matrix inference technique either produces better estimates of discriminability than do
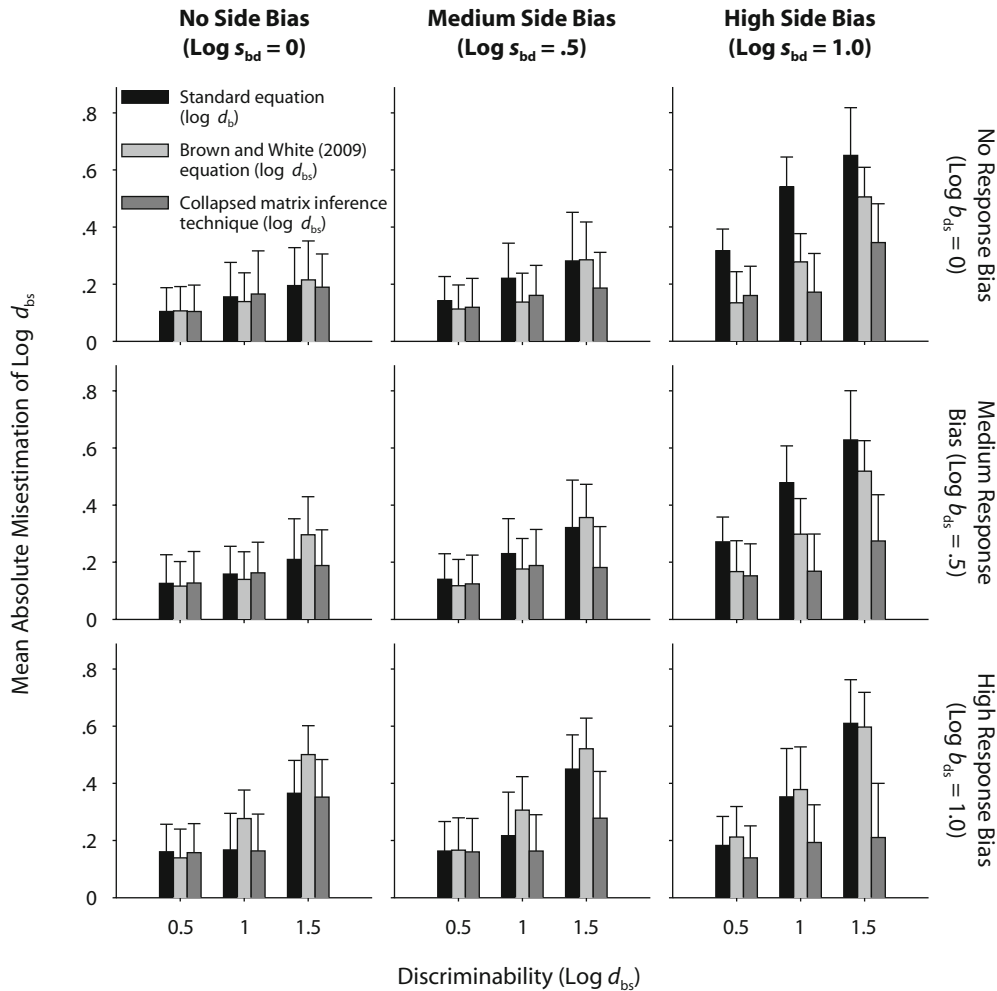
Figure 2. Simulated misestimations of discriminability for a small number of trials (16 of each of the four types) for the standard discriminability measure (log d_b; black bars), Brown and White's (2009) multidimensional discriminability measure (log d_{bs}; light gray bars), and log d_{bs} calculated using the present collapsed matrix inference technique. The layout and underlying methodology are otherwise the same as those in Figure 1.

the equations for log d_b or log d_{bs} or, at worst, has no clear disadvantages.[2]

The only significant cost that we have identified with the collapsed matrix inference technique is its greater computational overhead. Once the technique has been incorporated into an automated electronic format, however, its time cost is negligible. For example, in the Microsoft Excel workbook available as an online supplement to this article, parameter values can be calculated almost instantaneously and at the click of a button. We therefore suggest that the present technique can be used in place of Brown and White's (2009) direct equations even when it is not necessary to do so. Doing so incurs few costs but can, in some circumstances, greatly improve measurements of discriminability.

DISCUSSION

When choice in conditional discriminations is influenced by more than one dimension (e.g., both color and side), standard estimates of discriminability and bias tend to be less accurate than those produced by Brown and White's (2009) multidimensional equations. The limitation of Brown and White's (2009) equations is that, with low trial frequencies, they have a lower measurement ceiling and, hence, are more susceptible to the effects of extreme discriminability and bias (Brown & White, 2005). The collapsed matrix inference technique removes this limitation. Its discriminability estimates are consistently comparable to, or more accurate than, those produced by either the standard equations or Brown and White's (2009) multidimensional equations.

The conclusion applies both to estimates of discriminability and to estimates of bias (e.g., response bias and side bias): Their relationship is symmetrical, and the form of their equations is identical. Furthermore, the conclusion applies to other dimensions of choice (side bias is merely one example). For example, Brown and White (2009) suggested that other relevant choice dimensions may be congruency of a priming stimulus with the sample (congruent vs. incongruent) and match to the prior trial's

sample (same vs. different). The latter dimension is relevant in studies of proactive interference (D'Amato, 1973; Grant, 1975, 1982). Both are amenable to the analyses presented here. They can be incorporated into Brown and White's (2009) general multidimensional equation:

$$\log y = \frac{1}{N} \cdot \sum_{k=1}^{N} \log_{10}\left(\frac{y_1 \text{ Responses in Trial Type } k}{y_2 \text{ Responses in Trial Type } k}\right), (12)$$

where y is the choice dimension (e.g., correct vs. incorrect, red vs. green, left vs. right, or same vs. different), y_1 is one value of that choice dimension (e.g., correct, red, left, or same), and y_2 is its complementary value (e.g., incorrect, green, right, or different). N represents the number of trial types, with each type individually denoted by a value of k, and is determined by the number of choice dimensions. The additional dimension or dimensions can be incorporated into the collapsed matrix inference technique by following the principles outlined in the present article. The collapsed matrix inference technique thus represents a general technique for estimating bias and discriminability for any feasible number of choice dimensions.

Although the present technique works well, there may be different techniques that produce better discriminability estimates. One promising candidate is maximum likelihood estimation (e.g., Myung, 2003). It specifically estimates the discriminability and bias parameters that are most likely to have produced observed data. We have attempted to use this technique in the present context. In our simulations to date, maximum likelihood estimation produced many discriminability estimates that were slightly more accurate than those of the collapsed matrix inference technique. Unfortunately, many more of its estimates were rather inaccurate, particularly when discriminability and bias were strong. We believe that this problem might be overcome by using a more complex optimization tool than Solver, one that can resolve global, instead of local, maxima. Potential optimization techniques include simulated annealing, tabu search, and evolutionary algorithms (Michalewicz & Fogel, 2004). Such techniques are more difficult to implement than the easily available Solver. The collapsed matrix inference technique, on the other hand, is a relatively simple technique for producing good estimates of discriminability and bias when there are multiple dimensions of choice.

AUTHOR NOTE

REFERENCES

Blough, D. S. (1959). Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior*, **2**, 151-160.

Brown, G. S., & White, K. G. (2005). The optimal correction for estimating extreme discriminability. *Behavior Research Methods*, **37**, 436-449.

Brown, G. S., & White, K. G. (2009). Measuring discriminability when there are multiple sources of bias. *Behavior Research Methods*, **41**, 75-84.

D'Amato, M. R. (1973). Delayed matching and short-term memory in monkeys. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 7, pp. 227-269). New York: Academic Press.

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*, **51**, 291-315.

Goodman, L. A. (1970). The multivariate analysis of qualitative data: Interactions among multiple classifications. *Journal of the American Statistical Association*, **65**, 226-256.

Grant, D. S. (1975). Proactive interference in pigeon short-term memory. *Journal of Experimental Psychology: Animal Behavior Processes*, **3**, 207-220.

Grant, D. S. (1982). Samples of stimuli, responses, and reinforcers: Effect of incongruent sample type, serial position, and mode of presentation. *Animal Learning & Behavior*, **10**, 7-14.

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

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d'. *Behavior Research Methods, Instruments, & Computers*, **27**, 46-51.

Jones, B. M. (2003). Quantitative analyses of matching-to-sample performance. *Journal of the Experimental Analysis of Behavior*, **79**, 323-350.

Kadlec, H. (1999). Statistical properties of d' and β estimates of signal detection theory. *Psychological Methods*, **4**, 22-43.

Katz, J. L. (1989). Two types of bias in psychophysical detection and recognition procedures: Nonparametric indices and effects of drugs. *Psychopharmacology*, **97**, 202-205.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 103-189). New York: Wiley.

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

McNicol, D. (1972). *A primer of signal detection theory.* London: Allen & Unwin.

Michalewicz, Z., & Fogel, D. B. (2004). *How to solve it: Modern heuristics* (2nd ed.). New York: Springer.

Miller, J. (1996). The sampling distribution of d'. *Perception & Psychophysics*, **58**, 65-72.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, **47**, 90-100.

Nevin, J. A. (1981). Psychophysics and reinforcement schedules: An integration. In M. L. Commons & J. A. Nevin (Eds.), *Quantitative analyses of behavior: Vol. 1. Discriminative properties of reinforcement schedules* (pp. 3-27). Cambridge, MA: Ballinger.

Nevin, J. A., & Grosch, J. (1990). Effects of signaled reinforcer magnitude on delayed matching-to-sample performance. *Journal of Experimental Psychology: Animal Behavior Processes*, **16**, 298-305.

White, K. G., Ruske, A. C., & Colombo, M. (1996). Memory procedures, performance and processes in pigeons. *Cognitive Brain Research*, **3**, 309-317.

NOTES

1. Technically, "correct" versus "incorrect" is a dimension of choice additional to red versus green. For the sake of clarity and simplicity, we largely avoid referring to it as a choice dimension in the main text. This dimension is, however, inherent in both the traditional measures of discriminability and bias and Brown and White's (2009) multidimensional equations.

2. The simulations reveal similar patterns regardless of the number of trials (cf. Brown & White, 2005), but the advantage of the collapsed matrix inference technique over Brown and White's (2009) direct equations tends to become smaller as the number of trials increases and the range of obtainable measurements improves.

SUPPLEMENTAL MATERIAL

A worked example of the analysis in this study may be downloaded from brm.psychonomic-journals.org/content/supplemental.

APPENDIX
Application to Signal Detection Theory

The present measures and techniques may be translated into terms more familiar to users of signal detection theory. Brown and White (2009, Appendix A) provide the multidimensional equations for d' and response criteria that correspond to the present Equations 3, 4, and 5. The signal detection equations that are direct mathematical translations of the present Equations 6–11 appear below.

Discriminability or sensitivity, as measured from Table 5, is

$$d'_b = \Phi^{-1}\left(\frac{A+B}{A+B+C+D}\right) + \Phi^{-1}\left(\frac{G+H}{E+F+G+H}\right), \tag{A1}$$

where d'_b measures sensitivity, taking response bias into account. Terms A through H represent response counts in Table 1. This equation corresponds to the standard equation for d'. The related equation for the standard response criterion along the response bias dimension that takes sensitivity into account, $c_{d(b)}$, is

$$c_{d(b)} = -\frac{1}{2} \cdot \left(\Phi^{-1}\left(\frac{A+B}{A+B+C+D}\right) + \Phi^{-1}\left(\frac{E+F}{E+F+G+H}\right)\right). \tag{A2}$$

The equivalent measures from Table 6, in which responses are collapsed over color, are

$$d'_s = \Phi^{-1}\left(\frac{A+G}{A+D+F+G}\right) + \Phi^{-1}\left(\frac{B+H}{B+C+E+H}\right), \tag{A3}$$

where d'_s measures sensitivity, taking side bias into account, and

$$c_{d(s)} = -\frac{1}{2} \cdot \left(\Phi^{-1}\left(\frac{A+G}{A+D+F+G}\right) + \Phi^{-1}\left(\frac{C+E}{B+C+E+H}\right)\right), \tag{A4}$$

where $c_{d(s)}$ measures the response criterion along the side dimension, taking sensitivity into account.

The two bias equations from Table 7 are

$$c_{s(b)} = -\frac{1}{2} \cdot \left(\Phi^{-1}\left(\frac{A+E}{A+D+E+H}\right) + \Phi^{-1}\left(\frac{B+F}{B+C+F+G}\right)\right), \tag{A5}$$

where $c_{s(b)}$ is an estimate of the response criterion along the response bias dimension, taking side bias into account. Similarly,

$$c_{b(s)} = -\frac{1}{2} \cdot \left(\Phi^{-1}\left(\frac{A+E}{A+D+E+H}\right) + \Phi^{-1}\left(\frac{C+G}{B+C+F+G}\right)\right) \tag{A6}$$

is an estimate of the response criterion, $c_{b(s)}$, along the side dimension, taking response bias into account.

Other than these equations, the most important difference between applying the present technique to log d and to d' is the method used to program known parameters into the extended signal detection matrix (see Tables 2 and 4). With log d, known parameters can be used in equations to generate proportions in the extended signal detection matrix. Owing to the nature of its calculation, we are not aware of a corresponding method for d'. Our alternative is to adjust the proportions in the programmed extended signal detection matrix directly. We use a program like Microsoft Excel's Solver to adjust, within suitable bounds (i.e., proportions greater than 0 and less than 1), the values of four cells that are independent of each other in the extended signal detection matrix, such as the four cells on the left of Table 3. We use starting values of .5, representing indifference. The goal of the adjustment is to minimize the sum-of-squares difference between the values of Equations A1 through A6 derived from the programmed matrix and the corresponding values derived from the empirical matrix. (Note that for this optimization, the values of the sensitivity Equations A1 and A3 should be halved so that they are scaled the same as the bias Equations A2, A4, A5, and A6.) The technique is laid out in a Microsoft Excel spreadsheet that can be downloaded in the online supplement.