Models of Unforced Choice

Víctor H. Cervantes and Aaron S. Benjamin

University of Illinois Urbana-Champaign

Author Note

Víctor H. Cervantes https://orcid.org/0000-0001-9525-3454

Correspondence concerning this article should be addressed to Víctor H. Cervantes, Department of Psychology, University of Illinois Urbana-Champaign, 603 E Daniel St., Champaign, IL 61820. E-mail: victorhc@illinois.edu

We would like to dedicate this article to the memory of Thomas D. Wickens. Tom contributed as much as anyone in advancing the use of multivariate geometry in the development of signal detection theory, trained generations of students who loved him dearly at UCLA and Berkeley, and passed away much too young. VC is also grateful to Ehtibar N. Dzhafarov for many constructive conversations about contextuality and selective influences which influenced the developments of the covariance matrix structure.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the University of Illinois.

Abstract

Unforced choice tasks are ones in which the responder has the option of selecting from a limited array of choices or rejecting the entire set. Such tasks are common in perceptual and cognitive research, but models of decision-making for unforced choice are sparse in the literature and lack a unifying mathematical framework. Such tasks are important for theoretical development because they contain elements of relative, criterion-independent decision making—in which response options are compared to *one another*—and also criterion-dependent decision making, in which options are compared to a *decision criterion* determined by the observer. We provide a generic signal-detection multivariate framework for developing models of unforced choice that draws lessons from the geometry of multivariate statistics, from multidimensional signal-detection theories, and from psychophysical models of visual search. We show how this framework can accommodate all extant models of unforced choice that have been applied to the specific case of *lineup memory* tasks for eyewitnesses. Exact derivations, some of which have proven elusive to this point, are provided for each model in both an unrestricted form—in which variances and covariances are relatively free to vary across signal and noise—and in various restricted forms, in which constraints are applied to variances and covariances. Using the formalizations presented here, we show that all of the current models of lineup memory have severe limitations that render some models challenging to directly compare to one another and other models unidentifiable. Overall, the multivariate framework will aid in the testing of current models and in the development of new models of lineup memory and other tasks involving unforced choice.

Keywords: signal detection theory, SDT, multivariate SDT, unforced choice, non-forced choice, forced choice, lineup memory, independent observations model, dependent observations model, ensemble model, psychophysics, visual search

Models of Unforced Choice

Detection, discrimination, and identification tasks in which subjects choose from a set of predetermined options are among the most popular and widely used tools in psychology and related fields. One common implementation uses *forced-choice* responding, which obliges the subject to choose from that limited set to the best of their ability. In research on perception and cognition, the use of these forced-choice tasks benefits from a straightforward application of Signal Detection Theory (SDT; Green & Swets, 1966), a tool that aids understanding of behavior in tasks that involve decision-making under conditions of uncertainty. This fruitful marriage of popular empirical paradigms with well-grounded theoretical tools dates back to some of the earliest papers on SDT (e.g., Green, 1964), which developed and presented a natural relationship between yes/no response tasks and forced-choice response tasks. That relationship is based on simple principles in multivariate geometry and has held up well over time (e.g., Green & Moses, 1966; Jang, Wixted, & Huber, 2009; Jesteadt & Bilger, 1974; Wickelgren, 1968). The resultant models have been a major success for SDT and also a driving force underlying the development of more complex, higher-dimensional models of forced-choice decision-making.

One way to think about the difference between a forced-choice task and a yes/no response task is the nature of the decision rule. The key assumption of SDT is the idea that yes/no judgments are made by comparing a noisy percept or noisy memory to a stable decision criterion (cf. Benjamin, Diaz, & Wee, 2009); if the evidence yielded by that percept surpasses the criterion, a positive decision is made. Forced-choice tasks simplify the theoretical situation because observers simply choose the stimulus out of the response set that elicits the most evidence. There is no need for the responder to derive, set, or maintain a decision criterion (Benjamin & Bawa, 2004; Rotello & Macmillan, 2007).

The simplicity of this conceptual apparatus and the success of the models relating yes/no and forced-choice recognition have led to a notable gap in the literature, however.

Here we call these tasks *unforced choice*.[1] Any task in which the observer has to select from among a limited set of options, or to reject the entire set of options altogether, is a task of unforced choice. They are neither fish nor fowl in the language of standard SDT because they simultaneously rely on criterion-independent decision-making (in comparing the response options with one another) and criterion-dependent decision-making (in choosing whether to select an option or to reject the array).

Unforced choice tasks are widely used in psychology. In visual search paradigms, for example, it is common to ask the observer to search for one of a designated set of objects, with some trials including none of those objects (e.g., Apelt & Peitgen, 2008; Perry & Barron, 2013; Steinmetz, Zatka-Haas, Carandini, & Harris, 2019; Zatka-Haas, Steinmetz, Carandini, & Harris, 2021). A baggage screener faces a similar task: to identify whether one of a small set of contraband objects is present in a suitcase—but, in most cases, none of those objects are there (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004).

Although tasks involving unforced choice are common and related in spirit to tasks requiring forced choice, the application of SDT to these tasks has been incomplete and lacks an overarching theoretical framework. Models have been proposed both outside of the aegis of signal-detection theory (Clark, 2003; Wolfe, 1994; Wolfe, Cave, & Franzel, 1989), and within it (e.g., Kaernbach, 2001; Phelps, Rand, & Ryan, 2006; Taylor & Fraser, 1966; Watson, Kellogg, Kawanishi, & Lucas, 1973); most recently, advances in models of eyewitness memory have motivated bespoke signal-detection models for lineup procedures (Akan, Robinson, Mickes, Wixted, & Benjamin, 2021; Duncan, 2006; Wixted & Mickes, 2014; Wixted, Vul, Mickes, & Wilson, 2018).

Here, we provide a general multivariate framework by extending the well validated principles of forced-choice SDT models to cases involving unforced choice. Along the way, we highlight the mathematical challenges that accompany this transition. The technical

---

[1] These tasks are also sometimes denoted *non-forced choice*, *simultaneous detection and identification*, or *compound decision* tasks, though there is considerable variability in the use of these and related terms.

complexity notwithstanding, the principles of the generalization are quite generic and the resulting framework can accommodate a large family of theories of unforced choice. We feature the most prominent examples of such theories in the third major portion of this paper.

We use as a model case throughout this paper the example of a *lineup* as a test of eyewitness memory—the witness may select one of the faces in the lineup as the perpetrator of a previously witnessed crime but also understands that the perpetrator may not be in the lineup and so may reject the entire array. Recent theoretical advances in signal-detection models of eyewitness memory help provide the organization of this paper; within the multivariate framework we present, we derive all of the extant models of eyewitness lineup memory and present characteristics of those models that will aid in adjudicating among them empirically. We derive important relationships among the competing models that reduce the set of unique alternatives, and also point the way forward to the development of new models using this framework.

The overarching goal of this work is to bring tasks of unforced-choice into the language of SDT, so that research using those tasks can enjoy the same mutually beneficial relationship with SDT as do models of forced-choice. In so doing, we draw from relevant literatures on the geometry of multivariate statistics (Tong, 1990; Wickens, 2014), from multidimensional models of signal detection (Ashby, 2014; Ashby & Soto, 2015; Banks, 2000; Thomas, Altieri, Silbert, Wenger, & Wessels, 2015), from psychophysical models of visual search (Eckstein, 1998; Palmer, 1994), and, most recently, from models of lineup memory for eyewitnesses (Akan et al., 2021; Duncan, 2006; Wixted et al., 2018).

We begin this article with a brief review of relevant principles on the multivariate geometry of forced-choice tasks. We then present a generalization of these principles to models of unforced choice and present the simplest possible *independent observations* model that this framework can support. In the third section, we further derive other extant models, including the *dependent observations* model (Akan et al., 2021), the *integration*

model (Duncan, 2006; Graham, Kramer, & Yager, 1987; Wixted et al., 2018), and the *ensemble* model (Wixted et al., 2018). Throughout, we present exact expressions for each of the models we review, though most derivations are presented in appendices so that they can be conveniently accessed or bypassed at the will of the reader.

Multivariate geometry of n-alternative forced-choice

In standard signal-detection theory, the task for the observer is to discriminate between *signal* events and *noise* events. Any distinguishing characteristic can serve to differentiate these two classes, though it is common for that characteristic to be a perceptual one (for perception tasks) or reflect an aspect of prior exposure (for memory tasks). Signal events arise from a population of stimuli for which the to-be-detected target property is present, though it is worth noting that this property may be evident at the population level but not within each individual stimulus. Brown bears are on average larger than black bears, but the distributions of size overlap considerably, making identification solely on the basis of size challenging. Similarly, cognitive and perceptual tasks sometimes characterize signal events by their true physical status (say, orientation or luminance), with the idea that representational noise arises uniquely from perceptual mechanisms, and sometimes employ stimuli that are derived from populations that differ in the target dimension but are individually impossible to discriminate based on physical characteristics (e.g., Mueller & Weidemann, 2008), like in the example with bears. The case is even more complex for memory tasks, in which the nature of the underlying representations is truly unknown.

In standard forced-choice tasks, the observer is assumed to select the stimulus that yields the highest amount of evidence (Luce, 1963, pp. 138–139). It is useful to frame this decision using multivariate geometry, in which the probability distributions of signal and noise events are represented in a multidimensional space in which each dimension corresponds to one of the n stimuli in the presented set. Figure 1 shows an example of this

model in two-dimensional space, corresponding to a 2-alternative forced-choice task. Just as in standard SDT, the distribution corresponding to noise events is centered by fiat at the origin. The location of the signal distribution is offset in one dimension in that space. The shape of the distribution represents the correlation of evidence across stimuli within a set; in the example, the fact that it is an ellipsoid whose major axis aligns with the identity diagonal indicates a positive correlation[2].

With this simple structure, sophisticated models of forced choice across any set size can be easily derived. The probability of selecting the target and the probability of selecting one of the $n-1$ lures are represented by regions in this space. In 2-alternative forced-choice, these regions correspond to each of the two half-planes separated by the identity line. More generally, in n-alternative forced-choice, the regions are found by the partition of the space by hyperplanes into n symmetric regions around the identity line of that space (Luce, 1963). These models have been tremendously useful in research on perception, in which they have served as the backbone for some of the strongest theories in tasks related to the ones we pursue here, like visual search (Eckstein, 1998; Palmer, 1994). More extensive treatments of multivariate signal-detection theory and its application to forced-choice responding are provided by Macmillan and Creelman (2004) and Wickens (2001).

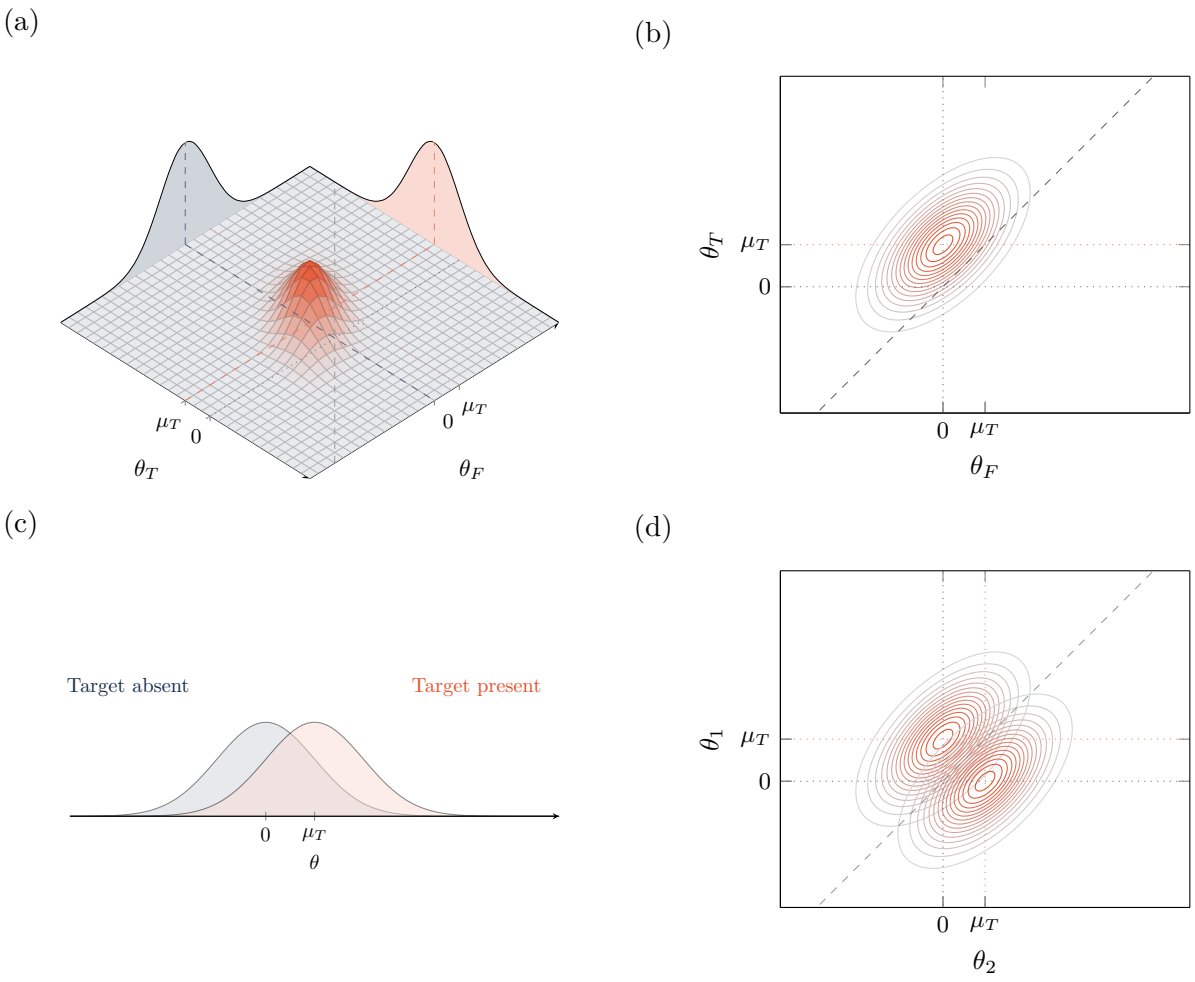
At a conceptual level, extending signal-detection models into the domain of unforced choice requires the introduction of a decision criterion that divides the space yet further, including a region in which none of the stimuli are judged to provide sufficient evidence for an endorsement. At a mathematical level, this means that the domain of the involved integrals reflects both an absolute criterion and the relative comparison among stimuli; the computations become considerably more complex but can still be derived.

---

[2] That is, the line where $y = x$ or, in the example in Figure 1b, where $\theta_T = \theta_F$.

Figure 1

Bivariate representation of 2-alternative forced-choice



Note. A joint distribution of two internal responses, one to a stimulus from the noise population (θ_F), and one to a stimulus from the signal population (θ_T). Dashed lines that run parallel to the axes indicate the locations of the population means, at 0 for the noise population, and at μ_T for the target population. The dashed diagonal line shows the identity line that partitions the space into the regions where the target stimulus is chosen (to the left of the line), and where the lure is chosen (to the right). Panel (a) presents the joint density of the responses together with their marginal densities. Panel (b) shows the corresponding contour plot. Panel (c) shows the two marginal distributions superimposed on the same response axis. Panel (d) superimposes the contour plots of the distributions when the target is present in one or the other observation area.

Structure of the variance-covariance matrix

One important step in providing a multivariate formulation for signal-detection theories is facing the question of how to restrict the variance-covariance matrix that describes the relationships among the stimuli within an array. In univariate SDT, this problem is limited to the question of whether the signal and noise distributions are presumed to have the same variance or are allowed to differ. This question leads directly to the primary distinction among extant univariate models, namely, whether they utilize the mathematics of equal-variance signal-detection theory (EVSD) or of unequal-variance signal-detection theory (UVSD; e.g., Jang et al., 2009). UVSD requires estimation of an additional parameter and so is undetermined in the many experimental designs that elicit only a single hit and false alarm rate. However, in direct comparisons using designs that use confidence ratings (Glanzer, Kim, Hilford, & Adams, 1999) or base-rate manipulations (Ratcliff, Sheu, & Gronlund, 1992), UVSD has generally proven to be the superior model.

In multivariate applications, technical challenges usually motivate use of the equal-variance assumption. This assumption usually carries along with it the assumption of equal covariance among stimuli. For example, the model of visual search presented by Eckstein (1998) assumes the variance of the distributions of evidence to be equivalent for target and lure items within and across displays, and further assumes that the samples are independent of one another (thus assuming a common covariance of 0).

The full complement of model restrictions for variance-covariance structures can be seen in Figure 2, which shows contour plots for a two-item display. The least restrictive model allows the distributions to be correlated and to have unequal variance (panel [a]), and the most restrictive model (shown in panel [e]) disallows both of these characteristics (and so corresponds to the EVSD model). Panels [c] and [d] show the cases for partial restriction among these assumptions. Panel [b] shows a new formulation, introduced here, in which the correlation between the target item and members of the filler population is a function of the variance of the signal distribution and of the correlation between filler items

themselves. This formulation adds additional flexibility to the model with equal variance and a common correlation across array members (panel [c]) in that the correlation between target and fillers is not constrained to be the same value as the correlation among the fillers. It is, however, restricted in important but not dramatic ways. The theoretical mechanism underlying this modification is that the two correlations, regardless of condition, have a single common source. The basis for this restriction is presented in detail in Appendix A.

The juxtaposition of the empirical success of the unequal-variance model in univariate applications with the ubiquity of the equal-variance assumption in multivariate applications is noteworthy. Whether the equal-variance assumption survives in multivariate models is a matter to be resolved by empirical inquiry, and we do not pursue that agenda here. We do, however, provide formulations of multivariate models that adopt the traditional equal-variance assumption (in the next section) and versions that do not (in the appendices). The models that employ the equal-variance (and equal covariance) assumption are more directly compatible with extant models and are easier to use. The less restricted models are more generalizable but may not be implementable for all experimental designs.

With no restrictions on the structure of the variance-covariance matrix, derivations of the general model would likely prove impossible. Thankfully, the basic assumptions of SDT inform simple, uncontroversial restrictions that meaningfully simplify the models and do not constrain the range of plausible models. These arguments and the accompanying proofs are provided in Appendix A, though this material can be safely skipped without impeding understanding of the restricted models presented in the next section. Derivations of the standard forced-choice model based on these principles are presented in Appendix B.

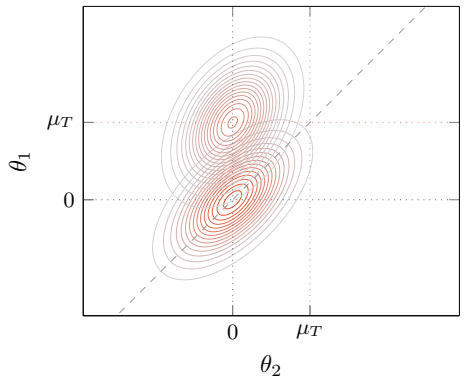
Signal detection models of n-alternative unforced choice

We start by posing a general version of the unforced-choice task that maps well onto the empirical measurement of lineup memory and possesses the basic features needed for generalization to any unforced-choice task. For each of the n stimuli that are presented

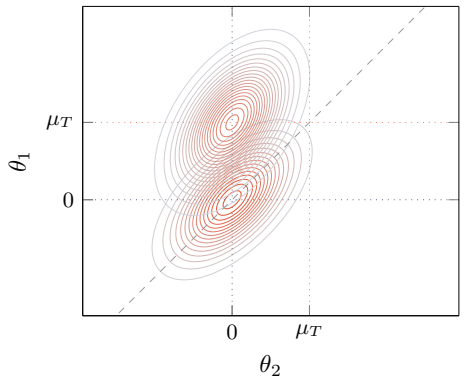
Figure 2

Bivariate representation of different variance-covariance structures

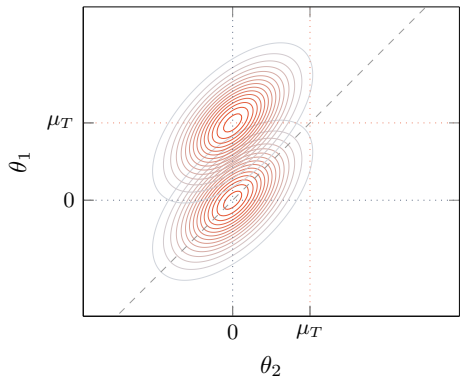
(a) *Unequal variance and correlation*



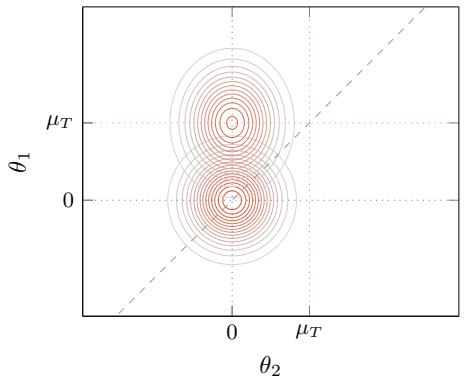
(b) *Unequal variance and correlation per common source*



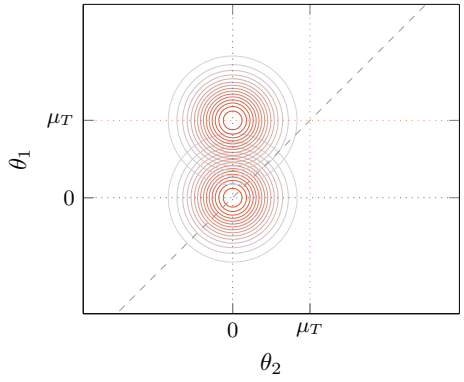
(c) *Equal variance and correlation*



(d) *Unequal variance with independence*



(e) *Equal variance with independence*



Note. Classes of assumptions for the bivariate model with respect to restrictions on variance (equal/unequal) and covariance (correlated/independent).

within each trial, the characteristics of that stimulus depend only on the parent population from which it is sampled. In this case, the n stimuli in the sample are selected in one of two ways:

1. Target-absent trials (TA): All stimuli are sampled from the noise population for which the target characteristic is absent.

2. Target-present trials (TP): One target stimulus is drawn from the signal distribution, and the remaining $n - 1$ lure stimuli are drawn from the noise distribution.

Other combinations are possible, in which several or even all the stimuli are sampled from the signal population. Those variants are not presented here, but they can be developed from the general framework and the derivations included in this paper.

Data structures in unforced choice

We adopt again the case of a criminal lineup as an example for illustrating how to characterize the data that arise in an unforced-choice experiment. In a lineup procedure, a suspect is presented together with a number of fillers. The fillers are known to be innocent of the investigated crime. The suspect may be the actual perpetrator (in a TP lineup) or may be innocent (in a TA lineup). In all lineup types, there are three possible responses: the suspect may be identified, the lineup may be rejected, or a filler may be identified. These choices correspond to different outcomes that depend on the trial type and are summarized in Table 1.

If the eyewitnesses had perfect memory, then they would always choose to reject the lineup when the suspect is innocent, and they would always select the suspect when they are the true perpetrator. These outcomes are the correct responses in TA and TP trials, respectively.

In TA trials, the selection of any of member of the lineup is an incorrect response. Within the witness identification literature, however, the incorrect selection of the suspect when the suspect is innocent is of greater consequence than the incorrect selection of a

Table 1

Terminology for the outcomes of different responses to target-present (TP) and target-absent (TA) arrays.

	TP	TA
Suspect identification	Hit	FA
Filler identification	FI_{TP}	FI_{TA}
Reject lineup	Miss	Correct rejection

filler (Colloff, Wilson, Seale-Carlisle, & Wixted, 2021; Gronlund & Benjamin, 2018). We define the probability of selecting an innocent suspect as the perpetrator as the *false-alarm rate* (FAR), though, following the lead of Smith, Yang, and Wells (2020) and Wells, Smith, and Smalarz (2015), we also consider the identification of fillers an important variable and provide derivations for *filler-identification* rates (FIR) as well. We define the probability of correctly rejecting the lineup as the *rejection rate* of the TA array (RR_{TA}). For a fair lineup, in the absence of response bias, the probability of choosing a lineup member instead of rejecting the lineup in a TA lineup is simply n times the FAR.

In a TP lineup, where the suspect is truly guilty, there are two different incorrect responses. The eyewitness may incorrectly reject the lineup, an outcome we designate as a *miss* and measure it as the rejection rate of the TP array (RR_{TP}). Alternatively, they may incorrectly identify a filler as the perpetrator, with a corresponding FIR. The probability that they correctly identify the guilty suspect is denoted as the *Hit rate* (HR).

In TP lineups, the FIR can be computed from the other rates as $\text{FIR}_{\text{TP}} = 1 - \text{HR} - \text{RR}_{\text{TP}}$; in TA lineups, it equals $\text{FIR}_{\text{TA}} = 1 - \text{FAR} - \text{RR}_{\text{TA}}$. Fair lineup construction ensures that the suspect in a TA trial can be characterized by the evidence distribution from the noise population; in that case, the filler-identification rate can also be computed as $\text{FIR}_{\text{TA}} = (n-1)\text{FAR}$. In our presentation, we assume that the arrays are

generated fairly—that is, that an innocent suspect will bear no more similarity to the target than the remaining foils in the lineup. When lineup construction is presumed to be unfair, the innocent suspect of a TA trial should be modeled as arising from a different population as the fillers, and different models are necessary. Such models are compatible with the structural assumptions we have made; together, they imply that the distribution of the memory trace for the innocent suspect in a TA lineup follows a normal distribution with some positive mean μ_I—most reasonably such that $\mu_I < \mu_T$. Hence, in a TA trial of this kind all rates would be defined in the same manner as in a TP trial, but with μ_I instead of μ_T.

Model parameterization

As in most applications of signal detection theory, we will assume that the internal representation of evidence is aptly represented by a normal distribution. We will denote the internal representation of evidence for a stimulus by θ_i, $i = 1, \ldots, n$. For a stimulus sampled from the noise population, the distribution of each θ_i is fixed, without loss of generality, to be a standard normal with a mean of 0 and standard deviation of 1. For a stimulus from the target population, the parameters of the corresponding normal distribution of θ_i will be denoted μ_T and σ_T, for the mean and standard deviation, respectively.

Model with restricted variance and covariance. In this section, we adopt the equal-variance and equal-covariance assumptions; hence, $\sigma_T^2 = 1$ and the common covariance reduces to the correlation ρ. From this framework we advance to more flexible and more complete models in which these parameters are also free to vary across individuals or conditions. In this scheme, we characterize a single trial as a vector of n internal representations, $\theta_1, \ldots, \theta_n$, each with a marginal normal distribution and

$$\vec{\theta}_n \sim \mathcal{N}_n(\vec{\mu}_n, \mathbf{\Sigma}_{n\times n}) \tag{1}$$

in which the n memory traces are multivariate normally distributed with mean vector $\vec{\mu}_n$ and covariance matrix $\mathbf{\Sigma}_{n\times n}$. Under the assumptions of equal variances and covariances,

the additional structural assumptions are expressed as

$$\vec{\mu}_n = \begin{pmatrix} \mu_1 \\ \cdots\cdots \\ \vec{0}_{n-1} \end{pmatrix} \tag{2}$$

$$\mathbf{\Sigma}_{n\times n} = (1-\rho)\mathbf{I}_{n\times n} + \rho\vec{1}_n\vec{1}_n^{\intercal} \tag{3}$$

where $\mu_1 = 0$ for stimuli for which the signal is absent (in our examples, a display that does not include a target, or a lineup without the perpetrator), and $\mu_1 = \mu_T > 0$ for stimuli for which the target signal is present (an array that includes a single target stimulus).[3] The same models, with relaxed constraints on both variance and covariance, are presented in appendices. Due to the assumption of equal variance for all event types, the sensitivity index $d' = \mu_T$.

The general rule in standard forced-choice models is that the observer selects the stimulus that yields the highest amount of evidence, which can be expressed as

$$\left(\vec{1}_{n-1} \vdots -\mathbf{I}_{(n-1)\times(n-1)}\right)\vec{\theta}_n > \vec{0}_{n-1}, \tag{4}$$

where the inequality must be read component-wise. If inequality (4) is satisfied, then the first stimulus elicits the highest amount of evidence, and would be chosen in a forced-choice task. For each of the other stimuli i between 2 and n, we can write an expression analogous to (4). For example, if

$$\left(-\mathbf{I}_{(n-1)\times(n-1)} \vdots \vec{1}_{n-1}\right)\vec{\theta}_n > \vec{0}_{n-1} \tag{5}$$

---

[3] The following notation conventions are used in expressions (1)–(3) and throughout the text.

1. A vector with n components is denoted by the symbol naming the vector, the size of the vector as a subscript, and an arrow hat. For example, $\vec{\theta}_n$ is the vector of size n of evidence values, $\vec{1}_n$ is the vector with n ones, and $\vec{0}_n$ is the vector with n zeroes.

2. A matrix of size $m \times n$ is denoted by the boldface symbol naming the matrix, and the size of the matrix as a subscript. For example, $\mathbf{A}_{m\times n}$ is the matrix $\mathbf{A}$ which has m rows and n columns, and $\mathbf{I}_{n\times n}$ is the identity matrix of size n.

3. For writing block matrices and block vectors, the partitions are separated by dotted lines.

is satisfied, then the n-th stimulus has elicited the highest amount of evidence. Only one of the n expressions thus produced can be satisfied in a given trial.[4] Because the labeling of the stimuli is arbitrary, it suffices to work only with the single inequality (4), which identifies the region where the (arbitrarily) first stimulus yields the highest response. Throughout this paper we will refer to this first stimulus as the target, though it should be understood that in actual empirical applications the target would be distributed among the positions in the choice array across trials or subjects.

Expression (4) embodies the criterion-independent aspect of the decision-making; shortly we will turn to the question of how to augment this representation with the criterion-dependent aspect that additionally allows the observer to reject the array. In models of eyewitness identification, this criterion-independent decision rule is commonly referred to as the MAX rule. Figure 3 illustrates in two-dimensional space the MAX rule partition of the space corresponding to 2-alternative unforced-choice. Notably, unforced choice tasks deviate from standard forced-choice tasks in the manner in which they carve up response space—three responses are available to the observer, rather than two. Throughout this paper, we accompany each of the models we introduce with an illustration of the segregation of the decision space into the various outcomes of an experiment utilizing an unforced-choice procedure, as described in Table 1.

This formulation and the derivations that follow are general for any size array—that is, for any dimensionality—but we depict each case in two-dimensional space in the accompanying figures. When the evidence distributions are of greater dimensionality, so are the required partitions; the lines in our figures become planes and hyperplanes in higher dimensions.

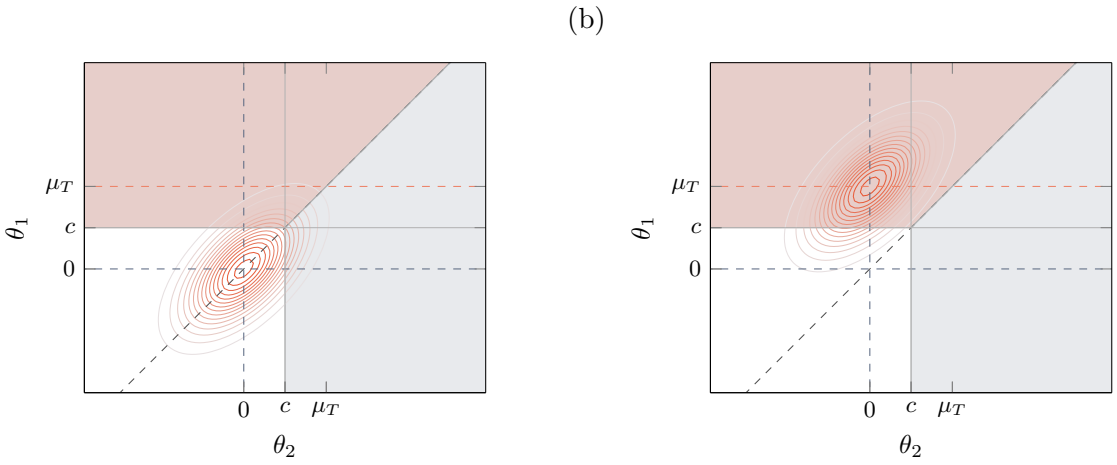
Note also that this model formulation also assumes no bias in selection. That is, responders are presumed not to favor certain positions in the array over others. Bias can be easily introduced by replacing the vector $\vec{0}_{n-1}$ with a vector $\vec{b}_{n-1}$ of bias parameters in

[4] Under the assumption of normality, a tie of two or more responses has probability zero.

the appropriate expression, though doing so increases the difficulty of model estimation and we do not pursue that case here. DeCarlo (2012) presents Bayesian algorithms for estimating n-alternative forced-choice models in the presence of position bias.

Figure 3

Bivariate representation of 2-alternative unforced-choice



Note. Example of the joint distribution of two internal responses in 2-alternative unforced-choice. The dashed diagonal line shows the identity line that partitions the space into the regions where the trace elicited by the first stimulus is higher than that elicited by the second. The solid lines perpendicular to the axes at the value c show the decision criteria for the *Dependent Observations* model that separate the region where the set is rejected (without shading) from regions where a selection is made. The region shaded in orange corresponds to the selection of the first stimulus and the region shaded in blue corresponds to the second stimulus. Each panel shows the contour plots for the joint distributions of two memory traces. Panel (a) presents the case of a TA trial, where both stimuli are drawn from the noise population. Panel (b) presents contours of the joint distribution of a TP trial in which the first stimulus is drawn from the target population and the second is drawn from the noise population.

We expand the model into the realm of unforced choice behavior by incorporating a decision criterion that divides the space into two regions. One region represents the space in which all internal responses fail to reach criterion and the decision is to reject the whole set. The second region represents the space in which the MAX rule determines which of

the stimuli is chosen. The most general way to introduce this decision criterion is by defining a set of functions of $\vec{\theta}_n$ such that

$$f_j(\vec{\theta}_n) < c_j \tag{6}$$

defines the region wherein the set is rejected. Though the framework is complete with the introduction of the functions f_j, with flexibility in the choice of f_j, the most tractable family of functions is that in which each f_j is a linear combination of the components of $\vec{\theta}_n$. That is, for each f_j, there is some vector of coefficients $\vec{k}_n$ such that

$$f_j(\vec{\theta}_n) = \vec{k}_n^{\intercal}\vec{\theta}_n.$$

This representation is not at all restrictive; in fact, all extant models of eyewitness identification can be expressed within this class of functions, as we will show in the next sections. Moreover, by using linear combinations, the models resulting from combining the criterion-independent MAX rule and the criterion-dependent decision rule (6) are tractable using the multivariate normal distribution to compute the relevant probabilities.

Standard results in SDT are easily understood as special cases of this framework. When the stimulus set is of size 1, the MAX rule plays no role and the model reduces to the standard SDT model for yes/no tasks under any (strictly increasing monotonic) choice of f_1. Similarly, for any choice of f_j, when the criterion is set to $-\infty$, the framework for unforced-choice reduces to the model for n-alternative forced-choice.

Limiting behavior. The fact that the behavior of forced-choice models can be understood as a special case of unforced-choice model behavior reveals that unforced-choice models exhibit a limit on endorsement rates and that, conditional upon a particular joint distribution of evidence, the family of all possible unforced choice models share the same limits.

Those limits provide maxima for the HR and FAR in n-alternative unforced choice, and those maxima directly influence the treatment and interpretation of data from unforced-choice experiments. We return to this point shortly, when we introduce different

forms for representing unforced-choice behavior in the form of receiver-operating characteristics. Appendix B presents the derivation of these limits, which are summarized here.

Letting φ and Φ denote the standard normal density function and the standard normal cumulative probability function, respectively, expression (7) provides the upper limit on HR and expression (8) provides the upper limit on FAR.

$$\mathrm{HR} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho}} - x \right) \right]^{n-1} dx \tag{7}$$

$$\mathrm{FAR} = \frac{1}{n}. \tag{8}$$

An intuitive understanding of the origin of these limits can be gained by considering the FAR in unforced-choice. Consider a decision-maker facing a two-item array with no target. Any selection of an item in the display is by definition an error, but the selection among the two options in the array is essentially random. For that reason, there is only a 50% chance that the selected item will be the one arbitrarily designated to yield a false alarm (as opposed to a filler identification). This is a consequence of the treatment of all nontarget stimuli as *essentially the same*, a topic covered in detail in Appendix A.

The same principles apply in a more complex manner to limits on HR. The straightforward limiting behavior of FAR measures in unforced choice models has been noted and understood but to our knowledge has not been derived for HR measures except under the very limited case of complete independence (see e.g., Hacker & Ratcliff, 1979; Macmillan & Creelman, 2004). Figure 4 presents HR values as a function of discriminability, the correlation between array element strengths, and the size of the array.

When $n = 2$, the maximum hit rate takes the form

$$\Phi\left(\frac{d'}{\sqrt{2(1-\rho)}} \right), \tag{9}$$

since we assume unit variance for the noise and target distributions. This expression further simplifies to the well known equation for percentage correct (PC) in 2-alternative

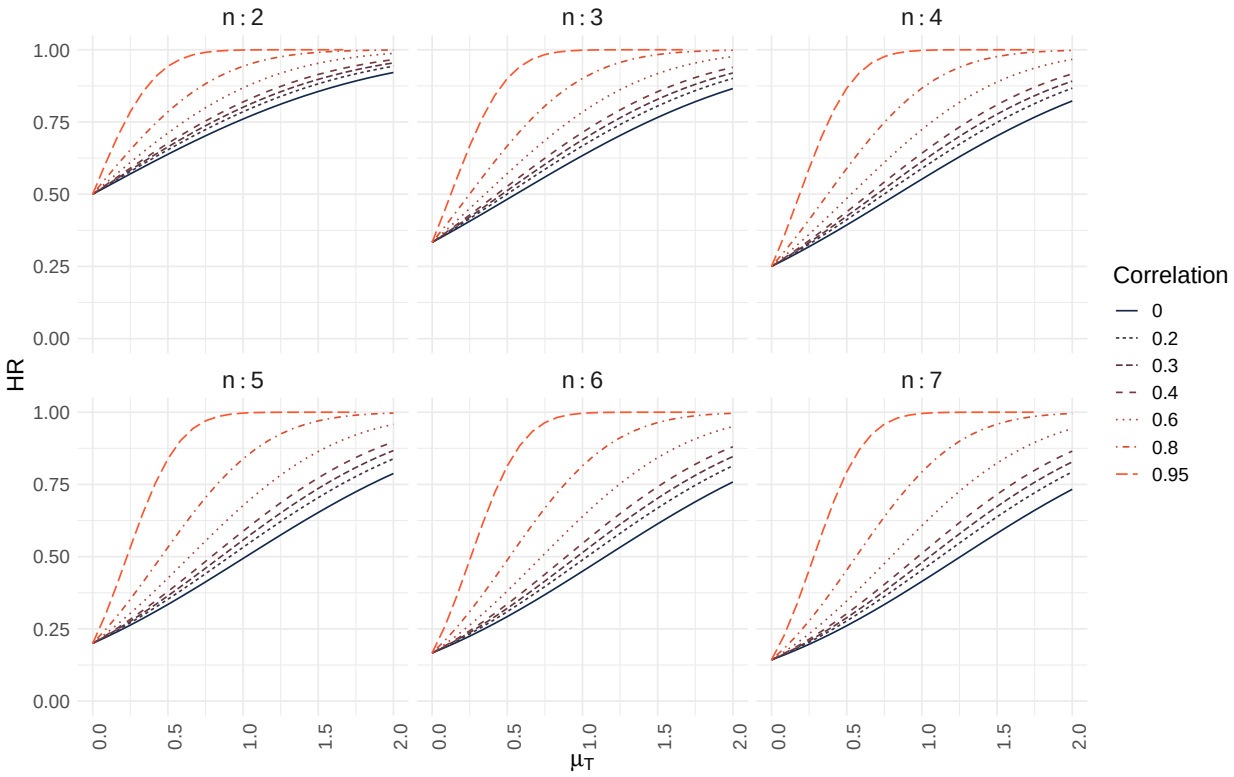
forced-choice

$$\Phi\left(\frac{d'}{\sqrt{2}}\right) \tag{10}$$

when evidence for the two traces is assumed to be independent.

Figure 4

Maximum hit rates for n-alternative unforced-choice



Note. Maximum Hit rate (HR) of for n-alternative forced choice as a function of μ_T for different array sizes n and correlations between element strengths. The ordinate at $\mu_T = 0$ on each panel indicates the maximum of FAR for the corresponding array size.

Independent-Observations model

The simplest signal-detection model in use in the analysis of unforced-choice decision making is the *Independent Observations* (IO) model. This model assumes that the internal responses are stochastically independent, and that the observer compares each

internal response to a criterion value, as if they performed n yes/no tasks. If none of the comparisons exceed the criterion, the observer rejects the whole set; if any comparison produces a detection, then the MAX rule is applied and the stimulus that yielded the highest response is chosen. The model is extremely useful as a starting point for theorizing because its assumptions are transparent and commonplace in statistical modeling. It is also simple to work with because, by the assumption of independence, the joint probability functions are found through the product of the distribution functions of each of the n internal responses. It provides an instructive starting point for thinking about models of unforced choice and also for developing the terminology and mathematics for more advanced, and more realistic, models.

The Independent Observations model was first presented as a model for eyewitness identification by Duncan (2006), following the presentation by Macmillan and Creelman (2004). Outside the context of witness identification, the 2-alternative case was presented by Watson et al. (1973) within traditional psychophysics, and by Phelps et al. (2006) in the context of mate choice and species recognition. Within psychophysics, the assumptions underlying this model appear in models of multidimensional detection (Thomas et al., 2015) and undergird a prominent theory of visual search (Eckstein, 1998; Palmer, 1994).

The rejection rates on both TA and TP trials require knowledge only of the probability of a single event failing to exceed the criterion. The probability that the response to a filler does not exceed the criterion is $\Phi(c)$, and the probability that a target event does not exceed the criterion is $\Phi(c - \mu_T)$. Since the array is rejected if none of the stimuli elicits a response that exceeds the criterion, and those responses are independent of one another, the Rejection rate for TA trials is

$$\mathrm{RR}_{\mathrm{TA}} = [\Phi(c)]^n \,, \tag{11}$$

and for TP trials is

$$\mathrm{RR}_{\mathrm{TP}} = \Phi(c - \mu_T)\,[\Phi(c)]^{n-1} \,. \tag{12}$$

To find the HR and FAR of the Independent Observations model, we additionally need to use the law of total probability for continuous variables. Note that a correct identification in a TP trial occurs when the internal response for the target stimulus is larger than the $n-1$ responses for the fillers. By independence, the likelihood of such an event when the evidence elicited by the target stimulus equals x is

$$\varphi\left(x-\mu_T\right)\left[\Phi(x)\right]^{n-1}.$$

From this formulation, we can compute the HR, using the law of total probability, by integrating over the range of values where the response to the suspect surpasses the criterion. This leads to expression (13) for the HR.

$$\mathrm{HR}=\int_c^{\infty}\varphi\left(x-\mu_T\right)\left[\Phi(x)\right]^{n-1}dx. \tag{13}$$

Similarly, expression (14) provides the FAR.

$$\mathrm{FAR}=\int_c^{\infty}\varphi\left(x\right)\left[\Phi(x)\right]^{n-1}dx. \tag{14}$$

The simplicity of the Independent Observations model allowed us to derive the expressions to compute these rates without referring to the multivariate geometry of the evidence distributions. Nonetheless, we can describe the model using the language of the multivariate framework. The assumption of normally distributed independent memory traces can be described in terms of a multivariate normal with mutually independent, and hence uncorrelated, components. The comparison of each memory trace to a given criterion value defines the set of linear combinations and decision criterion for the criterion-dependent part of the model. Lastly, the requirement that the memory trace for the selection of the guilty suspect or of a lure in the computations of the HR and the FAR corresponds to the application of the MAX rule. This framing is formally completed in the next section as part of deriving the *Dependent Observations* model, for which the Independent Observations is obtained as a special case where $\rho=0$. The full derivation from axiomatic principles of both the Independent Observations and Dependent Observations models is presented in Appendix C.

Receiver-operating characteristics

One way of characterizing a detection-theoretic model is by its predictions about the relationship between HR and FAR over the set of all possible decision criteria. All models share the same fundamental starting points: when the decision criterion is conservative, hits and false alarms are both infrequent; when the decision criterion is liberal, hits and false alarms are both common. The function that relates the decision criterion to these essential statistics is called a receiver-operating characteristic (ROC); here we review a number of variants of the ROC for the case of unforced choice that vary in content. By tradition, the ROC is plotted with FAR on the abscissa and HR on the ordinate. The shape and location of the ROC is used to represent the theory and parameter set from which it was generated. Individual points on that function specify a unique decision criterion.

One of the useful aspects of the ROC is that it provides an intuitive means of summarizing the discriminability of a situation or observer. When discriminability is weak, the HR and FAR will be very close to one another and the ROC will lie on or near the positive diagonal. Under those conditions, the area under that curve (AUC) will be around 0.5. When discriminability is high, the HR will be considerably higher than the FAR, yielding a function with an AUC approaching 1.0. An additional appeal of the ROC as a measure of discriminability is that the AUC derived from a yes/no decision-making task predicts 2-alternative forced-choice accuracy quite well (Green & Moses, 1966). These data, among others, support the validity of SDT as a model of decision making in the presence of noise (Wixted, 2020).

One challenge in developing models of unforced choice is that the standard apparatus for moving between probability distributions and the ROC is incomplete. This can be most easily understood by noting that the four responses typically gathered in a yes/no decision experiment, which can be summarized by two independent performance statistics, render hits and false alarms a complete summary of each response category or confidence level for each subject or condition. Yet, as shown in Table 1, unforced choice

tasks yield six, rather than four, response outcomes. The consequence of this complexity is that a single ROC in two dimensions is insufficient to account for the totality of performance in an unforced-choice task (cf. Wells et al., 2015).

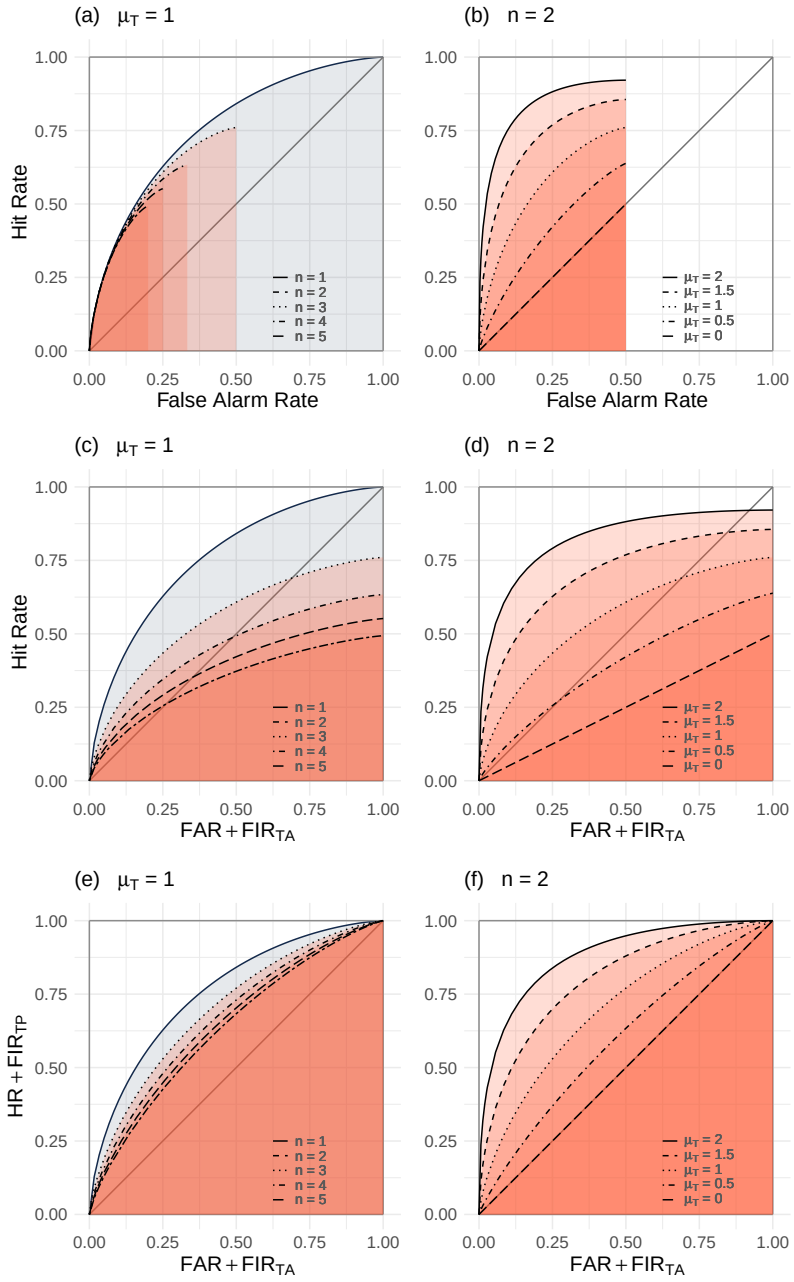
There are a number of ways of extracting an ROC from a table of responses in an unforced-choice experiment, and no one choice is optimal for all purposes. Here, we show how each model can be summarized with three different ROC functions, each of which highlights different aspects of performance.

The target operating characteristic. The first ROC that we consider features identification of the target, or a designated replacement for the target, within the ROC. It maps HR—the identification of the target in a target-present array—against FAR—the identification of a foil that replaces the target in a target-absent array, and so maintains a strong parallel with a "traditional" ROC from a yes/no experiment. In this paper, we refer to this function as a *TOC*, highlighting the centrality of the target in a TP array and its replacement in a TA array. Though the axes enjoy a straightforward mapping to a traditional ROC (HR and FAR), the resultant function is quite different. The top two panels of Figure 5 show examples from the IO model when the array size is varied (panel [a]) and when discriminability varies (panel [b]). These ROCs are *irregular*, in that they do not connect the lower leftmost point in the space with the upper rightmost point in the space. This is a consequence of the limiting behavior of the model discussed in the previous section. A decision-maker might successfully choose to reject the array—in which case the outcome is not tallied in this particular ROC. If they do mistakenly choose an option, there is only a 50% (in a 2-item array) chance that it will count as a false alarm (as opposed to a filler identification). This is why the data abruptly stop along the abscissa at a value of 0.5 (more generally, at $1/n$, as explained above). The truncated nature of these ROCs have led also to their designation as partial ROCs (pROC).

A partial ROC introduces a challenge for assessing discriminability, because the maximum area under that curve is no longer 1. As a practical matter, researchers have

Figure 5

Three receiver-operating characteristics for n-alternative unforced choice under the IO model.



Note. Panels [a] and [b] show *target-operating characteristics* (TOC); panels [c] and [d] show *identification-operating characteristics* (IOC); and panels [e] and [f] show *ensemble-operating characteristics* (EOC). Left panels show the effect of array size (n) on the functions. Right panels show the effect of discriminability (μ_T) on the functions.

attempted to circumvent this problem by comparing areas under the partial ROC (pAUC), though a comparison between conditions commonly requires either extrapolation of a shorter ROC to match the length of a longer one—in which case the theory underlying that extrapolation might be wrong—or elimination of the tailing end of the longer ROC—in which case data are being discarded and the procedure is *de facto* inefficient.

The target operating characteristic enjoys straightforward interpretation in experimental designs in which a singular foil in a TA array enjoys special status as a replacement for the target in a TP array. Such a situation is common in lineup memory experiments, in which the truly guilty suspect is replaced by a foil who is not guilty, but is different from the other foils in that they are at especial risk of prosecution if they are selected by the witness.

Identification-operating characteristic. The second set of plots shown in Figure 5 (panels [c] and [d]) plot HR against FAR + FIR. Here, we call this function an *identification-operating characteristic* (IOC). This means of plotting the data has the appeal of placing the correct response for a target-present trial on the ordinate and all possible incorrect responses from a target-absent trial on the abscissa—an alternative parallel with a standard ROC. This choice of axes extends the ROC horizontally, but not entirely vertically, yielding a partially regular function. This regularity comes at a cost, however—the functions now verge into the region of space previously understood as representing below-chance performance (below the positive diagonal). It is easy to see how this happens: though the ordinate reflects the probability of endorsement for a single item in a target-present lineup, the abscissa now reflects the endorsement of any of n items in a target-absent lineup. In essence, we have one dimension of the IOC that is regular and one that is irregular. This plot enables straightforward comparisons of conditions but eliminates the meaningful boundary value of the area under the curve of 0.5 as representing chance performance. Pure guessing will yield performance on the line $\mathrm{HR} = {}^{\mathrm{FAR}}/_{n}$, yielding an ROC with a slope of ${}^{1}/_{n}$ and AUC of ${}^{1}/_{2n}$. In the absence of array bias (in which some

members are systematically more alluring than others), the AUC for an IOC enjoys a simple relationship with the AUC for a TOC: $\mathrm{AUC}_{\mathrm{IOC}} = n\mathrm{AUC}_{\mathrm{TOC}}$.

An IOC is a straightforward tool for designs in which there is homogeneity among the members of a TA array. This is the case in many perceptual experiments in which a target is to be detected against a background of distractors, none of which have any status as the "replacement" for a target. This is not to say that the members of a TA need to be identical to one another, only that the experimenter does not wish to feature in the analysis any single one of them as a stand-in for a target.

Ensemble-operating characteristic. The final means of plotting performance in unforced choice is shown in panels [e] and [f] of Figure 5. Here, we call this function an *ensemble operating characteristic* (EOC).[5] Endorsements of any member of a target-present array are plotted on the ordinate ($\mathrm{HR} + \mathrm{FIR}_{\mathrm{TP}} = 1 - \mathrm{RR}_{\mathrm{TP}}$), and endorsements of any member of a target-absent array are plotted on the abscissa ($\mathrm{FAR} + \mathrm{FIR}_{\mathrm{TA}} = 1 - \mathrm{RR}_{\mathrm{TA}}$). This approach yields a regular ROC that obeys the AUC boundaries of a standard ROC of 0.5 and 1.0, remedying the problems faced by the receiver-operating characteristics summarized above, but it is also not without cost. Namely, the discriminability measured here is at the level of the ensemble (or array) and does not reveal anything about the identity of specific members of that array. So the AUC of the EOC indexes the degree to which an individual can discriminate ensembles with a target from ensembles without a target, but does not speak to their ability to select or reject the specific target within that ensemble. This feature makes the EOC an appropriate tool for describing detection of targets in the absence of localization or identification. For that reason it is poorly suited to eyewitness identification but appropriate for many basic tasks of detection and discrimination.

---

[5] This function has also been designated an ROC by Macmillan and Creelman (2004) and Meyer-Grant and Klauer (2021). It is a true ROC for only for detection at the level of the ensemble, which is why we have renamed it here.

Models of unforced choice applied to eyewitness identification

In this section, we present signal detection models of eyewitness memory that have been proposed as alternatives to the Independent Observations model. We derive each one as specifications of the multivariate framework introduced in the previous section. For each model, we highlight a simplified version in the text in which signal and noise variance are assumed to be equal, as is the covariance across those classes of items. Separate appendices for each model present the fully general case in which variance and covariances are free to vary within the minimal constraints provided by the analysis in Appendix A. For each model, we also present examples of all three types of receiver-operating characteristic as a means of visualizing aspects of the model's predictions.

Dependent observations model

Akan et al. (2021) extended the Independent Observations model to allow for covariation of the internal responses and called the resultant model the *Dependent Observations* (DO) model. The main complication in deriving closed-form solutions for the outcomes of the DO model is that the lack of independence across signals prevents the use of the product of the marginal densities and cumulative probability functions to compute the relevant joint probabilities.

We frame the model using expressions (1)–(3), assuming that ρ has some non-negative value and that the criterion-independent part of the model follows the MAX rule. The functions used to define the criterion-dependent decision rule of the Dependent Observations model can be written as $f_j(\vec{\theta}_n) = \theta_j$ since each response is compared with the criterion. Each function f_j, $j = 1, \ldots, n$, is easily written as a linear combination of $\vec{\theta}_n$ by taking $\vec{k}_n = \vec{e}_{j,n}$, where $\vec{e}_{j,n}$ is the vector with all but the j-th coefficient equal to zero, and a 1 as its j-th component. The regions corresponding to the decisions to reject the lineup and to choose the first person in the lineup can be written as

$$\mathbf{I}_{n\times n}\ \vec{\theta}_n < c\vec{1}_n \tag{15}$$

and

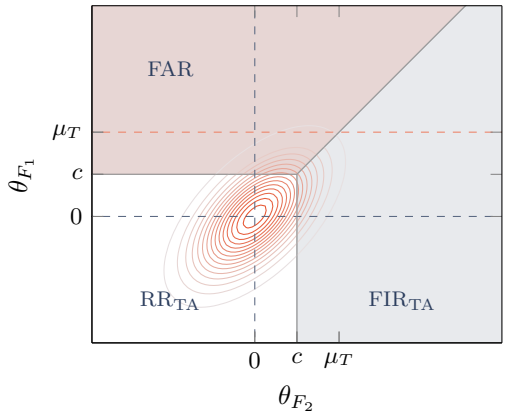
$$\begin{pmatrix} 1 & \vdots & \vec{0}_{n-1}^{\intercal} \\ \cdots & \vdots & \cdots \\ \vec{1}_{n-1} & \vdots & -\mathbf{I}_{(n-1)\times(n-1)} \end{pmatrix} \vec{\theta}_n > \begin{pmatrix} c \\ \cdots \\ \vec{0}_{n-1} \end{pmatrix}, \tag{16}$$

respectively. These regions are illustrated for the 2-dimensional case in Figure 6. Figure 6a illustrates the regions of the space in a TA trial, where two fillers are presented. Figure 6b shows the regions in a TP lineup where a guilty suspect is presented along with one innocent filler.

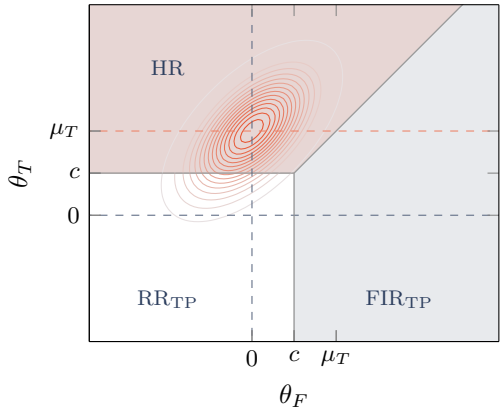
Figure 6

Decision regions for the Dependent Observations model

(a) *Outcomes for a TA array*



(b) *Outcomes for a TP array*



Note. The axes represent the evidence elicited by two nontarget stimuli (θ_{F_1} and θ_{F_2} in panel [a]), and by one target and one nontarget stimulus (θ_T and θ_F in panel [b]). The ellipses represent isoprobability contours of the bivariate normal distribution of $\vec{\theta}_2$. The horizontal and vertical dashed lines at μ_T and at 0 indicate the means of the target and noise populations. The horizontal and vertical lines at c indicate the boundary between the regions where $\theta_{F_1} > c$ and where $\theta_{F_2} > c$. The diagonal line indicates the boundary between the regions where $\theta_{F_1} > \theta_{F_2}$ and where $\theta_{F_1} < \theta_{F_2}$. These boundaries separate the space into three response regions: rejection of the lineup (RR), selection of the (arbitrarily) first stimulus (HR or FAR), and selection of the second stimulus (FIR).

Using the properties of the multivariate normal distribution, we derive the following expressions for the HR, FAR, $\mathrm{RR_{TA}}$, and $\mathrm{RR_{TP}}$ for the Dependent Observations model in

Appendix C. Below we present the outcomes of those derivations and present some simplifying cases for practical use. Expressions (17) and (18) present the HR and the FAR for the Dependent Observations model.

$$\begin{aligned} \mathrm{HR} = \int_{-\infty}^{\infty} \varphi(x)\, \Phi\!\left(\frac{\mu_1 - c}{\sqrt{\rho}} + \sqrt{\frac{1-\rho}{\rho}}x \right) \times \\ \left[\Phi\!\left(\frac{\mu_1}{\sqrt{1-\rho}} + x \right) \right]^{n-1} dx. \end{aligned} \tag{17}$$

$$\mathrm{FAR} = \int_{-\infty}^{\infty} \varphi(x)\, \Phi\!\left(\frac{-c}{\sqrt{\rho}} + \sqrt{\frac{1-\rho}{\rho}}x \right) [\Phi(x)]^{n-1}\, dx \tag{18}$$

These two expressions can be manipulated to arrive at the corresponding expressions for the Independent Observations model as the correlation parameter tends to zero. However, to facilitate derivation we also present alternative expressions for the Hit and False alarm rates for the Dependent Observations model. The following expressions can be derived by first conditioning upon the values of the stimulus that contains the target, in TP arrays, or upon the values of one of the noise stimuli, in TA arrays, then finding an expression for the distribution of the remaining noise stimuli, and, lastly, using the law of total probability.

$$\begin{aligned} \mathrm{HR} = \int_{c}^{\infty} \varphi\left(x - \mu_T\right) \times \\ \int_{-\infty}^{\infty} \varphi(t) \left[\Phi\!\left(\mu_T \sqrt{\frac{\rho^2}{1-\rho}} + x\sqrt{1-\rho} - \sqrt{\rho}t \right) \right]^{n-1} dt\, dx. \end{aligned} \tag{19}$$

$$\mathrm{FAR} = \int_{c}^{\infty} \varphi\left(x\right) \int_{-\infty}^{\infty} \varphi(t) \left[\Phi\!\left(x\sqrt{1-\rho} - \sqrt{\rho}t \right) \right]^{n-1} dt\, dx. \tag{20}$$

The Rejection Rate is computed according to expression (21) for TP arrays and expression (22) for TA arrays.

$$\begin{aligned} \mathrm{RR_{TP}} = \int_{-\infty}^{\infty} \varphi(x)\ \Phi\!\left(\frac{c - \mu_T}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}x \right) \times \\ \left[\Phi\!\left(\frac{c}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}}x \right) \right]^{n-1} dx \end{aligned} . \tag{21}$$

$$\mathrm{RR_{TA}} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{c}{\sqrt{1-\rho}} + \sqrt{\frac{\rho}{1-\rho}} x \right) \right]^n dx. \tag{22}$$

Figure 7 shows ROC curves for the Dependent Observations model at different values of μ_T and ρ.

Yet more compact expressions of these rates can also be obtained specifically for the case of 2-alternative unforced-choice. They can be represented using $\Phi_2(x, y, r)$, the cumulative (quadrant) probability function of the bivariate standard normal distribution with correlation coefficient r:

$$\begin{aligned} \mathrm{HR} =& \int_{c-\mu_T}^{\infty} \varphi(x)\ \Phi\left(\frac{\mu_T}{\sqrt{1+\rho^2}} + \sqrt{\frac{1-\rho}{1+\rho}} x \right)\ dx \\ =& \Phi\left(\frac{\mu_T}{\sqrt{2(1-\rho)}} \right) - \Phi_2\left(\frac{\mu_T}{\sqrt{2(1-\rho)}}, c-\mu_T, -\sqrt{\frac{1-\rho}{2}} \right) \end{aligned} \tag{23}$$

$$\begin{aligned} \mathrm{FAR} =& \int_{c}^{\infty} \varphi(x)\ \Phi\left(\sqrt{\frac{1-\rho}{1+\rho}} x \right)\ dx \\ =& \frac{1}{2} - \Phi_2\left(0, c, -\sqrt{\frac{1-\rho}{2}} \right) \end{aligned} \tag{24}$$

$$\mathrm{RR_{TP}} = \Phi_2(c-\mu_T, c, \rho) \tag{25}$$

$$\mathrm{RR_{TA}} = \Phi_2(c, c, \rho) \tag{26}$$

Independent Observations model, revisited. From expressions (19)–(22), it is straightforward to obtain the expressions for the Independent Observations model by setting $\rho = 0$. For illustration, we show the derivation for the HR. The derivation of the other rates follows the same logic. For completeness, we present them below again without their derivations.

$$\begin{aligned} \mathrm{HR} =& \int_{c}^{\infty} \varphi\left(x-\mu_T\right) \times \\ & \int_{-\infty}^{\infty} \varphi(t) \left[\Phi\left(\mu_T \sqrt{\frac{\rho^2}{1-\rho}} + x\sqrt{1-\rho} - \sqrt{\rho} t \right) \right]^{n-1} dt\, dx \\ =& \int_{c}^{\infty} \varphi\left(x-\mu_T\right) \int_{-\infty}^{\infty} \varphi(t) \left[\Phi(x)\right]^{n-1} dt\, dx \\ =& \int_{c}^{\infty} \varphi\left(x-\mu_T\right) \left[\Phi(x)\right]^{n-1} \int_{-\infty}^{\infty} \varphi(t)\, dt\, dx \\ =& \int_{c}^{\infty} \varphi\left(x-\mu_T\right) \left[\Phi(x)\right]^{n-1} dx. \end{aligned} \tag{27}$$

$$\text{FAR} = \int_c^{\infty} \varphi\left(x\right) \left[\Phi(x)\right]^{n-1} dx. \qquad \text{(14: revisited)}$$

$$\text{RR}_{\text{TA}} = \left[\Phi(c)\right]^n. \qquad \text{(11: revisited)}$$

$$\text{RR}_{\text{TP}} = \Phi(c - \mu_T) \left[\Phi(c)\right]^{n-1}. \qquad \text{(12: revisited)}$$

From expressions (11) and (12) it is possible to obtain formulas for the criterion value and the mean of the target distribution for the Independent Observations model assuming equal variances. These are

$$\hat{c} = \Phi^{-1}(\text{RR}_{\text{TA}}^{1/n}). \qquad (28)$$

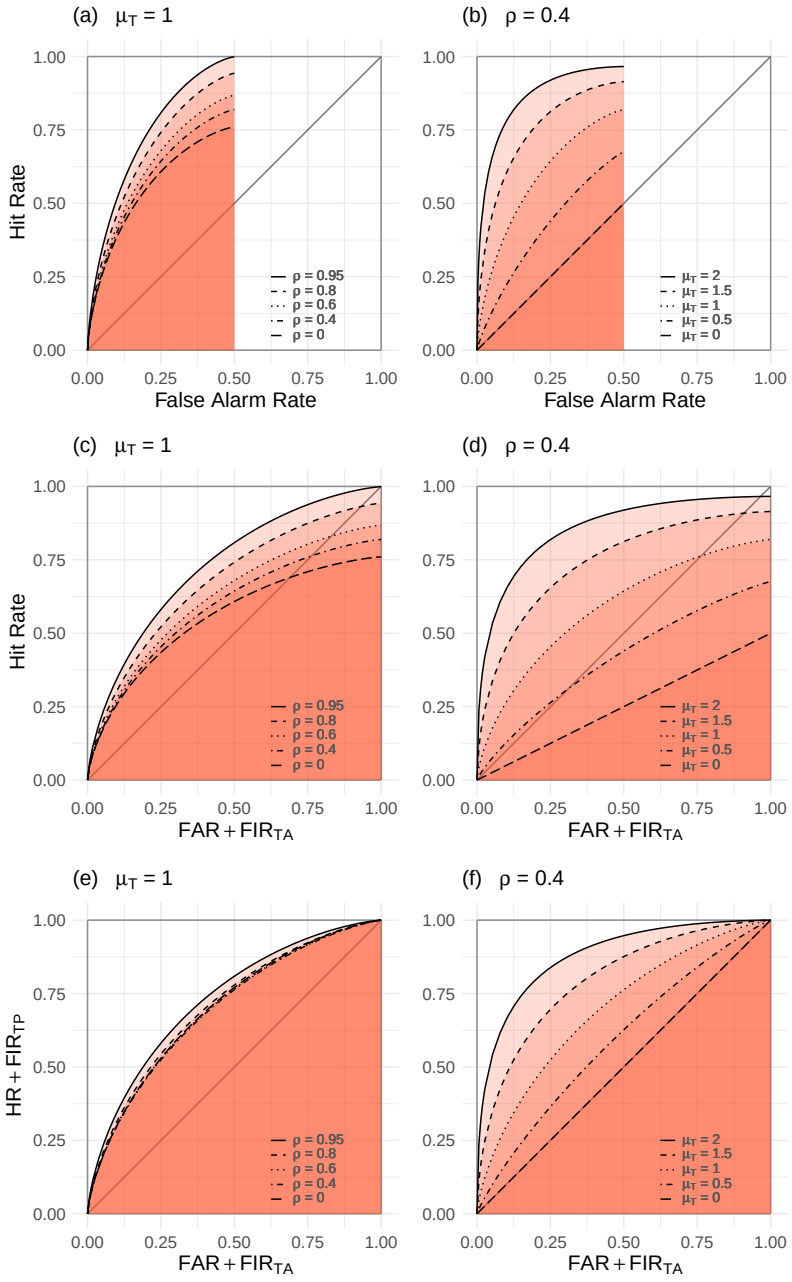
$$\hat{\mu}_T = \hat{c} - \Phi^{-1}(\text{RR}_{\text{TP}} \cdot \text{RR}_{\text{TA}}^{n/(n-1)}) \qquad (29)$$

Integration model

Graham et al. (1987) proposed a set of SDT models in which compound stimuli elicit multiple independent internal responses, just like in the independent-observations model. One unique model that they proposed used the summed output of the channels and compared that *integration* to a decision criterion. Both Duncan (2006) and Wixted et al. (2018) considered the Integration model as a model of lineup memory; Wixted et al. (2018) even extended the model to include the possibility of correlated signals across channels. Though that model did not fare well in accounting for human behavior in eyewitness memory, it has proven useful in basic psychophysics, where tasks involving integration (instead of discrimination) are more common (Graham et al., 1987).

Figure 7

Three receiver-operating characteristics for n-alternative unforced choice under the DO model.



Note. TOCs (panels [a] and [b]), IOCs (panels [c] and [d]), and EOC curves (panels [e] and [f]) for the Dependent Observations model. Left panels show the effect of varying the correlation (ρ) among stimuli. Right panels show the effect of varying discriminability (μ_T).

In the multivariate framework, the function to define a criterion-dependent integration rule is

$$f_1(\vec{\theta}_n) = \sum_{i=1}^{n} \theta_i.$$

Hence, the region where the lineup is rejected is represented as

$$\begin{aligned} \sum_{i=1}^{n} \theta_i <& c \\ \vec{1}_n^{\intercal} \vec{\theta}_n <& c, \end{aligned} \tag{30}$$

and the region where the first (target) stimulus is chosen is

$$\left(\begin{array}{c:c} 1 & \vec{1}_{n-1}^{\intercal} \\ \hdashline \vec{1}_{n-1} & -\mathbf{I}_{(n-1)\times(n-1)} \end{array}\right) \vec{\theta}_n > \begin{pmatrix} c \\ \cdots \\ \vec{0}_{n-1} \end{pmatrix} \tag{31}$$

These regions are illustrated for the 2-dimensional case in Figure 8. Figure 8a illustrates these regions superimposed on the distribution of evidence in a TA trial, and Figure 8b shows the same regions superimposed on the distributions of evidence in a TP trial.

The expressions for the RR are given by

$$\mathrm{RR}_{\mathrm{TP}} = \Phi\left(\frac{c - \mu_T}{\sqrt{n\left[1 + (n-1)\rho\right]}}\right) \tag{32}$$

$$\mathrm{RR}_{\mathrm{TA}} = \Phi\left(\frac{c}{\sqrt{n\left[1 + (n-1)\rho\right]}}\right) \tag{33}$$

The HR and FAR for the integration model are given by expressions (34) and (35).

$$\begin{aligned} \mathrm{HR} = \Phi\left(\frac{\mu_T - c}{\sqrt{n\left[1 + (n-1)\rho\right]}}\right) \times \\ \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho}} - x\right)\right]^{n-1} dx, \end{aligned} \tag{34}$$

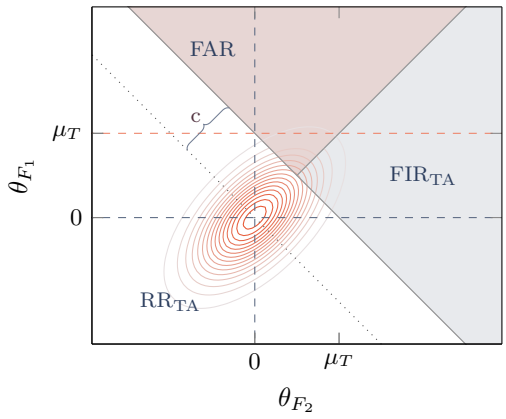
$$\mathrm{FAR} = \frac{1}{n}\Phi\left(\frac{-c}{\sqrt{n\left[1 + (n-1)\rho\right]}}\right). \tag{35}$$

The full derivation of this model is presented in Appendix D, and Figure 9 shows the ROC curves predicted by this model under the same parameters used for the joint distributions in Figure 7.

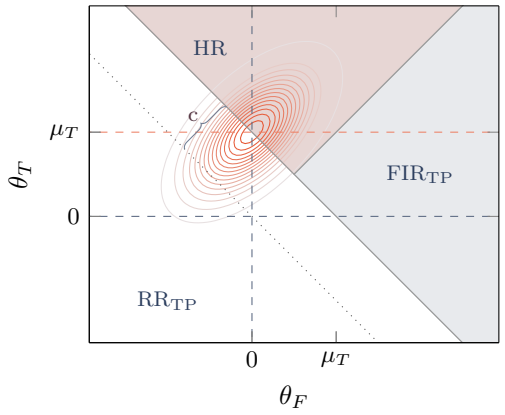
Figure 8

Decision regions for the Integration model

(a) *Outcomes for a TA array*



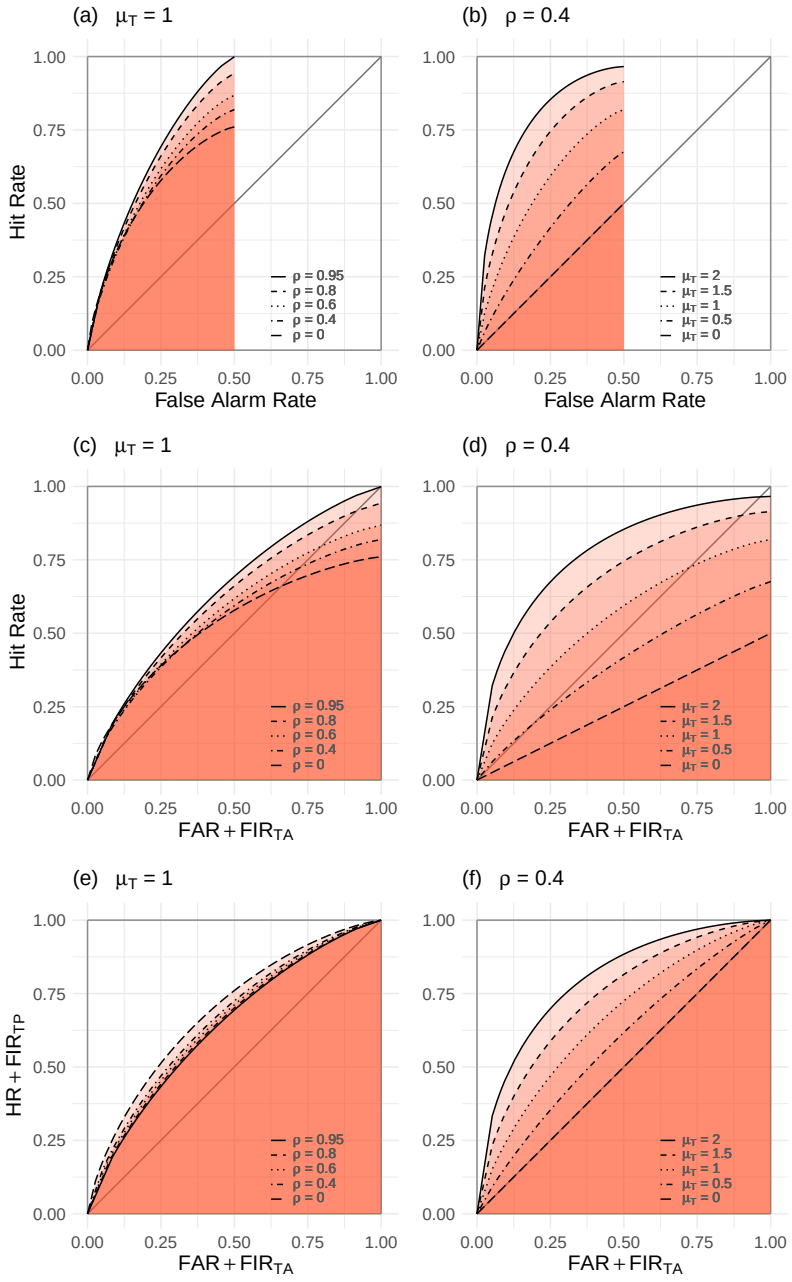
(b) *Outcomes for a TP array*



Note. The axes represent the evidence elicited by two nontarget stimuli (θ_{F_1} and θ_{F_2} in panel [a]), and by one target and one nontarget stimulus (θ_T and θ_F in panel [b]). The ellipses represent isoprobability contours of the bivariate normal distribution of $\vec{\theta}_2$. The diagonal line along the positive diagonal indicates the boundary between the regions where $\theta_{F_1} > \theta_{F_2}$ and where $\theta_{F_2} > \theta_{F_1}$. The dotted diagonal line of negative slope shows the boundary between the regions where $\theta_{F_1} + \theta_{F_2} > 0$ (above the line), and where $\theta_{F_1} + \theta_{F_2} < 0$ (below it). The diagonal solid line parallel to the dotted one marks the decision criterion boundary between the areas in which $\theta_{F_1} + \theta_{F_2} > c$ and $\theta_{F_1} + \theta_{F_2} < c$. The diagonal solid lines separate the three response regions: rejection of the lineup, selection of the (arbitrarily) first stimulus, and selection of the second stimulus. The horizontal (vertical) dashed lines show the location of the means of the target (μ_T) and noise populations (0).

Figure 9

Three receiver-operating characteristics for n-alternative unforced choice under the Integration model.



Note. TOCs (panels [a] and [b]), IOCs (panels [c] and [d]), and EOC curves (panels [e] and [f]) for the Integration model. Left panels show the effect of varying the correlation (ρ) among stimuli. Right panels show the effect of varying discriminability (μ_T).

Ensemble model

The *Ensemble* model (Wixted et al., 2018) operates with a different relative decision rule among the stimuli. It assumes that the evidence yielded by each stimulus in the array is compared to the average evidence elicited by the entire array of stimuli. If the largest of these differences exceeds a decision criterion, then the corresponding stimulus is chosen; conversely, if none of these differences surpasses the criterion, then the lineup is rejected. An earlier development of this theory was presented by Kaernbach (2001), whose model accounted for n-alternative unforced choice under the assumption of independence and also allowed for a *don't know* response by the observer. Though that response is substantively different than a rejection of the array, in which the observer is making a claim about a state of knowledge rather than the lack of one, the modeling is formally equivalent in the two cases. Kaernbach (2001) defined the criterion-dependent part of his model by identifying the hypersurfaces where the ratio between the likelihood of the target corresponding to a specific signal and the sum of the likelihoods that the target corresponds to each of the n signals is constant and equal to $c + 1/n$ (for n-alternative forced-choice; the model can only produce unbiased responding, with $c = 0$). Such a criterion is optimal on Bayesian logic (but perhaps not with human behavior), and requires the structural assumptions of unbiased arrays or fair lineups (i.e., that all nontarget stimuli are exchangable), variance homogeneity (that the variability of signal strength is equal to the variability of noise strength), and common correlation (that the covariance among members of an array is equivalent across specific examplars of the stimuli and across categories of signal and noise). The models we present in the appendices of this paper include versions that relax the second and third of these assumptions, making them considerably more flexible than ones currently in use.

This set of characteristics aligns this model closely with the ensemble model of Wixted et al. (2018), but the decision rules between the two models are notably different. Kaernbach (2001)'s model employs a criterion-dependent rule based on the likelihood ratio.

That assumption obliges the model to assume that the distribution of evidence in a TA trial is optimally separated, in a Bayesian sense, from the evidence distribution in a TP trial, a characteristic that does not comport to the decision structure in the ensemble model. In fact, under the assumption of independence, Kaernbach's model reduces to the Independent Observations model, not to the Ensemble model. When the assumption of independence is relaxed, it is different of all other models reviewed here and is in fact more related in spirit to models employed within the framework of General Recognition Theory using an optimal quadratic criterion (Ashby, 1992; Ashby & Soto, 2015).

Wixted et al. (2018) presented an approximation to the ensemble model based on a normal approximation to the distribution of a truncated normal and the use of the law of total probability. Wixted et al. (2018) assumed the same correlation structure we have used here (expression [3]), representing the models with a unidimensional decomposition into shared and unique variance of the joint distribution. Using these tools, they arrived at the following expression[6] to approximate the HR of the Ensemble model (with analogous expressions for the other rates):

$$\begin{aligned} \text{HR} \approx \int_{-\infty}^{\infty} \frac{1}{\sqrt{1-\rho}}\, \phi\!\left(\frac{x-\mu_T}{\sqrt{1-\rho}}\right) \left[\Phi\!\left(\frac{x}{\sqrt{1-\rho}}\right)\right]^{n-1} \times \\ \left\{1-\Phi\!\left(\frac{nc-(n-1)(x-\mu^*)}{\sigma^*\sqrt{n-1}}\right)\right\} dx, \end{aligned} \tag{36}$$

where

$$\mu^* = \mu_T - Z\sqrt{1-\rho} \text{ and } \sigma^* = \sqrt{(1-\rho)(1-Z\beta-Z^2)},$$

with

$$\beta = \frac{x}{\sqrt{1-\rho}} \text{ and } Z = \frac{\varphi(\beta)}{\Phi(\beta)}.$$

The last factor of the integrand of expression (36, in curly brackets) is the approximation to the probability

$$\Pr\left(\theta_1 - \frac{1}{n}\sum_{i=1}^{n}\theta_i > c \,\middle|\, \theta_1, \max(\vec{\theta}_n) = \theta_1\right).$$

[6] Here we use the notation and terms used in this paper, which differ from those used by Wixted et al. (2018).

The authors make the case that this approximation is reasonable (see unnumbered Figure on p. 109, Wixted et al., 2018). However, a risk yet to be assessed is whether the approximation yields biased estimation of the different rates or of the model parameters. Moreover, the decomposition into the factors of the integrand of expression (36) depends on the assumption that all covariances are equal. While the structural assumptions of equal variance and equal correlations that we have we used here for the restricted versions of our models also pose this requirement, the framework can be used to explore and derive corresponding expressions when a different covariance structure, such as one in which correlations are equal but variances are unequal, is assumed. Detailed presentations of these different versions of ensemble models are provided in Appendix E.

The assumptions that are made by Wixted et al. (2018) to develop their approximation also pose a strong requirement on the specific values of the correlations between stimulus members within a class of items. Specifically, it requires both that the correlation between nontarget evidence values is the same (ρ_F) across all pairs (as we have done here as well), and that the correlation between targets and fillers is the ratio of the correlation between fillers to the standard deviation of the target population (ρ_F/σ_1). This assumption is implicit in the reexpression of the likelihoods of each response in terms of common and unique variance presented in Wixted et al.'s Appendix B.2. This latter assumption is employed for mathematical convenience; the development of the ensemble model we present here poses no such restriction. As a reminder, Appendix A includes a detailed discussion of the various considerations that can be used to constrain the covariance matrix.

Exact expressions for the outcomes of the Ensemble model can be derived within the multivariate framework. We derive these expressions in Appendix E and present them below. Using the exact expression, we can also verify the accuracy of the approximation in expression (36) under the assumptions of equal variances and covariances. We return later to examine the appropriateness of the approximation after the presentation of the model's

exact expressions.

The ensemble criterion-dependent rule of taking the difference between each stimulus and the mean of the ensemble means that the appropriate functions f_j are

$$f_j(\vec{\theta}_n) = \theta_j - \frac{1}{n}\sum_{i=1}^{n}\theta_i,$$

for $j = 1, \ldots, n$. Consequently, the region where the lineup is rejected is defined by simultaneously satisfying the following n inequalities.

$$\theta_j - \frac{1}{n}\sum_{i=1}^{n}\theta_i < c, \qquad j = 1, \ldots, n \tag{37}$$

To express the functions f_j in the form of linear combinations, it is convenient to transform the inequalities as follows.

$$\begin{aligned}
\frac{1}{n}\left[n\theta_j - \sum_{i=1}^{n}\theta_i\right] &< c, & j &= 1, \ldots, n,\\
n\theta_j - \sum_{i=1}^{n}\theta_i &< nc, & j &= 1, \ldots, n,\\
n\vec{e}_{j,n}^{\intercal}\vec{\theta}_n - \vec{1}_n^{\intercal}\vec{\theta}_n &< nc, & j &= 1, \ldots, n,
\end{aligned}$$

where $\vec{e}_{j,n}$ is the vector with all but the j-th coefficient equal to zero, and a 1 as its j-th component. In this manner, the regions corresponding to the decisions to reject the lineup and to choose the first person in the lineup can be written as

$$\left[n\mathbf{I}_{n\times n} - \vec{1}_n\vec{1}_n^{\intercal}\right]\vec{\theta}_n < nc\vec{1}_n \tag{38}$$

and

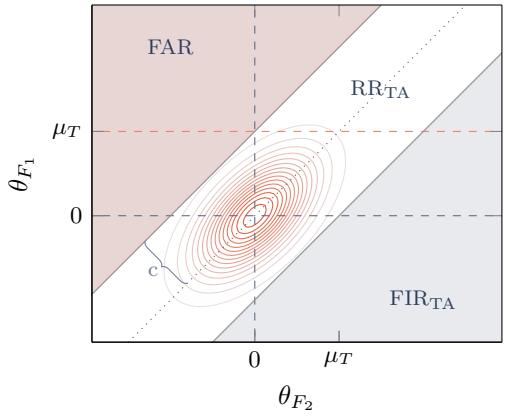
$$\left(\begin{array}{c:c} n-1 & -\vec{1}_{n-1}^{\intercal} \\ \hdashline \vec{1}_{n-1} & -\mathbf{I}_{(n-1)\times(n-1)} \end{array}\right)\vec{\theta}_n > \left(\begin{array}{c} nc \\ \hdashline \vec{0}_{n-1} \end{array}\right), \tag{39}$$

respectively. These regions are illustrated for the 2-dimensional case in Figure 10. Figure 10a illustrates the region of the space in which the lineup is rejected along with the joint distribution of the internal responses in a TA trial; Figure 10b shows the region where the guilty suspect is identified in a TP trial.

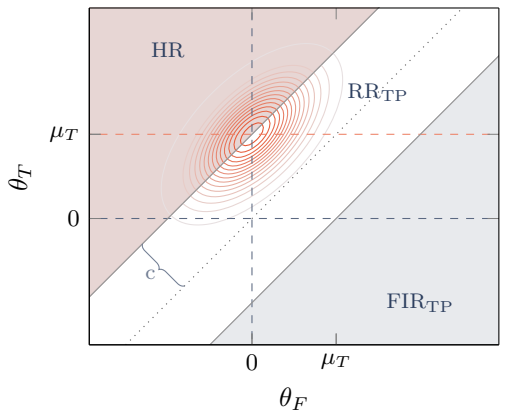
Figure 10

Decision regions for the Ensemble model

(a) *Outcomes in a TA trial*



(b) *Outcomes in a TP trial*



Note. The axes represent the memory traces elicited by each of two stimuli (θ_{F_1} and θ_{F_2} in panel (a), θ_T and θ_F in panel (b)). The ellipses represent isoprobability contours of the bivariate normal distribution of $\vec{\theta}_2$. The solid diagonal lines show the boundaries between the regions where $\theta_T - \frac{1}{2}(\theta_T + \theta_F) > c > 0$, left, and $\theta_F - \frac{1}{2}(\theta_T + \theta_F) > c > 0$, right. The dotted diagonal line shows the boundary between the regions where $\theta_{F_1} > \theta_{F_2}$ and where $\theta_{F_1} < \theta_{F_2}$. The diagonal solid lines separate the three response regions: rejection of the lineup, selection of the first stimulus, and selection of the second stimulus.

The Hit, False Alarm, and Rejection Rates for the Ensemble model are derived in Appendix E. The HR is presented in expression (40).

$$\text{HR} = \int_{nc}^{\infty} \frac{1}{\sqrt{n(n-1)(1-\rho)}}\, \varphi\left(\frac{x-(n-1)\mu_T}{\sqrt{n(n-1)(1-\rho)}}\right) \times \int_{a_1}^{b_1} \cdots \int_{a_{n-2}}^{b_{n-2}} \prod_{k=1}^{n-2} \varphi(z_k)\, d\vec{z}_{n-2} dx \tag{40}$$

where

$$\begin{aligned} a_1 &= -\frac{x}{\sqrt{(n-1)(n-2)(1-\rho)}} \\ a_k &= \sqrt{\frac{n-k}{n-k-1}} \left[\sum_{j=1}^{k-1} \sqrt{\frac{1}{(n-j)(n-j-1)}} z_j - \frac{x}{(n-1)\sqrt{1-\rho}}\right], \\ k &= 2, \ldots, (n-2) \end{aligned} \tag{41}$$

and

$$\begin{aligned} b_1 &= -(n-2)\frac{x}{\sqrt{(n-1)(n-2)(1-\rho)}} \\ b_k &= -(n-k-1)a_k, \qquad k = 2, \ldots, (n-2). \end{aligned} \tag{42}$$

The FAR is computed as shown in expression (43) with the limits of integration a_k and b_k, $k = 1, \ldots, n-2$, defined in the same manner.

$$\begin{aligned} \text{FAR} = &\int_{nc}^{\infty} \frac{1}{\sqrt{n(n-1)(1-\rho)}}\, \varphi\left(\frac{x}{\sqrt{n(n-1)(1-\rho)}}\right) \times \\ &\int_{a_1}^{b_1} \cdots \int_{a_{n-2}}^{b_{n-2}} \prod_{j=1}^{n-2} \varphi(z_j)\, d\vec{z}_{n-2} dx \end{aligned} \tag{43}$$

These two expressions simplify for the case of $n = 2$ to

$$\Phi\left(\frac{\mu_T - 2c}{\sqrt{2(1-\rho)}}\right) \tag{44}$$

for the Hit Rate, and to

$$\Phi\left(\frac{-2c}{\sqrt{2(1-\rho)}}\right) \tag{45}$$

for the False Alarm Rate.

The Rejection Rates for the Ensemble models, RR_{TA} and RR_{TP}, can be computed using expression (46). In the expression, one sets $\mu_1 = \mu_T$ for TP trials and $\mu_1 = 0$ for TA trials.

$$\text{RR} = \int_{a_1}^{b_1} \cdots \int_{a_k}^{b_k} \cdots \int_{a_{n-1}}^{b_{n-1}} \prod_{k=1}^{n-1} \varphi(z_k)\, d\vec{z}_{n-1} \tag{46}$$

where

$$\begin{aligned} b_1 &= \frac{nc - (n-1)\mu_1}{\sqrt{n(n-1)(1-\rho)}} \\ b_k &= \sqrt{\frac{n-k+1}{n-k}} \left[\frac{nc + \mu_1}{n\sqrt{1-\rho}} + \sum_{j=1}^{k-1} \sqrt{\frac{1}{(n-j+1)(n-j)}} z_j\right], \\ &\quad k = 2, \ldots, (n-1), \end{aligned} \tag{47}$$

and

$$\begin{aligned} a_1 &= -(n-1)\frac{nc + \mu_1}{\sqrt{n(n-1)(1-\rho)}} \\ a_k &= -(n-k)b_k, \qquad k = 2, \ldots, (n-1). \end{aligned} \tag{48}$$

When lineup size is $n = 2$, expression (46) simplifies to

$$\Phi\left(\frac{2c - \mu_1}{\sqrt{2(1-\rho)}}\right) - \Phi\left(\frac{-2c - \mu_1}{\sqrt{2(1-\rho)}}\right) \tag{49}$$

Figure 11 shows the ROC curves predicted by this model under the same parameters for the evidence joint distributions as in Figure 7.

BEST-Rest model. To complete the presentation of the Ensemble model, we consider the BEST−Rest (BEST) model (Clark, 2003; Clark, Erickson, & Breneman, 2011). This model is similar to the Ensemble model, and, as noted by Wixted et al. (2018), the two models are linearly related. In the BEST model, it is assumed that each memory trace is first compared to the average of the traces elicited by the *remaining* stimuli. If the largest of these differences exceeds the decision criterion, then the corresponding stimulus is chosen; conversely, if none of these differences surpasses the criterion, then the lineup is rejected. The relationship between the two models is made apparent if we write the functions f_j and their corresponding representation as linear combinations of $\vec{\theta}_n$. The functions f_j are

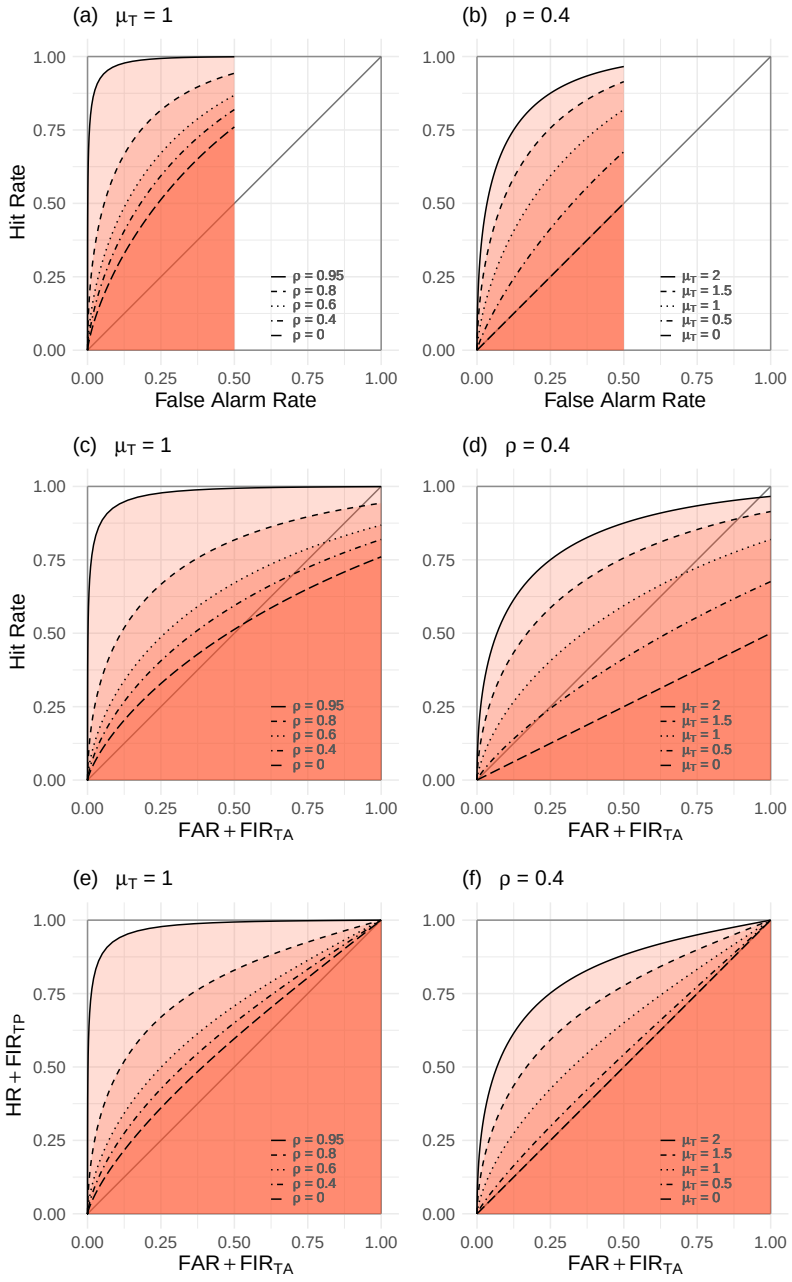
$$f_j(\vec{\theta}_n) = \theta_j - \frac{1}{n-1}\sum_{\substack{i=1 \\ j \neq j}}^{n} \theta_i,$$

for $j = 1, \ldots, n$. Consequently, the region where the lineup is rejected is defined by simultaneously satisfying the following n inequalities, with their transforms in terms of linear combinations as follows:

$$\begin{aligned}
\theta_j - \frac{1}{n-1}\sum_{\substack{i=1 \\ j \neq j}}^{n} \theta_i &< c, && j = 1, \ldots, n, \\
\frac{1}{n-1}\left[(n-1)\theta_j - \sum_{\substack{i=1 \\ j \neq j}}^{n} \theta_i\right] &< c, && j = 1, \ldots, n, \\
\frac{1}{n-1}\left[n\theta_j - \sum_{i=1}^{n} \theta_i\right] &< c, && j = 1, \ldots, n, \\
n\theta_j - \sum_{i=1}^{n} \theta_i &< (n-1)c, && j = 1, \ldots, n, \\
n\vec{e}_{j,n}^{\intercal}\vec{\theta}_n - \vec{1}_n^{\intercal}\vec{\theta}_n &< (n-1)c, && j = 1, \ldots, n,
\end{aligned}$$

Figure 11

Three receiver-operating characteristics for n-alternative unforced choice under the Ensemble model.



Note. TOCs (panels [a] and [b]), IOCs (panels [c] and [d]), and EOC curves (panels [e] and [f]) for the Ensemble model. Left panels show the effect of varying the correlation (ρ) among stimuli. Right panels show the effect of varying discriminability (μ_T).

where $\vec{e}_{j,n}$ is the vector with all but the j-th coefficient equal to zero, and a 1 as its j-th component. Therefore, the region corresponding to the lineup rejection can be written as

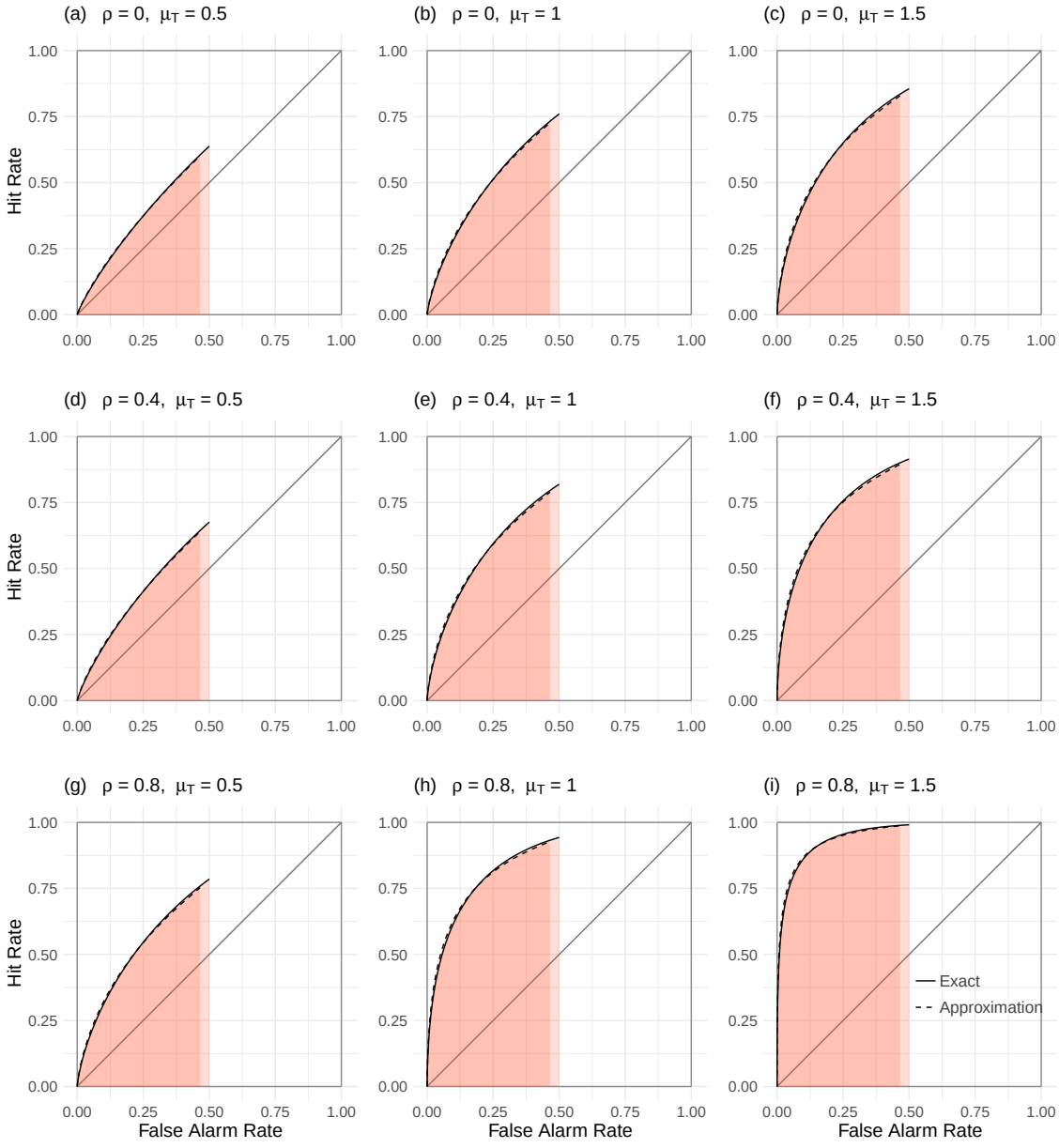
$$\left[n\mathbf{I}_{n\times n} - \vec{1}_n\vec{1}_n^{\intercal}\right]\vec{\theta}_n < (n-1)c\vec{1}_n, \tag{50}$$

The comparison of expressions (50) and (38) shows clearly that the BEST model and the Ensemble model are mathematically equivalent. If the Ensemble model predicts a set of probabilities given a criterion value c_E, the BEST model makes precisely the same set of predictions for the criterion value $c_B = \frac{n}{n-1}c_E$.

Approximation, revisited. Figure 12 presents the TOC curves of the ensemble model computed using the exact expressions, (40) and (43), and the approximation presented in Expression (36). The panels show the TOC curves as the correlation changes, by row, and as discriminability changes, by column, for fixed array size of $n = 2$. We chose this set size because larger set sizes, with consequently shorter TOCs, make it difficult to see the differences between the two approaches. Two important effects are apparent in the figure. First, the curves overlap to a striking degree, substantiating the view that the approximation captures the tradeoff between HR and FAR appropriately. Second, the approximation fails to account for the limiting behavior of the model described previously. This can be seen in the fact that all curves for the approximation end at values lower than the limits from Expressions (7) and (8). The immediate consequence of this shortcoming of the approximation is that the AUC of the TOC is biased. Figure 13 presents the values, computed using the approximation and the exact expressions, of the AUC for several set sizes and magnitudes of discriminability as the common correlation varies from 0 to 1. The bias of the AUC decreases as the correlation among ensemble members increases and also as the array size increases. Since the TOC is generated by the variation of the decision criterion, the fact that the curves overlap but differ in length means that the values of the FAR and the HR predicted by the approximation and the exact expressions cannot be the same for the same criterion value. This fact is illustrated in Figure 14 for several array sizes and correlations. Taking all of these factors together, it appears that there is some

Figure 12

TOC curves for the exact and the approximation of the Ensemble model

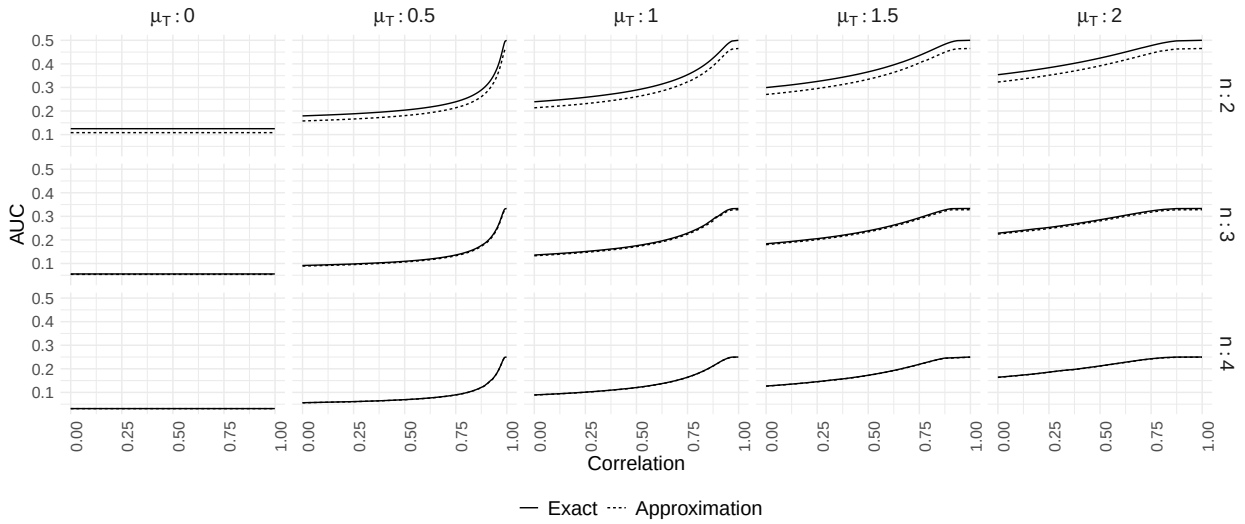


Note. Varying correlations (ρ) by row. Varying μ_T by column.

advantage to using the exact derivation of the ensemble model, especially for parameter estimation and for tasks with small array sizes.

Figure 13

Comparison of AUC for Exact Derivation and Approximation of the Ensemble model



Note. Values of the AUC of the TOC for the Ensemble model computed using the exact expressions for the HR and FAR derived in Appendix E and using the approximation derived by Wixted, Vul, Mickes, and Wilson (2018).

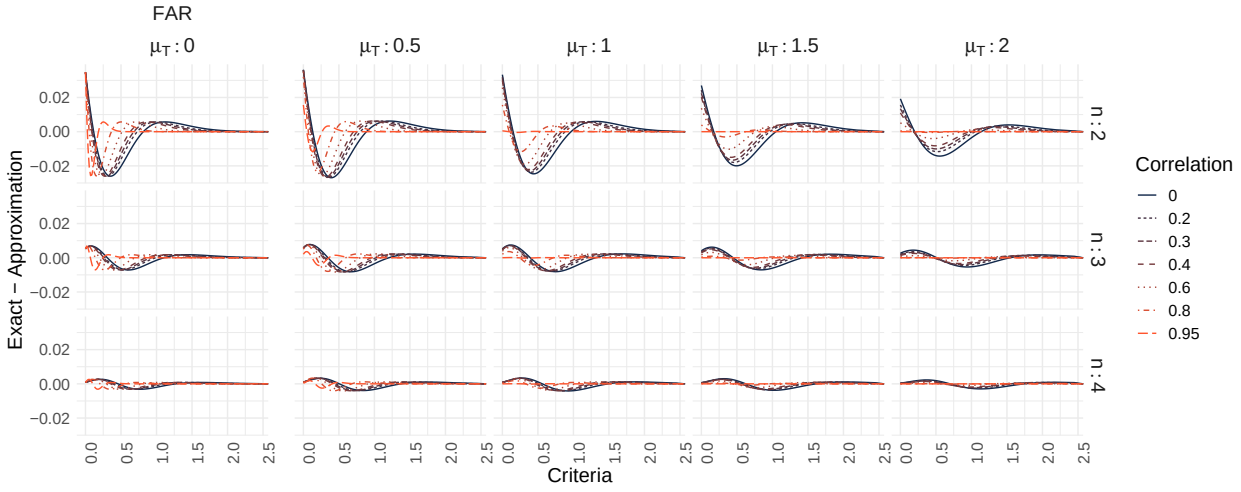
Discussion

Tasks involving unforced choice are abundant in psychology. We have introduced several important cases from research in perception, attention, and memory, with a particular focus on *lineup memory* tasks, in which a witness can indicate a particular member of a lineup as a previously seen perpetrator of a crime, or reject the lineup as failing to contain that perpetrator. The fact that that task involves a *relative* comparison among stimuli, as well as an *absolute* comparison between those stimuli and a memory trace, makes it a task of unforced choice.

Models of lineup memory inspired by multivariate signal-detection theory have proliferated in recent years (Akan et al., 2021; Duncan, 2006; Smith et al., 2020; Wixted

Figure 14

Difference in FAR and HR between Exact and Approximated computations



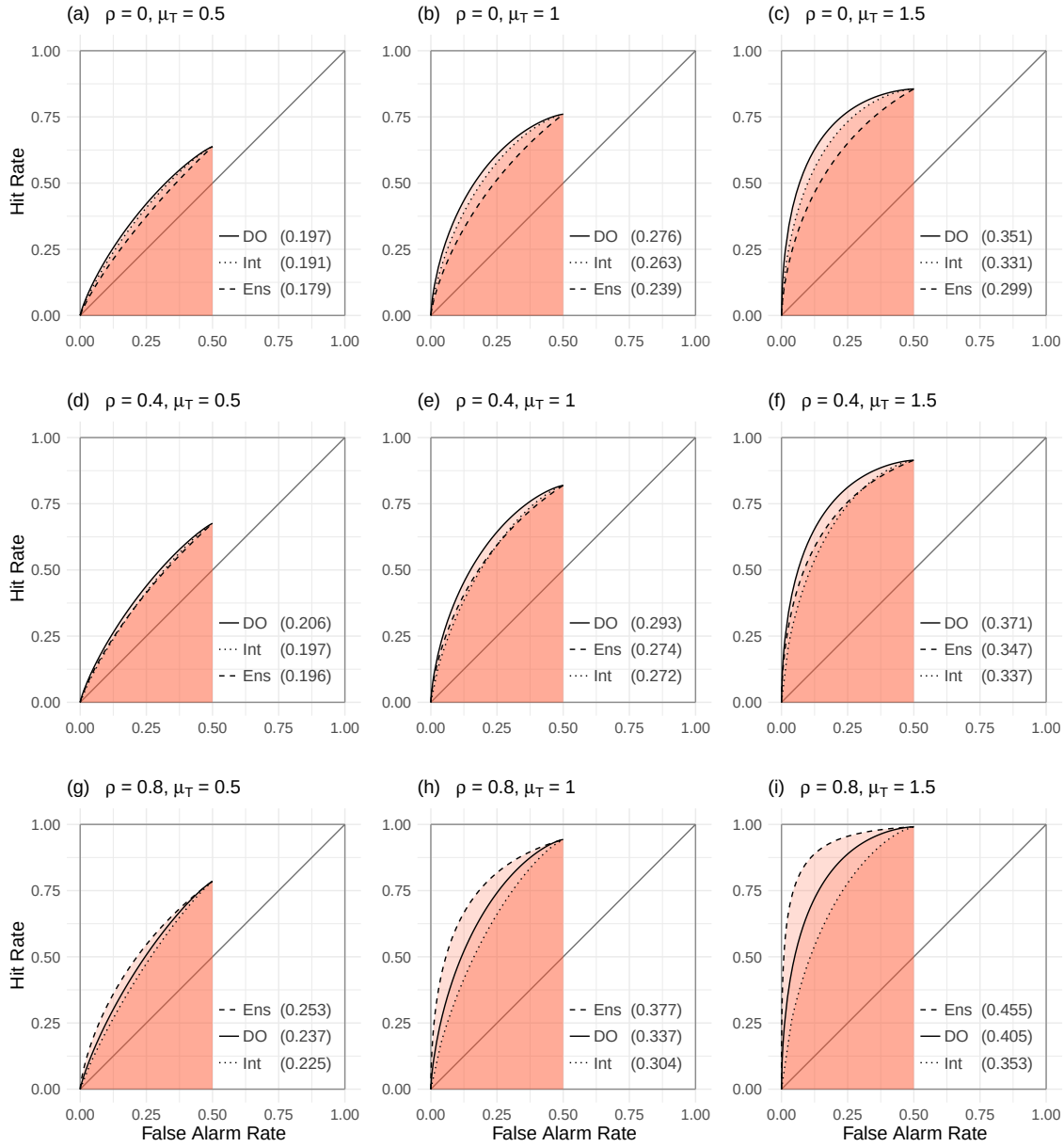
Note. Difference in predicted probability between the exact and the truncated normal approximation for the Ensemble model. FAR predictions correspond to $\mu_T = 0$, on the far left side of each function. HR predictions correspond to $\mu_T > 0$ across different set sizes (rows) and μ_T (columns).

et al., 2018), and have supported advances in both theoretical and empirical research in eyewitness memory. This article aims to take those theories and place them on solid mathematical footing, an act that should facilitate comparison among theories, additional model development, and generalization to a wide variety of tasks both within and adjacent to psychology.

The vehicle we have used to establish this framework is multivariate geometry (Wickens, 2014). Though the particular means by which we have developed and derived results in this paper may be unfamiliar to some, the general principles harken back to some of the earliest development in signal-detection theory (Green, 1964), and the general principles should be comfortable to anyone working within these domains.

Figure 15

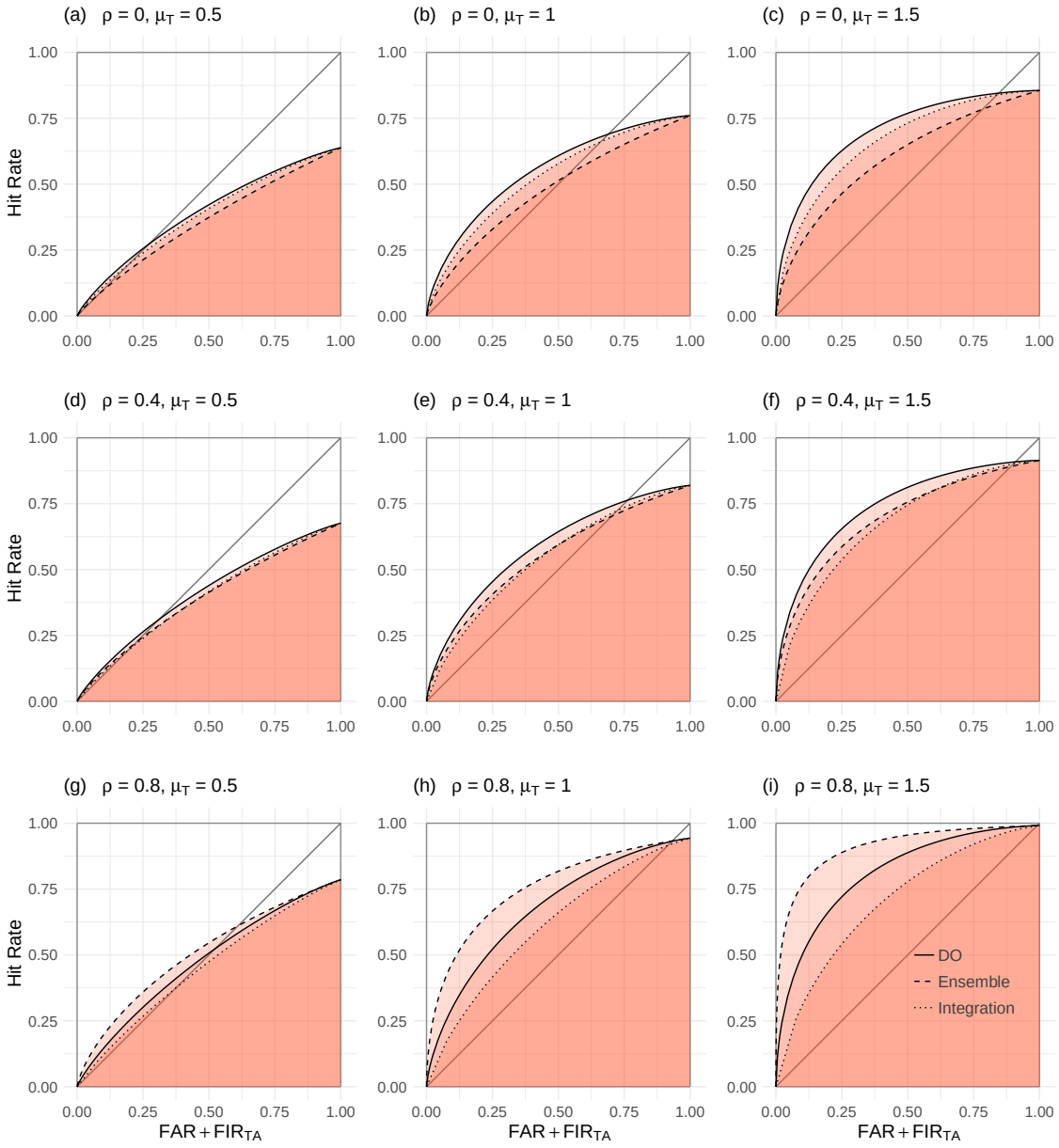
TOC curves for three models of unforced choice.



Note. Correlations among stimuli (ρ) vary across rows. Signal strength μ_T varies across columns. DO denotes the TOC for the *Dependent Observations* model (*Independent Observations* for $\rho = 0$); *Int* for the Integration model; and *Ens* for the Ensemble model. The numbers within parentheses indicate the area under the curve of the corresponding TOC.

Figure 16

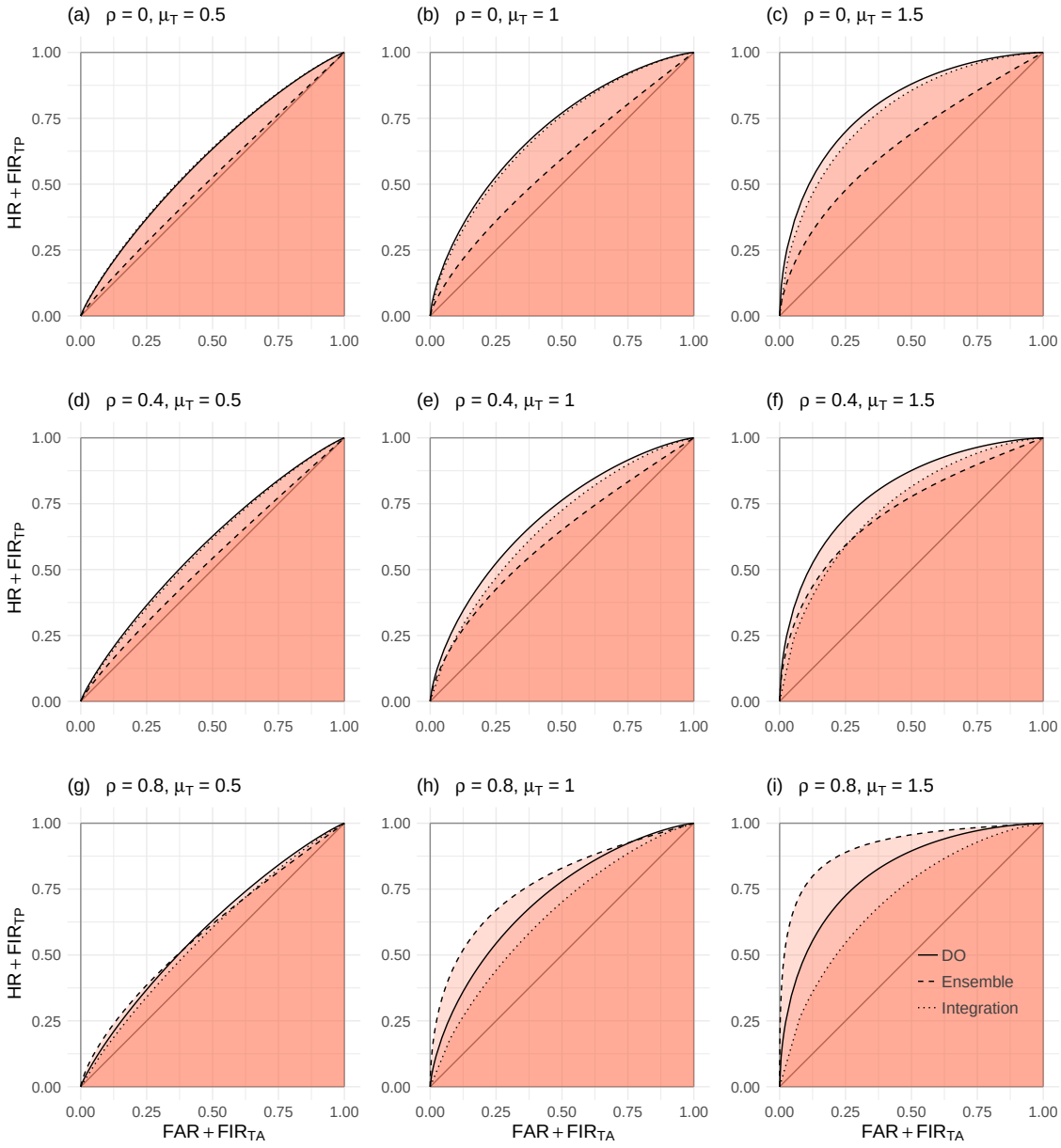
IOC curves for three models of unforced choice



Note. Correlations among stimuli (ρ) vary across rows. Signal strength μ_T varies across columns. DO denotes the IOC for the *Dependent Observations* model (*Independent Observations* for $\rho = 0$); *Int* for the Integration model; and *Ens* for the Ensemble model.

Figure 17

EOC curves for three models of unforced choice



Note. Correlations among stimuli (ρ) vary across rows. Signal strength μ_T varies across columns. DO denotes the EOC for the *Dependent Observations* model (*Independent Observations* for $\rho = 0$); *Int* for the Integration model; and *Ens* for the Ensemble model.

Model comparison

With exact formulations in hand, it is possible to draw more precise comparisons among models than before. Figures 15, 16, and 17 show all three types of ROC for each of the three major maodels presented in this paper: *Dependent Observations* in the top row (with *Independent Observations* as a special case in the leftmost panel), *Integration* in the middle row, and the *Ensemble* model in the bottom row.

With the three types of ROCs and the exact derivations of models presented here, the models can be meaningfully compared to determine their uniqueness and identifiability. This section provides a first attempt to do so in broad strokes, though we emphasize that extensive model comparison is beyond the goals of this paper and is left for future work.

Model mimicry. Figure 18 shows the results of an analysis of model mimicry using the major models presented in the paper. For each model, we generated one "simple" data set in which the distributions were of equal variance and the correlation between all array members was equal (with a nonzero value for all models except the IO model). We also simulated a more "complex" data set with the distributions having unequal variance and the correlation between target and fillers (ρ_1) being different than the correlation between fillers and fillers (ρ_0). Exact values are provided in the figure caption.

Each model was then fit to the two data sets generated by the four decision models. Fitting was constrained under three regimes. In the top panels, all parameters were estimated freely. In the middle panel, the structural relationship between the correlation parameters ($\rho_1 = \rho_0/\sigma_T$) was used as a constraint and all other parameters were estimated. Finally, in the bottom panel, the variance of the signal distribution and the relationship between the two correlation parameters ($\rho_1 = \rho_0/\sigma_T$) was provided and the remaining parameters were estimated.

The outcomes of this fitting exercise are shown as heatmaps, with the left panels indicating χ^2 values of deviation between the original parameter values and the recovered values, and the right panels indicating the p-value associated with that value of χ^2.

Notably, the DO and Integration model were completely able to mimic every model in the set of generating models, including the Ensemble model. The IO model also exhibited substantial flexibility, being able to mimic data from all of the models except the Ensemble model. The Ensemble model, in contrast, successfully recovered parameters only generated by the Ensemble model; it could not mimic data sets generated by the other models.

This outcome should be concerning for researchers who employ the IO, DO, or Integration models. It is not to say that those models are without merit—the parameters provided by a fitting exercise may be informative for interpreting the outcome of an experimental manipulation or relationship with a psychometric variable. However, it does seem that those models can not be meaningfully adjudicated among using model fit. Of course, there may be constraints beyond the very few employed here in which those models can be successfully differentiated, though we see no obvious way of determining what those constraints might be.

Model identifiability. Figure 19 shows the target-operating characteristics for the DO, Integration, and Ensemble models under a variety of parameter values for μ_T and ρ_0. In the upper panels, the signal and noise distributions are assumed to have equal variance, and the correlation among all array members is the same, regardless of whether they are distractors or the target. In the bottom panels, the variance of the signal distribution is allowed to vary across individual ROC functions, and the correlation between the target and the distractors differs from the correlation among the distractors themselves.

All of the TOCs within each panel of Figure 19 are constrained to have the same ending point. The goal of this exercise is to evaluate whether there is a unique TOC for each set of distribution parameters; a prerequisite for this comparison is that those TOCs end at the same point.

For both the DO and Integration models, it can be seen that different combinations of parameters yield different TOCs that nonetheless end at the same point. However, this

is not true for the ensemble model. In fact, the ending point of the TOC completely determines the position of the TOC in space, and that TOC can be reached via an infinite combination of distribution parameters. This is shown in the rightmost panels of Figure 19 for four different combinations of parameters. This outcome reveals that TOCs (and ROCs more generally) are an inadequate tool for estimating parameters within the ensemble model, and that the Ensemble model will need to be augmented with additional constraints to render it identifiable.

Model viability. The lessons from model comparison are sobering. None of the models have completely salutary characteristics. The DO and Integration models exhibit worrisome flexibility, a fact that compromises model comparison within this realm. The Ensemble model does not produce unique ROCs for each set of parameter values, making parameter recovery challenging if not impossible.

Theoretical development using these models will have to move beyond comparison of ROCs, even beyond comparison of combinations of different types of ROCs. Additional model constraints may be necessary; those constraints might be properly informed by the domain of study in which the models are being applied. Alternatively, model comparison using techniques beyond ROC analysis may be beneficial (e.g. Kellen & Klauer, 2014, 2015; Ratcliff & Starns, 2009)

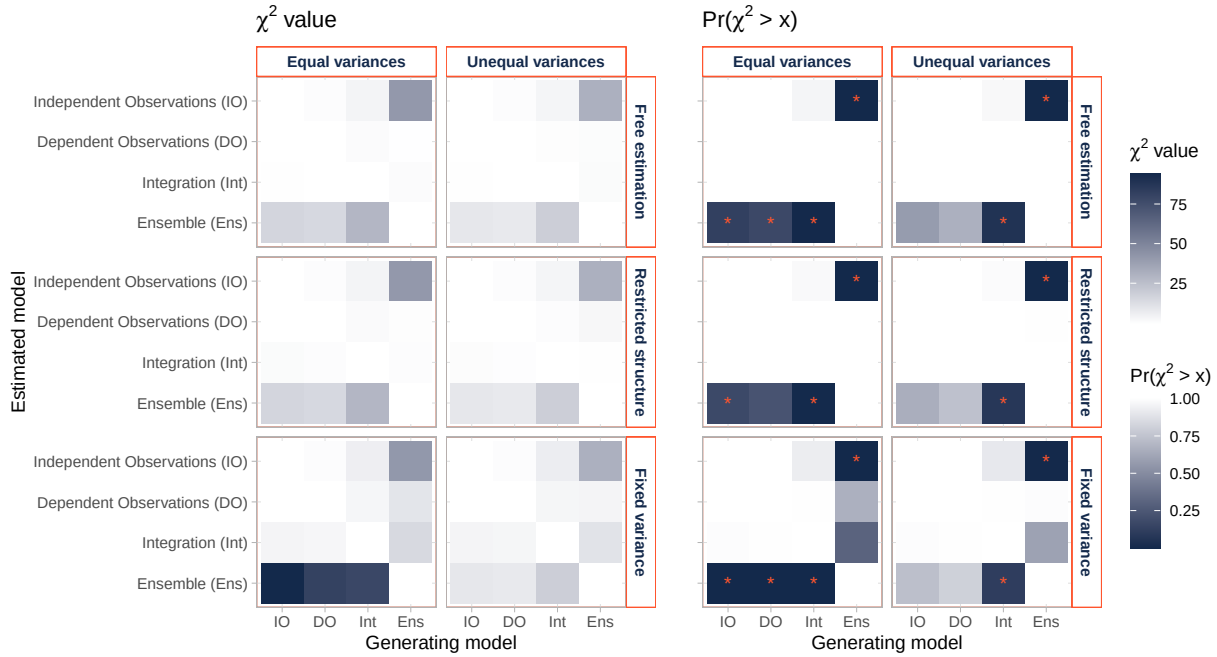
Software for model fitting

A software package accompanies this paper that is designed to enable use of the models presented here. The software can be accessed at: https://github.com/herulor/unforcedChoice. In this section we provide an overview of that software with guidance for its use in practical settings. The software has two purposes, though it has the potential for expansion to meet the demands of future users.

For model exploration, the user can enter model parameters and visualize target-operating characteristics (TOC), identification-operating characteristics (IOC), and

Figure 18

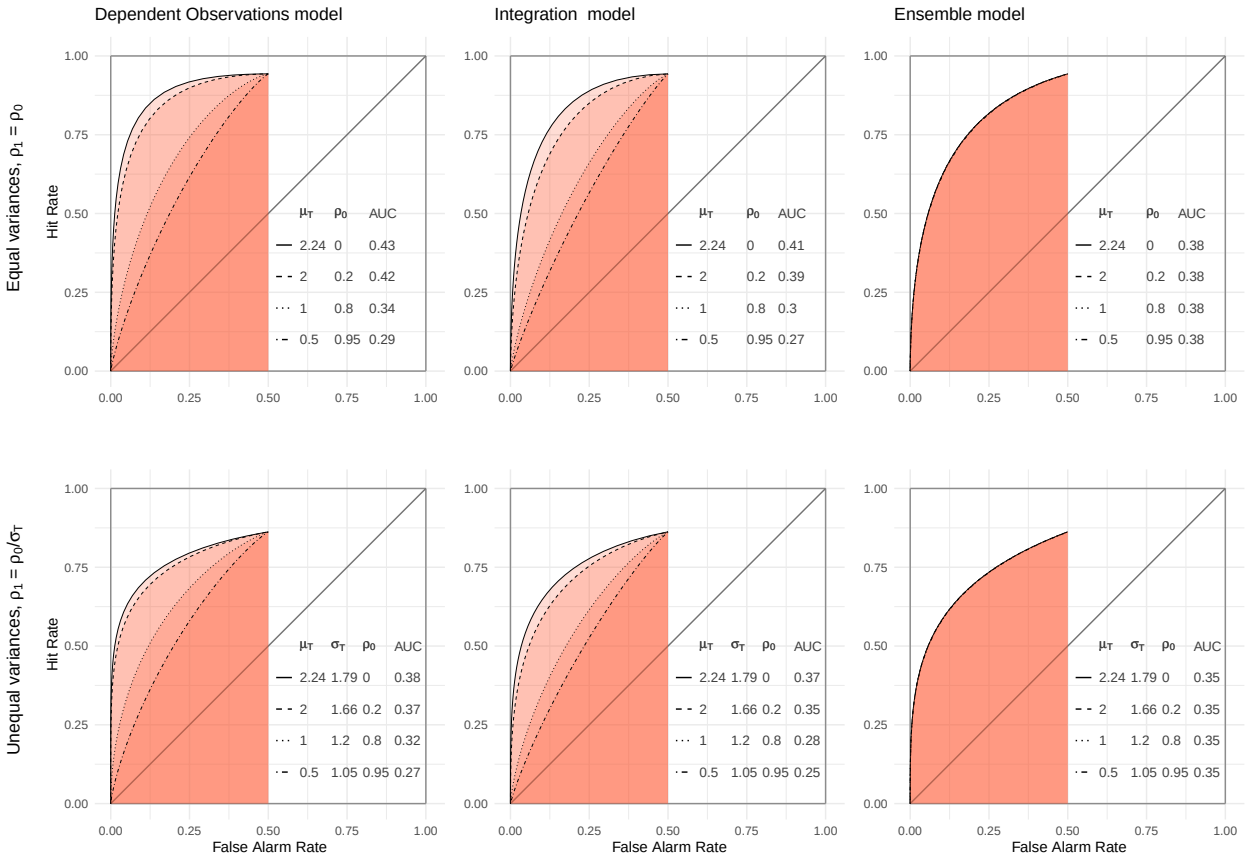
Analysis of model mimicry



Note. Two data sets were generated for each of the models (IO, DO, Integration, Ensemble). In the *equal variances* data set, the distribution parameters were $\mu = 1, \sigma_T = 1, \rho_0 = \rho_1 = .2$. In the *unequal variances* data set, they were $\mu = 1, \sigma_T = 1.1, \rho_0 = .22, \rho_1 = .2$. The distributions were assumed to satisfy the relation $\rho_1 = \rho_0/\sigma_T$. *Free estimation* indicates no constraints on parameter estimation. *Restricted structure* indicates that ρ_1 was not estimated and its relation with ρ_0 and σ_T was used to inform estimation and compute its implied value. *Fixed variance* estimation was computed assuming that the value of σ_T was known in addition to the structural information about correlations. The $*$ on some of the right panel tiles indicates where the probability of observing a χ^2 value under the null hypothesis as large as the one that was found is smaller than 0.2.

Figure 19

Analysis of model identifiability using TOCs



Note. TOCs for different combinations of model constraints, model parameters, and specific model. The left column shows the TOCs from the DO model, the middle column from the Integration model, and the right column from the Ensemble model. For each decision model, the upper panel shows TOCs for an equal-variance model with an equivalent correlation across array members. The bottom panel shows TOCs for models with unequal variance and a differential correlation between target and distractors as between distractors themselves. Parameter values for the four TOCs within each panel and indicated in the right side of each panel.

ensemble-operating characteristics (EOC) for all of the models presented in this paper. These different forms of the receiver-operating characteristics are presented in an earlier section in this paper. In addition, when criterion are specified by the user, HR, FAR, and Rejection Rates are also provided. The user also has the choice between working with the restricted versions of the models presented in the body of this article or the flexible versions in which variances and correlations among stimuli are free to vary.

For model fitting, the user can enter performance data and estimate model parameters and fits for any of the models presented in this paper. Again, both the restricted and flexible versions can be selected among. The software will also generate model ROCs with the data superimposed in a format similar to the figures presented in this paper. Those figures can be cut-and-pasted (with appropriate attribution) in other manuscripts and presentations.

Conclusions

Unforced-choice tasks are a valuable part of the experimentalist's arsenal, with strong connections to many important applied scenarios. The recent flurry of theoretical development in this area is a sign of a maturing field, one that we see in some need of organization. The framework presented in this paper is an attempt to bring that organization to models of unforced choice.

We have summarized here the tools for model development and for characterizing empirical data and model predictions in a manner that respects the approaches currently being taken and unifies those approaches within well validated principles of signal-detection theory. As part of that development, all of the models have been laid out in a relatively unconstrained form throughout the Appendices, yielding more flexible versions than those that have been worked with to date. Those models can be additionally constrained to yield new variants.

The major models summarized here all appear to have serious flaws. These flaws

would not have been evident without the derivations that follow from the multivariate framework. Model comparison is undermined by the ability of some models (DO, Integration) to mimic others. Parameter estimation is undermined by the unidentifiability of others (Ensemble). New techniques, new models, or new constraints may be needed to move forward in this domain, and it is out hope that the framework presented here can aid researchers in that quest.

References

Akan, M., Robinson, M. M., Mickes, L., Wixted, J. T., & Benjamin, A. S. (2021). The effect of lineup size on eyewitness identification. *Journal of Experimental Psychology: Applied*, *27*(2), 369–392. https://doi.org/10.1037/xap0000340

Apelt, D., & Peitgen, H.-O. (2008). Contrast sensitivity in mammographic softcopy reading – determination with psychophysical procedures. In E. A. Krupinski (Ed.), *Digital mammography* (pp. 756–763). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70538-3_104

Ashby, F. G. (1992). Multidimensional models of categorization. *Multidimensional models of perception and cognition.* (pp. 449–483). Lawrence Erlbaum Associates, Inc.

Ashby, F. G. (2014). *Multidimensional models of perception and cognition*. Psychology Press.

Ashby, F. G., & Soto, F. A. (2015). Multidimensional Signal Detection Theory. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford handbook of of computational and mathematical psychology*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199957996.013.2

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, *11*(4), 267–273. https://doi.org/10.1111/1467-9280.00254

Benjamin, A. S., & Bawa, S. (2004). Distractor plausibility and criterion placement in recognition. *Journal of Memory and Language*, *51*(2), 159–172. https://doi.org/10.1016/j.jml.2004.04.001

Benjamin, A. S., Diaz, M., & Wee, S. (2009). Signal detection with criterion noise: Applications to recognition memory. *Psychological Review*, *116*(1), 84–115. https://doi.org/10.1037/a0014351

Clark, S. E. (2003). A memory and decision model for eyewitness identification. *Applied Cognitive Psychology*, *17*(6), 629–654. https://doi.org/10.1002/acp.891

Clark, S. E., Erickson, M. A., & Breneman, J. (2011). Probative value of absolute and relative judgments in eyewitness identification. *Law and Human Behavior*, *35*(5), 364–380. https://doi.org/10.1007/s10979-010-9245-1

Colloff, M. F., Wilson, B. M., Seale-Carlisle, T. M., & Wixted, J. T. (2021). Optimizing the selection of fillers in police lineups. *Proceedings of the National Academy of Sciences*, *118*(8), e2017292118. https://doi.org/10.1073/pnas.2017292118

DeCarlo, L. T. (2012). On a signal detection approach to -alternative forced choice with bias, with maximum likelihood and Bayesian approaches to estimation. *Journal of Mathematical Psychology*, *56*(3), 196–207. https://doi.org/10.1016/j.jmp.2012.02.004

Duncan, M. (2006). *A signal detection model of compound decision tasks* (Technical Report).

Dzhafarov, E. N., Cervantes, V. H., & Kujala, J. V. (2017). Contextuality in canonical systems of random variables. *Philosophical Transactions of the Royal Society A*, *375*(2106), 20160389. https://doi.org/10.1098/rsta.2016.0389

Dzhafarov, E. N., & Gluhovsky, I. (2006). Notes on selective influence, probabilistic causality, and probabilistic dimensionality. *Journal of Mathematical Psychology*, *50*(4), 390–401. https://doi.org/10.1016/j.jmp.2006.03.003

Dzhafarov, E. N., & Kujala, J. V. (2012). Selectivity in probabilistic causality: Where psychology runs into quantum physics. *Journal of Mathematical Psychology*, *56*(1), 54–63. https://doi.org/10.1016/j.jmp.2011.12.003

Dzhafarov, E. N., & Kujala, J. V. (2014). On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Topics in Cognitive Science*, *6*(1), 121–128. https://doi.org/10.1111/tops.12060

Dzhafarov, E. N., Kujala, J. V., & Cervantes, V. H. (2016). Contextuality-by-Default: A brief overview of ideas, concepts, and terminology. In H. Atmanspacher, T. Filk, & E. Pothos (Eds.), *Quantum Interaction. LNCS* (pp. 12–23). Springer.

Dzhafarov, E. N., Kujala, J. V., Cervantes, V. H., Zhang, R., & Jones, M. (2016). On contextuality in behavioural data. *Philosophical Transactions of the Royal Society A*, *374*(2068), 20150234. https://doi.org/10.1098/rsta.2015.0234

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science*, *9*(2), 111–118. https://doi.org/10.1111/1467-9280.00020

Glanzer, M., Kim, K., Hilford, A., & Adams, J. K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 500.

Graham, N., Kramer, P., & Yager, D. (1987). Signal-detection models for multidimensional stimuli: Probability distributions and combination rules. *Journal of Mathematical Psychology*, *31*(4), 366–409. https://doi.org/10.1016/0022-2496(87)90021-6

Green, D. M. (1964). General prediction relating yes-no and forced-choice results. *The Journal of the Acoustical Society of America*, *36*(5), 1042–1042. https://doi.org/10.1121/1.2143339

Green, D. M., & Moses, F. L. (1966). On the equivalence of two recognition measures of short-term memory. *Psychological Bulletin*, *66*(3), 228–234. https://doi.org/10.1037/h0023645

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). Wiley New York.

Gronlund, S. D., & Benjamin, A. S. (2018). The new science of eyewitness memory. *Psychology of learning and motivation* (pp. 241–284). Elsevier. https://doi.org/10.1016/bs.plm.2018.09.006

Hacker, M. J., & Ratcliff, R. (1979). A revised table of d' for M-alternative forced choice. *Perception & Psychophysics*, *26*(2), 168–170. https://doi.org/10.3758/BF03208311

Jang, Y., Wixted, J. T., & Huber, D. E. (2009). Testing signal-detection models of yes/no and two-alternative forced-choice recognition memory. *Journal of Experimental Psychology: General*, *138*(2), 291–306. https://doi.org/10.1037/a0015525

Jesteadt, W., & Bilger, R. C. (1974). Intensity and frequency discrimination in one- and two-interval paradigms. *J. Acoust. Soc. Am.*, *55*(6), 11.

Kaernbach, C. (2001). Adaptive threshold estimation with unforced-choice tasks. *Perception & Psychophysics*, *63*(8), 1377–1388. https://doi.org/10.3758/BF03194549

Kellen, D., & Klauer, K. C. (2014). Discrete-state and continuous models of recognition memory: Testing core properties under minimal assumptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1795.

Kellen, D., & Klauer, K. C. (2015). Signal detection and threshold modeling of confidence-rating rocs: A critical test with minimal assumptions. *Psychological Review*, *122*(3), 542.

Kujala, J. V., & Dzhafarov, E. N. (2008). Testing for selectivity in the dependence of random variables on external factors. *Journal of Mathematical Psychology*, *52*(2), 128–144. https://doi.org/10.1016/j.jmp.2008.01.008

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

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport-security screening. *Psychological Science*, *15*(5), 302–306. https://doi.org/10.1111/j.0956-7976.2004.00673.x

Meyer-Grant, C. G., & Klauer, K. C. (2021). Monotonicity of rank order probabilities in signal detection models of simultaneous detection and identification. *Journal of Mathematical Psychology*, *105*, 102615.

Mueller, S. T., & Weidemann, C. T. (2008). Decision noise: An explanation for observed violations of signal detection theory. *Psychonomic Bulletin & Review*, *15*(3), 465–494. https://doi.org/10.3758/PBR.15.3.465

Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*(13), 1703–1721. https://doi.org/10.1016/0042-6989(94)90128-7

Perry, C. J., & Barron, A. B. (2013). Honey bees selectively avoid difficult choices. *Proceedings of the National Academy of Sciences*, *110*(47), 19155–19159. https://doi.org/10.1073/pnas.1314571110

Phelps, S. M., Rand, A. S., & Ryan, M. J. (2006). A cognitive framework for mate choice and species recognition. *The American Naturalist*, *167*(1), 28–42. https://doi.org/10.1086/498538

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using roc curves. *Psychological Review*, *99*(3), 518.

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological review*, *116*(1), 59.

Rotello, C. M., & Macmillan, N. A. (2007). Response bias in recognition memory. *Psychology of learning and motivation* (pp. 61–94). Elsevier. https://doi.org/10.1016/S0079-7421(07)48002-1

Schweickert, R., Fisher, D. D., & Kyongje, S. (2012). Selective influence of interdependent variables. *Discovering cognitive architecture by selectively influencing mental processes* (pp. 359–381). World Scientific Publishing Co.

Smith, A. M., Yang, Y., & Wells, G. L. (2020). Distinguishing between investigator discriminability and eyewitness discriminability: A method for creating full receiver operating characteristic curves of lineup identification performance. *Perspectives on Psychological Science*, *15*(3), 589–607. https://doi.org/10.1177/1745691620902426

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action and engagement across the mouse brain. *Nature*, *576*(7786), 266–273. https://doi.org/10.1038/s41586-019-1787-x

Taylor, M. M., & Fraser, W. C. G. (1966). A table of *d'* for a model of the unforced choice experiment. *Perceptual and Motor Skills*, *22*(1), 282–282. https://doi.org/10.2466/pms.1966.22.1.282

Thomas, R. D., Altieri, N. A., Silbert, N. H., Wenger, M. J., & Wessels, P. M. (2015). Multidimensional signal detection decision models of the uncertainty task: Application to face perception. *Journal of Mathematical Psychology*, *66*, 16–33. https://doi.org/10.1016/j.jmp.2015.03.001

Tong, Y. L. (1990). *The multivariate normal distribution*. Springer.

Watson, C. S., Kellogg, S. C., Kawanishi, D. T., & Lucas, P. A. (1973). The uncertain response in detection-oriented psychophysics. *Journal of Experimental Psychology*, *99*(2), 180–185. https://doi.org/10.1037/h0034736

Wells, G. L., Smith, A. M., & Smalarz, L. (2015). ROC analysis of lineups obscures information that is critical for both theoretical understanding and applied purposes. *Journal of Applied Research in Memory and Cognition*, *4*(4), 324–328. https://doi.org/10.1016/j.jarmac.2015.08.010

Wickelgren, W. A. (1968). Unidimensional strength theory and component analysis of noise in absolute and comparative judgments. *Journal of Mathematical Psychology*, *5*(1), 102–122. https://doi.org/10.1016/0022-2496(68)90059-X

Wickens, T. D. (2001). *Elementary signal detection theory*. Oxford university press.

Wickens, T. D. (2014). *The geometry of multivariate statistics*. Psychology Press. https://doi.org/10.4324/9781315806334

Wixted, J. T. (2020). The forgotten history of signal detection theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(2), 201.

Wixted, J. T., & Mickes, L. (2014). A signal-detection-based diagnostic-feature-detection model of eyewitness identification. *Psychological Review*, *121*(2), 262–276. https://doi.org/10.1037/a0035940

Wixted, J. T., Vul, E., Mickes, L., & Wilson, B. M. (2018). Models of lineup memory. *Cognitive Psychology*, *105*, 81–114. https://doi.org/10.1016/j.cogpsych.2018.06.001

Wolfe, J. M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 419–433. https://doi.org/10.1037/0096-1523.15.3.419

Zatka-Haas, P., Steinmetz, N. A., Carandini, M., & Harris, K. D. (2021). Sensory coding and the causal impact of mouse cortex in a visual decision. *eLife*, *10*, e63163. https://doi.org/10.7554/eLife.63163

Appendix A

Admissible covariance structures of n-alternative choice tasks

Formulating models of signal detection starts with assumptions about the distributions of evidence. In multivariate models, these assumptions are primarily captured in the nature of the variance-covariance matrix. The covariance structure is neither arbitrary nor it is completely constrained by theoretical assumptions. Here we start with rudimentary assumptions that are intended to be noncontroversial and show that the structure can be considerably constrained quite easily.

Under the assumption that events of a common type (either signal or noise) are drawn from a common distribution, the first restriction on admissible covariance structures $\mathbf{\Sigma}_{n\times n}$ is that variances of events drawn from the same population are equal. Moreover, because the variance of the noise population is always one, all entries in the diagonal of $\mathbf{\Sigma}_{n\times n}$ corresponding to a noise ensemble should be 1. Also note that relaxing the assumption of independence of a set of variables within a multivariate normal distribution is equivalent to allowing their correlations to differ from zero. We seek to identify how this correlation matrix can be constrained in a principled way. The main substantive consideration that guides the following discussion is that stimuli drawn from the same population are thought of as *essentially the same*.

Here, we present several means of formalizing this concept and show that each of these approaches leads to the same means of constraining the structure of $\mathbf{\Sigma}_{n\times n}$. The primary statistical vehicle for this work is the concept of exchangeability, whereas the primary theoretical consideration is the notion of selective influences. Under the constraints that follow from these varied approaches, the remaining Appendices present derivations for the unrestricted versions of the models (in which variances and covariances can vary) and the restricted versions of the models (in which signal and noise variance are assumed to be equal and a common covariance among all evidence types is assumed) that are discussed in the body of the paper.

Selectivity

It is natural to assume that for an n-alternative unforced choice task, the evidence elicited by each of the n stimuli is selectively influenced by the population to which each belongs. This assumption gives the most general interpretation of *essentially the same*: the definition of selective influences (Dzhafarov & Gluhovsky, 2006; Schweickert, Fisher, & Kyongje, 2012) states that the distribution of the evidence for each of the stimuli can be represented as some function of (p_i, U_i, C) for each $i = 1, \ldots, n$, where $p_i \in \{S, N\}$ indicates which population the i-th stimulus belongs to (signal or noise, respectively), and all U_i and C are independent random variables (or vectors) whose distribution does not depend on p_i. Under the assumption of (multivariate) normality of the distributions of the evidence elicited, without loss of generality, each U_i, $i = 1, \ldots, n$, can be taken to be standard normal and C to be multivariate normal, of some dimensionality. The variables U_i capture the variability unique to each of the $i = 1, \ldots, n$ stimuli and the vector C constitutes the source of any common variability among them.

The consequences of selectivity on the joint distributions of evidence are that, for any two stimuli i, j, the joint bivariate normal distribution of their evidence (θ_i, θ_j) has parameters $\mu(p_i)$, $\sigma(p_i)$, $\mu(p_j)$, $\sigma(p_j)$, and $\rho(p_i, p_j)$, and that $\rho(p_i, p_j)$ can be written as

$$\rho(p_i, p_j) = \sum_{k=1}^{m} a_k(p_i) a_k(p_j), \tag{51}$$

for some $m \geq 1$, where the functions a_k are subject to

$$\sum_{k=1}^{m} a_k^2(x) \leq 1 \tag{52}$$

(Kujala & Dzhafarov, 2008). Note that the functions $\mu(.)$ and $\sigma(.)$ can be chosen with additional constraints $\mu(N) = 0$ and $\sigma(N) = 1$, consistent with standard approaches in signal detection theory. Expressions (51) and (52) together are a necessary and sufficient condition for selectivity of the populations upon the evidence elicited when all target stimuli have the same normal distribution with mean μ_T and variance σ_T^2 and, analogously,

all noise stimuli have the same standard normal distribution (Propostitions 1 and 2, Kujala & Dzhafarov, 2008).

The form of the functions in Expression (51) means that the distribution of evidence for stimuli from the same distribution must all have the same (normal) distribution, and that the covariance structure for an array of n stimuli in an n-alternative unforced-choice task must satisfy the following structure

$$\mathbf{\Sigma}^{\mathrm{SI}}_{n\times n} = \begin{pmatrix} \sigma^2(p_1) & \vdots & \sigma(p_1)\rho(p_1,N)\vec{1}^{\intercal}_{n-1} \\ \cdots\cdots\cdots & \vdots & \cdots\cdots\cdots \\ \sigma(p_1)\rho(p_1,N)\vec{1}_{n-1} & \vdots & (1-\rho(N,N))\mathbf{I}_{(n-1)\times(n-1)} + \rho(N,N)\vec{1}_{n-1}\vec{1}^{\intercal}_{n-1} \end{pmatrix} \tag{53}$$

where the superscript SI indicates that the variables representing the evidence are selectively influenced by their respective population, signal or noise, and where $\rho(N,N) \geq 0$. That is, the correlation of evidence across all samples of noise stimuli is the same and is non-negative, and the correlation between evidence for the first stimulus (target or lure) with that of the remaining stimuli is constant. Additionally, if a trial consists of an array with more than one stimulus from the target population, the correlation between the evidence of any two of the target stimuli would be the same ($\rho(S,S)$) and non-negative.

Selectivity has an additional consequence on how the values of $\rho(p_i,p_j)$ relate to each other across different combinations of p_i and p_j. In fact, given the multivariate normality assumption, it is a criterion for the selectivity of the populations upon the evidence that does not require finding the functions a_k in equation (51). Letting $\rho_{p_ip_j}$ denote the correlation $\rho(p_i,p_j)$, there are only three different populations that are relevant to the n-alternative task: ρ_{NN}, $\rho_{NS} = \rho_{SN}$, and ρ_{SS}. The criterion for selectivity (cosphericity test, Kujala & Dzhafarov, 2008) can be written as

$$|\rho_{NS}(\rho_{NN} - \rho_{SS})| \leq \sqrt{1-\rho^2_{NS}}\left[\sqrt{1-\rho^2_{NN}} + \sqrt{1-\rho^2_{SS}}\right]. \tag{54}$$

This inequality can only be assessed when the three component correlations are known. However, a weaker converse that provides a sufficient condition based only on the two correlations known in Expression (53) can be found if several additional conditions are met.

If evidence is assumed to follow a multivariate normal distribution, with equal distribution of evidence within a population, and the covariance matrix of an array containing one target has the structure from Expression (53) with:

- $|\rho_{NS}| \leq \frac{\sqrt{2}}{2}$, or
- $\rho_{NN} \leq \frac{\sqrt{2}}{2}$ and $\frac{\sqrt{2}}{2} \leq |\rho_{NS}| \leq \cos\left(\frac{1}{2}\arccos\rho_{NN}\right)$, or
- $\rho_{NN} \geq \frac{\sqrt{2}}{2}$ and $\frac{\sqrt{2}}{2} \leq |\rho_{NS}| \leq \cos\left(\frac{1}{2}\arcsin\rho_{NN}\right)$,

then the evidence elicited by each stimulus is selectively influenced by the population to which it belongs. To prove this statement we note that, after algebraic manipulation, the inequality in Expression (54) is equivalent to

$$\max\left(|2\arcsin\rho_{NS} \pm \arcsin\rho_{NN} \mp \arcsin\rho_{SS}|\right) \leq \pi. \tag{55}$$

From Expression (55)[7], the regions described above are easily derived, and are shown in Figure A1. The main practical consequence of this result is that, provided the regular assumptions of multivariate normal distribution of evidence and equal distribution evidence for all stimuli from the same population, one can safely conclude that the population selectively influences the evidence in the task when the evidence from the same population is exchangeable (as defined below) with very mild restrictions on the correlations within and across populations.

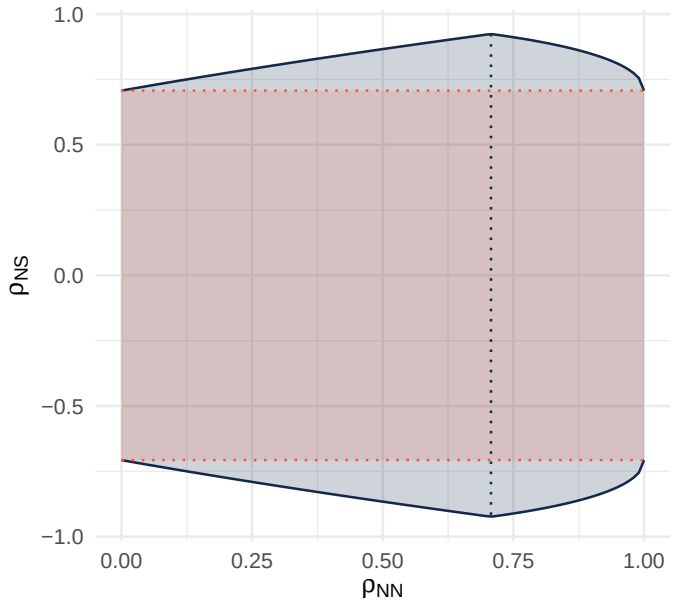
Exchangeability

From a statistical perspective, without assuming independence, the most restrictive interpretation of *essentially the same* is that the variables representing the evidence for

[7] This expression is a special case of the well known CHSH (Clauser-Horne-Shimony-Holt) inequalities for contextuality in quantum mechanics. The relation of this type of set of inequalities with selective influences in the psychological literature and with a more general theory of contextuality can be found in Dzhafarov, Cervantes, and Kujala (2017), Dzhafarov and Kujala (2012, 2014), Dzhafarov, Kujala, and Cervantes (2016), and Dzhafarov, Kujala, Cervantes, Zhang, and Jones (2016).

Figure A1

Selectivity regions for ρ_{NN} and ρ_{NS}



Note. The shaded regions show the values of ρ_{NN} and ρ_{NS} for which it is possible to affirm that the evidence elicited by each stimulus in the n-alternative choice task is selectively influenced by the population to which each belongs. The rectangle bounded by red dashed lines and the margin of the figure indicates a region where the value of ρ_{NS} is sufficient. Both values of ρ_{NS} and ρ_{NN} need to be considered in the adjacent regions. Values outside the shaded area indicate that nothing can be affirmed about selectivity without additional knowledge of ρ_{SS}. See the text for the functions that delimit these regions.

stimuli from the same population are exchangeable. Intuitively, a set of exchangeable random variables is one in which, if someone replaced one variable with another within the set, or if one stimulus was replaced by another from the same population, one would not be able to identify, in the long run, that a change occurred. Probabilistically, this means that the joint distribution is unchanged by any permutation of the random variables; that is, if $\{X_1, \ldots, X_n\}$ is a set of exchangeable random variables and F is their joint distribution function, then for any permutation π of the variables, $F(\pi(\vec{x}_n)) = F(\vec{x}_n)$. Formally, the definition of exchangeability involves infinite sequences of random variables. In addition to the symmetry with respect to permutations just mentioned, if there exists a sequence of variables $\{Y\}_{i=1}^{\infty}$ such that for every m the variables $\{Y_1, \ldots, Y_m\}$ are permutation-symmetric, and such that $\{Y_1, \ldots, Y_n\}$ have the same distribution as $\{X_1, \ldots, X_n\}$, then the variables $\{X_1, \ldots, X_n\}$ are exchangeable (see Section 5.3.1 of Tong, 1990).

In the present case, with multivariate normally distributed evidence, exchangeability simply requires that, in addition to the same mean and variance, any two variables from the same population have the same (non-negative) correlation (Theorem 5.3.1, p. 112, Tong, 1990). Within TA arrays, all stimuli are drawn from a common population and exchangeability implies that

$$\mathbf{\Sigma}_{n\times n} = (1 - \rho_{\mathrm{NN}})\mathbf{I}_{n\times n} + \rho_{\mathrm{NN}}\vec{1}_n\vec{1}_n^{\intercal},$$

where ρ_{NN} is the common correlation among fillers.

For TP arrays, the filler items are exchangeable but the target is not exchangeable with the fillers. Taken together, this pair of assumptions implies a covariance matrix in which

$$\mathbf{\Sigma}_{n\times n} = \begin{pmatrix} \sigma_T^2 & \vdots & \sigma_T\rho_{\mathrm{NS}}\vec{1}_{n-1}^{\intercal} \\ \cdots\cdots & \vdots & \cdots\cdots\cdots\cdots \\ \sigma_T\rho_{\mathrm{NS}}\vec{1}_{n-1} & \vdots & \left[(1-\rho_{\mathrm{NN}})\mathbf{I}_{(n-1)\times(n-1)} + \rho_{\mathrm{NN}}\vec{1}_{n-1}\vec{1}_{n-1}^{\intercal}\right] \end{pmatrix}$$

where σ_T is the variance of the target population and ρ_{NS} is the correlation between the target and any filler.

We can parametrize both covariance matrices within a single expression by letting $\rho_0 = \rho_{NN}$, and setting $\sigma_1 = 1$ and $\rho_1 = \rho_{\text{NN}}$ for TA arrays and $\sigma_1 = \sigma_T$ and $\rho_1 = \rho_{\text{NS}}$ for TP arrays:

$$\mathbf{\Sigma}^{\text{EW}}_{n\times n} = \begin{pmatrix} \sigma_1^2 & \vdots & \sigma_1\rho_1\vec{1}^{\intercal}_{n-1} \\ \cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sigma_1\rho_1\vec{1}_{n-1} & \vdots & (1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)} + \rho_0\vec{1}_{n-1}\vec{1}^{\intercal}_{n-1} \end{pmatrix} \tag{56}$$

where the superscript EW indicates that the variables representing the evidence are exchangeable only within their respective population (signal or noise).

More restrictive special cases of this covariance structure are found by making the correlation between all variables the same (that is, by setting $\rho_1 = \rho_0$), or by assuming equal variances for both populations, hence making $\sigma_T = 1$. Under both restrictions, the covariance matrix simplifies to

$$\mathbf{\Sigma}^{\text{AE}}_{n\times n} = (1-\rho_0)\mathbf{I}_{n\times n} + \rho_0\vec{1}_n\vec{1}^{\intercal}_n, \tag{57}$$

where AE stands for all variables having equal variance and covariance, which provides the structure for the restricted models presented in the body of the paper. For these models, selectivity is assured whenever the common correlation (and covariance, since variances all equal 1) is less than $\sqrt{3/4}$, and needs to be further investigated otherwise.

Next, we introduce some additional substantive considerations on how the covariance structure may be restricted by looking at two important special cases of selective influences.

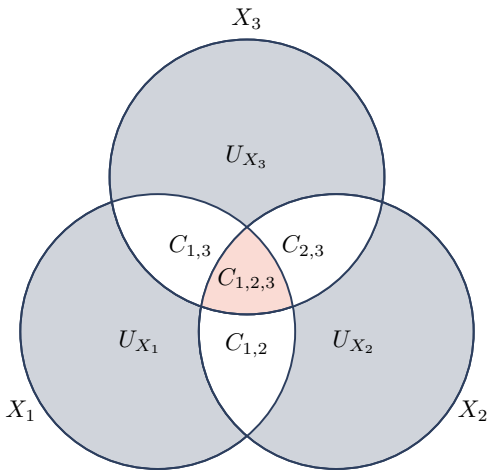
Partitioning of correlations into independent sources

It is a consequence of the definition of covariance that any two variables X and Y can be rewritten as the sum of three independent variables C, U_X, and U_Y, where C captures all the common variability of X and Y. Note that the specific C depends on both X and Y. One alternative instantiation of selectivity is that the same unidimensional source of common variability C is shared by all evidence coming from stimuli of the same

population, such that, conditioned on any value of $C = c$, all memory traces are independent. This framing is illustrated in Figure A2.

Figure A2

Variance decomposition into common and unique sources



Note. Illustration of the decomposition of the variance of three variables (X_1, X_2, and X_3) into variances from shared sources ($C_{1,2}$, $C_{2,3}$, $C_{1,3}$, and $C_{1,2,3}$) and unique variance (U_{X_1}, U_{X_2}, and U_{X_3}). The variance of the common source $C_{1,2,3}$ is shared by the three variables X_1, X_2, and X_3; variances of common sources $C_{i,j}$ are shared by each pair X_i and X_j only. Only if all three, $C_{1,2}$, $C_{2,3}$, and $C_{1,3}$, are the null set, is it possible to comprehensively model the three variables, X_1, X_2, and X_3, with a single common variance component.

If we assume that the evidence from the same population can be decomposed as having a single source of variance C, then one can write each evidence variable as:

$$\theta_i = \mu^*_{p_i} + C + U_i$$

with variance given by

$$\sigma^2_{\text{CS}} = \text{Var}(C) + \text{Var}(U_i),$$

where CS indicates that this variance has the same common source for any two variables. Note that this decomposition implies that the sources of unique variance for evidence from

the same population must have the same variance, i.e., that $\mathrm{Var}(U_i) = \mathrm{Var}(U_j)$, for any two θ_i, θ_j from the same population. For any two samples of evidence from the same population, the covariance is

$$\begin{aligned}\mathrm{Cov}(\theta_i, \theta_j) &= \mathrm{Cov}(C + U_i, C + U_j) \\ &= \mathrm{Var}(C);\end{aligned}$$

hence, their correlation is

$$\mathrm{Cor}(\theta_i, \theta_j) = \frac{\mathrm{Var}(C)}{\sigma^2_{\mathrm{CS}}}.$$

That is, if we assume this particular decomposition, then all pairwise sets of evidence drawn from a common population must have the same common correlation, and their corresponding variables are symmetric to permutations (as defined above). This decomposition in combination with symmetry to permutations is equivalent to the condition of exchangeability of evidence from the same population (Proposition 5.3.1, Tong, 1990), and it satisfies the definition of selectivity where the evidence can be written as

$$\theta_i = \sigma^*(p_i)\left[C_0 + \mathbb{I}(p_i)C_1 + (1 - \mathbb{I}(p_i))C_2 + Z_i\right] + \mu^*(p_i),$$

where C_0, C_1, C_2, and Z_i, $i = 1, \ldots, n$, are independent standard normally distributed random variables, and $\mathbb{I}(\cdot)$, $\mu^*(\cdot)$, and $\sigma^*(\cdot)$ are functions of the population with $\mathbb{I}(N) = 1 - \mathbb{I}(S) = 1$, $\sigma^*(N) = \sqrt{1/3}$, $\sigma^*(S) = \sqrt{\sigma^2_T/3}$, $\mu^*(N) = 0$, and $\mu^*(S) = \mu_T$.

This approach yields the same covariance structure as discussed in the Exchangeability subsection, and also provides a further means to explore the nature of the constraint on the covariance matrix.

Additional constraints on the covariance matrix from the independent-dependent sources formulation

There are two conceptual possibilities with respect to the question of shared variance across signal and noise populations. The simplest possibility is that the variance

shared by evidence within each population equals the variance shared by all evidence across populations—that is, across noise and signal populations, as shown in Figure A2. In that case, because the variance of the noise distribution is 1, the correlation between noise events is

$$\rho_0 = \mathrm{Cor}(\theta_{F_i}, \theta_{F_j}) = \mathrm{Var}(C),$$

the correlation between target events with standard deviation σ_T is

$$\mathrm{Cor}(\theta_{T_i}, \theta_{T_j}) = \frac{\mathrm{Var}(C)}{\sigma_T^2} = \frac{\rho_0}{\sigma_T^2},$$

and the correlation between a noise stimulus (filler) and a signal (target) is

$$\rho_1 = \frac{\mathrm{Var}(C)}{\sigma_T} = \frac{\rho_0}{\sigma_T}.$$

The alternative scenario is one in which variance is shared within a population of events but not across those populations. In that case, there are three additional possibilities, depending on which class of stimuli shares additional variance among their members than what they also share with the members of the other class. Let us denote as C_{FF} the variance shared by fillers but not between fillers and targets, and as C_{TT} the variance shared by targets but not by targets and fillers. The outcomes of these competing possibilities for the source of covariance are shown in Table A1.

Table A1

Possible sources of covariance within and between populations of events

	All variance is common	Some additional variance within type of stimuli		
	$C_{\mathrm{FF}} = 0$ and $C_{\mathrm{TT}} = 0$	$C_{\mathrm{FF}} = 0$ and $C_{\mathrm{TT}} > 0$	$C_{\mathrm{FF}} > 0$ and $C_{\mathrm{TT}} = 0$	$C_{\mathrm{FF}} > 0$ and $C_{\mathrm{TT}} > 0$
$\mathrm{Cov}(\theta_{\mathrm{T}_i}, \theta_{\mathrm{F}_j})$	ρ_0	ρ_0	$\rho_0 - C_{\mathrm{FF}}$	$\rho_0 - C_{\mathrm{FF}}$
$\mathrm{Cov}(\theta_{\mathrm{T}_i}, \theta_{\mathrm{T}_j})$	ρ_0	$\rho_0 + C_{\mathrm{TT}}$	ρ_0	$\rho_0 + C_{\mathrm{TT}} - C_{\mathrm{FF}}$

In all, if one assumes that all samples of evidence share the same common source of variance as the fillers, then all off-diagonal elements of $\mathbf{\Sigma}_{n\times n}$ are the same, and the matrix

is parametrized as

$$\mathbf{\Sigma}^{\text{CS}}_{n\times n} = \begin{pmatrix} \sigma_1^2 & \vdots & \rho_0 \vec{1}^{\intercal}_{n-1} \\ \cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \rho_0 \vec{1}_{n-1} & \vdots & \left[(1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)} + \rho_0 \vec{1}_{n-1}\vec{1}^{\intercal}_{n-1}\right] \end{pmatrix}. \tag{58}$$

It is worth to emphasize that the matrix $\mathbf{\Sigma}^{\text{CS}}_{n\times n}$ is a covariance matrix. That is, although ρ_0 is the correlation among fillers, expression (58) indicates that its value equals the covariance, not the correlation, between fillers and targets under the assumption of a single common source of shared variance. Hence, the assumption of a single common source of variance accounting for all correlations regardless of the population of origin implies a very specific relationship between the correlation of the target with the fillers. This correlation must equal the ratio between the correlation among fillers and the variance of the target population.

Factor analytical decomposition

A more general approach to formalize the common sources of variation introduced in the previous section is through the lens of factor analysis. Factor analysis decomposes a covariance matrix into a (smaller) set of common factors, such that each of the variables can be expressed as a linear combination of the factors plus a unique independent residual. A covariance matrix is perfectly factorizable if all off-diagonal elements are perfectly reproduced by the quadratic form of the factor loadings and the factors' covariance matrix. If the number of factors is not smaller than the number of variables whose covariance matrix is factor analyzed, any covariance structure is (trivially) perfectly factorizable.

Most reasonably, we should expect that the variables representing evidence from the same population to share a single factor, since it should account for any set size $n \geq 1$. Furthermore, any factor analyzable set of variables, in general, satisfies selective influence of the set of (experimental, for example) factors that determine the factors (dimensions) of the analysis: The functions are directly taken from the factor decomposition, whereas the common and unique factors play the role of C and U_i.

If one assumes that the population to which each stimulus belongs selectively influences their corresponding evidence, then the factor weights for all stimuli from the same population will be the same. It is easy to see that in that case, only a single factor (dimension) is needed in the factor representation of evidence from the same population. In this manner, the factor representation of the evidence yields the same constraint of exchangeability within populations as represented by $\mathbf{\Sigma}^{CS}_{n\times n}$ in Expression (58).

Hence, we see that under all considered approaches the covariance structure given by expression (56) already gives a reasonably general yet strongly constrained covariance structure for evidence in n-alternative choice. We will develop the models under this general structure, and also for the special cases of (1) a single common source of variance across populations, and (2) of a single common source plus equal variances of both target and noise populations. For all models, it is worth noting that the expressions for FAR and $\mathrm{RR}_{\mathrm{TA}}$ for TA arrays do not change when a more restricted covariance is considered, since the covariance structure within each population already corresponds to the most restricted one.

Appendix B

Forced-choice model

In the n-alternative forced-choice model, the region where the (arbitrarily) first stimulus is selected is defined by the following inequality:

$$\left(\vec{1}_{n-1} \vdots -\mathbf{I}_{(n-1)\times(n-1)}\right) \vec{\theta}_n > \vec{0}_{n-1}, \qquad \text{(4: revisited)}$$

where the inequality is read component-wise. We parameterize the means of the target and noise populations as μ_1, with $\mu_1 = \mu_T$ for TP arrays and $\mu_1 = 0$ for TA arrays.
Additionally, since the matrix in (4) represents the differences used to define the MAX rule used in all models, we notate it as

$$\mathbf{M}_{(n-1)\times n} = \left(\vec{1}_{n-1} \vdots -\mathbf{I}_{(n-1)\times(n-1)}\right) \tag{59}$$

Exchangeability

Under exchangeability within each population, the evidence vector has a multivariate normal distribution with mean and covariance matrix:

$$\vec{\mu}_n = \begin{pmatrix} \mu_1 \\ \cdots\cdots \\ \vec{0}_{n-1} \end{pmatrix} \qquad \text{(2: revisited)}$$

$$\mathbf{\Sigma}^{\mathrm{EW}}_{n\times n} = \begin{pmatrix} \sigma_1^2 & \vdots & \sigma_1\rho_1\vec{1}^{\intercal}_{n-1} \\ \cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \sigma_1\rho_1\vec{1}_{n-1} & \vdots & \left[(1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)} + \rho_0\vec{1}_{n-1}\vec{1}^{\intercal}_{n-1}\right] \end{pmatrix} \qquad \text{(56: revisited)}$$

Thus, the left-hand side of expression (4) has a multivariate normal distribution with mean vector $\vec{\mu}_{\mathrm{FC}}$ and covariance matrix $\mathbf{\Sigma}^{\mathrm{FC}}_{(n-1)\times(n-1)}$:

$$\vec{\mu}_{\mathrm{FC}} = \mathbf{M}_{(n-1)\times n}\vec{\mu}_n = \mu_1\vec{1}_{n-1} \tag{60}$$

$$\begin{aligned} \mathbf{\Sigma}^{\mathrm{FC}}_{(n-1)\times(n-1)} &= \mathbf{M}_{(n-1)\times n}\mathbf{\Sigma}^{\mathrm{EW}}_{n\times n}\mathbf{M}^{\intercal}_{n\times(n-1)} \\ &= (1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)} + (\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0)\vec{1}_{n-1}\vec{1}^{\intercal}_{n-1} \end{aligned} \tag{61}$$

From expression (61), we see that all differences $\theta_1 - \theta_i$, $i = 2, \ldots, n$, are identically normally distributed, with mean μ_1 and variance

$$\sigma^2_{\text{FC}} = \sigma_1^2 - 2\sigma_1\rho_1 + 1. \tag{62}$$

Therefore, all pairs of differences have the same correlation

$$\begin{aligned} \rho_{\text{FC}} =& \frac{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}{\sigma_1^2 - 2\sigma_1\rho_1 + 1} && (63) \\ =& \frac{\sigma^2_{\text{FC}} - (1 - \rho_0)}{\sigma^2_{\text{FC}}}. && (64) \end{aligned}$$

Therefore, the generic probability of selecting the (arbitrarily) first stimulus in an n-alternative forced-choice task is given by:

$$\begin{aligned} &\Pr(\theta_1 - \theta_2 > 0, \ldots, \theta_1 - \theta_n > 0 \,|\, \mu_1, \sigma_1, \rho_0, \rho_1) \\ &\quad = \Pr(\theta_2 - \theta_1 < 0, \ldots, \theta_n - \theta_1 < 0 \,|\, \mu_1, \sigma_1, \rho_0, \rho_1) \\ &\quad = \Pr(\theta_2 - \theta_1 + \mu_1 < \mu_1, \ldots, \theta_n - \theta_1 + \mu_1 < \mu_1 \,|\, \mu_1, \sigma_1, \rho_0, \rho_1) \\ &\quad = \Pr\left(Z_1 < \frac{\mu_1}{\sigma_{\text{FC}}}, \ldots, Z_{n-1} < \frac{\mu_1}{\sigma_{\text{FC}}} \,\middle|\, \mu_1, \sigma_1, \rho_0, \rho_1 \right) \\ &\quad = \int_{-\infty}^{\infty} \varphi(x) \prod_{i=1}^{n-1} \left[\Phi\left(\frac{\frac{\mu_1}{\sigma_{\text{FC}}} + \sqrt{\rho_{\text{FC}}} x}{\sqrt{1 - \rho_{\text{FC}}}} \right) \right] dx \\ &\quad = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\frac{\mu_1}{\sigma_{\text{FC}}} + \sqrt{\frac{\sigma^2_{\text{FC}} - 1 + \rho_0}{\sigma^2_{\text{FC}}}} x}{\sqrt{1 - \frac{\sigma^2_{\text{FC}} - 1 + \rho_0}{\sigma^2_{\text{FC}}}}} \right) \right]^{n-1} dx \\ &\quad = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_1}{\sqrt{1 - \rho_0}} + \sqrt{\frac{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}{1 - \rho_0}} x \right) \right]^{n-1} dx. \end{aligned} \tag{65}$$

The fifth line in (65) follows from Equation (8.3.1) in Tong (1990, p. 193), provided that ρ_{FC} is non-negative. As shown in the first row of Table A1, the covariance term $\sigma_1\rho_1$ of matrix $\boldsymbol{\Sigma}^{\text{EW}}_{n\times n}$ is less or equal to ρ_0; hence,

$$\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0 \geq \sigma_1^2 - \rho_0.$$

The right-hand side is non-negative whenever $\sigma_1 \geq \rho_0$. In most practical cases, it has been observed that the target variability is no smaller than the noise variability, that is, $\sigma^2 \geq 1$.

Hence, the non-negativity condition is always satisfied for TA arrays under the exchangeability assumptions, and is expected to also hold for TP arrays in most applications. Moreover, it should be noted that the correlation ρ_{FC} does not depend on the set size n. However, a correlation matrix with a common negative correlation among all variables must have correlations no smaller than $-\frac{1}{n-1}$. If ρ_{FC} is negative, then there exists some n such that $\rho_{\text{FC}} < -\frac{1}{n-1}$. Therefore, if in any practical application it is feasible that $\rho_{\text{FC}} < 0$, then one must reject the assumption of exchangeability of the evidence within populations.

We can specify expression (65) to compute the Hit and False alarm rates as:

$$\text{HR} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma^2 - 2\sigma\rho_1 + \rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx \tag{66}$$

and

$$\begin{aligned} \text{FAR} &= \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\sqrt{\frac{1-2\rho_0+\rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx \\ &= \int_{-\infty}^{\infty} \varphi(x) \left[\Phi(x) \right]^{n-1} dx \\ &= \frac{1}{n}, \end{aligned} \tag{67}$$

where the last step is completed by noting that

$$\frac{d}{dx} \left[\Phi(x) \right]^n = n\varphi(x) \left[\Phi(x) \right]^{n-1}.$$

Restricted cases

The HR expressions for the case in which a single common source of variance accounts for all the covariance among the evidence is obtained by replacing ρ_1 with ρ_0/σ Thus, for a single common source, the HR simplifies to:

$$\text{HR} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma^2 - \rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx, \tag{68}$$

and for the case with a single common source of variance and with equal variances across populations (setting $\sigma = 1$), the HR is:

$$\text{HR} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} - x \right) \right]^{n-1} dx. \tag{69}$$

Equation (69) has been derived from independence by Elliot (cited by Hacker & Ratcliff, 1979). A simple change of variables in expressions (66)–(69) reveals that it is impossible to identify whether the evidence is uncorrelated or not by exclusively using a forced-choice design.

Appendix C

Dependent (and Independent) Observations model

The regions corresponding to the decisions to reject the lineup and to choose the (arbitrarily) first stimulus in a lineup can be written as:

$$\mathbf{I}_{n\times n}\ \vec{\theta}_n < c\vec{1}_n, \tag{15: revisited}$$

and

$$\begin{pmatrix} 1 & \vdots & \vec{0}_{n-1}^{\intercal} \\ \cdots & \vdots & \cdots \\ \vec{1}_{n-1} & \vdots & -\mathbf{I}_{(n-1)\times(n-1)} \end{pmatrix} \vec{\theta}_n > \begin{pmatrix} c \\ \cdots \\ \vec{0}_{n-1} \end{pmatrix}, \tag{16: revisited}$$

respectively. Note that the matrix used for the rejection region is the identity matrix, $\mathbf{I}_{n\times n}$, and that the matrix that defines the selection of the first stimulus can be written as:

$$\mathbf{D}_{n\times n} = \begin{pmatrix} \vec{e}_{1,n}^{\intercal} \\ \cdots \\ \mathbf{M}_{(n-1)\times n} \end{pmatrix}, \tag{70}$$

where $\vec{e}_{1,n}$ is the vector with all but the first coefficient equal to zero, and $\mathbf{M}_{(n-1)\times n}$ as defined in (59).

Exchangeability

Under exchangeability within each population, the evidence vector has a multivariate normal distribution with mean and covariance matrix:

$$\vec{\mu}_n = \begin{pmatrix} \mu_1 \\ \cdots \\ \vec{0}_{n-1} \end{pmatrix} \tag{2: revisited}$$

$$\mathbf{\Sigma}^{\mathrm{EW}}_{n\times n} = \begin{pmatrix} \sigma_1^2 & \vdots & \sigma_1\rho_1\vec{1}_{n-1}^{\intercal} \\ \cdots & \vdots & \cdots \\ \sigma_1\rho_1\vec{1}_{n-1} & \vdots & \left[(1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)} + \rho_0\vec{1}_{n-1}\vec{1}_{n-1}^{\intercal}\right] \end{pmatrix} \tag{56: revisited}$$

By the general assumptions of the models, the correlations among the fillers are all equal to ρ_0 and the correlations between each filler and the target are all equal to ρ_1. Therefore, we can write them as $\rho_0 = \lambda_j\lambda_k$ and $\rho_1 = \lambda_1\lambda_j$, for $j, k = 2, \ldots, n$, with $\lambda_1 = {}^{\rho_1}/_{\sqrt{\rho_0}}$ and $\lambda_j = \sqrt{\rho_0}, j = 2, \ldots, n$. Based on this decomposition of the correlations, we

can find the general expression for the probability of rejecting the lineup under the Dependent Observations model as follows:

$$\begin{aligned}\Pr(\text{Reject}) &= \Pr(\theta_1 < c, \ldots, \theta_n < c) \\ &= \Pr(\theta_1 - \mu_1 < c - \mu_1, \theta_2 < c, \ldots, \theta_n < c) \\ &= \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{{}^{(c-\mu_1)}/_{\sigma_1} + \lambda_1 x}{\sqrt{1-\lambda_1^2}} \right) \prod_{j=2}^{n} \left[\Phi\left(\frac{c + \lambda_j}{\sqrt{1-\lambda_j^2}} \right) \right] dx \\ &= \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{\sqrt{\rho_0}(c-\mu_1)}{\sigma_1 \sqrt{\rho_0 - \rho_1^2}} + \frac{\rho_1}{\sqrt{\rho_0 - \rho_1^2}} x \right) \left[\Phi\left(\frac{c}{\sqrt{1-\rho_0}} + \sqrt{\frac{\rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx\end{aligned} \tag{71}$$

The third line in (71) follows from Equation (8.2.13) in Tong (1990, p. 193), provided that $|\rho_1| \leq \sqrt{\rho_0}$. As discussed in the "Exchangeability" section of Appendix B regarding the non-negativity of ρ_{FC}, the condition $|\rho_1| \leq \sqrt{\rho_0}$ is satisfied under the exchangeability assumption.

For TP arrays, expression (71) specifies the Rejection rate:

$$\begin{aligned}\text{RR}_{\text{TP}} = \int_{-\infty}^{\infty} \varphi(x) \Phi\left(\frac{\sqrt{\rho_0}(c-\mu_T)}{\sigma \sqrt{\rho_0 - \rho_1^2}} + \frac{\rho_1}{\sqrt{\rho_0 - \rho_1^2}} x \right) \times \\ \left[\Phi\left(\frac{c}{\sqrt{1-\rho_0}} + \sqrt{\frac{\rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx.\end{aligned} \tag{72}$$

For TA arrays the rejection rate becomes

$$\text{RR}_{\text{TA}} = \int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{c}{\sqrt{1-\rho_0}} + \sqrt{\frac{\rho_0}{1-\rho_0}} x \right) \right]^{n} dx. \tag{73}$$

With respect to the calculations for the probabilities of choosing the (arbitrarily) first stimulus, we note that

$$\mathbf{D}_{n\times n} \vec{\theta}_n$$

has a multivariate normal distribution with mean and covariance matrix:

$$\vec{\mu}_{\text{DC}} = \mathbf{D}_{n\times n} \vec{\mu}_n = \mu_1 \vec{1}_n, \tag{74}$$

and

$$\begin{aligned}\boldsymbol{\Sigma}^{\mathrm{DC}}_{n\times n} &= \mathbf{D}_{n\times n}\boldsymbol{\Sigma}^{\mathrm{EW}}_{n\times n}\mathbf{D}^{\intercal}_{n\times n} \\ &= \begin{pmatrix} \vec{e}^{\intercal}_{1,n} \\ \cdots\cdots\cdots \\ \mathbf{M}_{(n-1)\times n} \end{pmatrix} \boldsymbol{\Sigma}^{\mathrm{EW}}_{n\times n} \begin{pmatrix} \vec{e}_{1,n} \vdots \mathbf{M}^{\intercal}_{(n-1)\times n} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1^2 & \vdots & (\sigma_1^2 - \sigma_1\rho_1)\vec{1}^{\intercal}_{n-1} \\ \cdots\cdots\cdots & \vdots & \cdots\cdots\cdots\cdots\cdots \\ (\sigma_1^2 - \sigma_1\rho_1)\vec{1}_{n-1} & \vdots & \boldsymbol{\Sigma}^{\mathrm{FC}}_{(n-1)\times(n-1)} \end{pmatrix}. \end{aligned} \tag{75}$$

We note further that the covariances in the first row and column of matrix $\boldsymbol{\Sigma}^{\mathrm{DC}}_{n\times n}$ can be written

$$\sigma_1^2 - \sigma_1\rho_1 = \rho_{DC}\sigma_1\sqrt{\sigma_1^2 - 2\sigma_1\rho_1 + 1},$$

where ρ_{DC} is the corresponding correlation. Thus, we see that the correlations in the first row and column of the correlation matrix implied by $\boldsymbol{\Sigma}^{\mathrm{DC}}_{n\times n}$ are:

$$\rho_{\mathrm{DC}} = \frac{\sigma_1 - \rho_1}{\sqrt{\sigma_1^2 - 2\sigma_1\rho_1 + 1}}. \tag{76}$$

The other correlations in the matrix equal ρ_{FC} as defined in expression (63). From these two expressions, we can express the correlations as $\rho_{\mathrm{DC}} = \lambda_{\mathrm{DC}}\lambda_{\mathrm{FC}}$ and $\rho_{\mathrm{FC}} = \lambda^2_{\mathrm{FC}}$ with

$$\lambda_{\mathrm{DC}} = \frac{\sigma_1 - \rho_1}{\sqrt{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}}$$

and $\lambda_{\mathrm{FC}} = \sqrt{\rho_{\mathrm{FC}}}$. Thus, we can use Equation (8.2.13) from Tong (1990), as in the computation of the Rejection rate in expression (71), to obtain the generic probability of choosing the first stimulus in a lineup under the Dependent Observations model. Its

expression is

$$\begin{aligned}
\Pr(\text{Selection}) &= \Pr(\theta_1 > c, \theta_1 - \theta_2 > 0, \ldots, \theta_1 - \theta_n > 0) \\
&= \Pr(-\theta_1 < -c, \theta_2 - \theta_1 < 0, \ldots, \theta_n - \theta_1 < 0) \\
&= \Pr(\mu_1 - \theta_1 < \mu_1 - c, \theta_2 - \theta_1 + \mu_1 < \mu_1, \ldots, \theta_n - \theta_1 + \mu_1 < \mu_1) \\
&= \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{{}^{(\mu_1 - c)}/_{\sigma_1} + \lambda_{\text{DC}} x}{\sqrt{1 - \lambda_{\text{DC}}^2}} \right) \left[\Phi\left(\frac{\mu_1}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx \\
&= \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{\sqrt{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}(\mu_1 - c)}{\sigma_1\sqrt{\rho_0 - \rho_1^2}} + \frac{\sigma_1 - \rho_1}{\sqrt{\rho_0 - \rho_1^2}} x \right) \times \\
&\qquad \left[\Phi\left(\frac{\mu_1}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma_1^2 - 2\sigma_1\rho_1 + \rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx
\end{aligned} \tag{77}$$

Expression (77) particularizes into the HR of the Dependent Observations model for TP arrays as

$$\begin{aligned}
\text{HR} = \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{\sqrt{\sigma^2 - 2\sigma\rho_1 + \rho_0}(\mu_T - c)}{\sigma\sqrt{\rho_0 - \rho_1^2}} + \frac{\sigma - \rho_1}{\sqrt{\rho_0 - \rho_1^2}} x \right) \times \\
\left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma^2 - 2\sigma\rho_1 + \rho_0}{1-\rho_0}} x \right) \right]^{n-1} dx
\end{aligned}, \tag{78}$$

and the FAR for TA arrays as:

$$\text{FAR} = \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{-c}{\sqrt{\rho_0}} + \sqrt{\frac{1-\rho_0}{\rho_0}} x \right) [\Phi(x)]^{n-1}\, dx \tag{79}$$

Restricted cases

If we assume that the same shared variance is shared among all evidence, we may replace ρ_1 by ${}^{\rho_0}/_{\sigma}$ in the expressions above. If, in addition, we assume equal variances, we can further simplify the expressions by replacing σ with 1.

For the case with the same shared variance, the Rejection rate for TP arrays and

the HR reduce to:

$$\begin{aligned} \mathrm{RR}_{\mathrm{TP}} = \int_{-\infty}^{\infty} \varphi(x)\Phi\left(\frac{c-\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\rho_0}{1-\rho_0}}x\right) \times \\ \left[\Phi\left(\frac{c}{\sqrt{1-\rho_0}} + \sqrt{\frac{\rho_0}{1-\rho_0}}x\right)\right]^{n-1} dx, \end{aligned} \tag{80}$$

$$\begin{aligned} \mathrm{HR} &= \int_{-\infty}^{\infty} \varphi(x)\,\Phi\left(\frac{\sqrt{\sigma^2-\rho_0}(\mu_T-c)}{\sqrt{\rho_0(1-\rho_0)}} + \frac{\sigma - {}^{\rho_0}/_{\sigma}}{\sqrt{\rho_0(1-{}^{\rho_0}/_{\sigma^2})}}x\right) \times \\ &\qquad \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma^2-\rho_0}{1-\rho_0}}x\right)\right]^{n-1} dx \\ &= \int_{-\infty}^{\infty} \varphi(x)\,\Phi\left(\frac{\sqrt{\sigma^2-\rho_0}(\mu_T-c)}{\sqrt{\rho_0(1-\rho_0)}} + \frac{\sigma^2-\rho_0}{\sqrt{\rho_0(1-\rho_0)}}x\right) \times \\ &\qquad \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + \sqrt{\frac{\sigma^2-\rho_0}{1-\rho_0}}x\right)\right]^{n-1} dx. \end{aligned} \tag{81}$$

Only the HR simplifies yet further when we additionally assume equal variances for both populations. In this case, the HR becomes:

$$\begin{aligned} \mathrm{HR} = \int_{-\infty}^{\infty} \varphi(x)\,\Phi\left(\frac{\mu_T-c}{\sqrt{\rho_0}} + \sqrt{\frac{1-\rho_0}{\rho_0}}x\right) \times \\ \left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}} + x\right)\right]^{n-1} dx. \end{aligned} \tag{82}$$

The derivations for the independent observations model (in which the covariances across items are assumed to be 0) are presented in the body of the paper.

Appendix D

Integration model

The rejection region for the Integration model is defined by the projection on a unidimensional subspace. It represents the sum of all evidence, which is straightforward to present in matrix notation as:

$$\vec{1}_n^{\intercal}\vec{\theta}_n < c. \tag{30'}$$

The region where the (arbitrarily) first stimulus is chosen is given by

$$\begin{pmatrix} 1 & \vdots & \vec{1}_{n-1}^{\intercal} \\ \cdots & \vdots & \cdots \\ \vec{1}_{n-1} & \vdots & -\mathbf{I}_{(n-1)\times(n-1)} \end{pmatrix} \vec{\theta}_n > \begin{pmatrix} c \\ \cdots \\ \vec{0}_{n-1} \end{pmatrix}. \tag{31: revisited}$$

The matrix that defines the region can be written as

$$\mathbf{N}_{n\times n} = \begin{pmatrix} \vec{1}_n^{\intercal} \\ \cdots \\ \mathbf{M}_{(n-1)\times n} \end{pmatrix}. \tag{83}$$

Exchangeability

The rejection rates are very simple to compute because the sum of normally distributed random variables is (univariate) normally distributed. The mean of their sum is the sum of the means; hence, it is 0 for TA arrays, and μ_T for TP arrays. The variance for TA arrays is

$$n\left[1 + (n-1)\rho_0\right],$$

and for TP arrays is

$$\sigma^2_{\text{IN}} = \sigma^2 + (n-1)[1 + 2\sigma\rho_1 + (n-2)\rho_0] \tag{84}$$

Therefore, the expressions for the RR are given by

$$\text{RR}_{\text{TP}} = \Phi\left(\frac{c - \mu_T}{\sqrt{\sigma^2 + (n-1)[1 + 2\sigma\rho_1 + (n-2)\rho_0]}}\right) \tag{32: revisited}$$

$$\text{RR}_{\text{TA}} = \Phi\left(\frac{c}{\sqrt{n\left[1 + (n-1)\rho_0\right]}}\right) \tag{33: revisited}$$

For the probability of selecting the (arbitrarily) first stimulus, the mean and covariance matrix of $\mathbf{N}_{n\times n}\vec{\theta}_n$ are given by:

$$\vec{\mu}_{\text{IC}} = \mathbf{N}_{n\times n}\vec{\mu}_n = \mu_1\vec{1}_n, \tag{85}$$

and

$$\begin{aligned}\mathbf{\Sigma}^{\text{IC}}_{n\times n} &= \mathbf{N}_{n\times n}\mathbf{\Sigma}^{\text{EW}}_{n\times n}\mathbf{N}^{\intercal}_{n\times n} \\ &= \left(\begin{array}{c|c} \sigma_1^2 + (n-1)[1 + 2\sigma_1\rho_1 + (n-2)\rho_0] & [\sigma_1^2 - (n-2)(\rho_0 - \sigma_1\rho_1) - 1]\vec{1}^{\intercal}_{n-1} \\ \hline [\sigma_1^2 - (n-2)(\rho_0 - \sigma_1\rho_1) - 1]\vec{1}_{n-1} & \mathbf{\Sigma}^{\text{FC}}_{(n-1)\times(n-1)} \end{array}\right).\end{aligned} \tag{86}$$

As was the case for the Dependent Observations model, it is possible to decompose the correlations implied by matrix $\mathbf{\Sigma}^{\text{IN}}_{n\times n}$ into λ_{IN} and λ_{FC}, satisfying the requirements of Equation (8.2.13) of Tong (1990). To facilitate the transparency of this derivation, we shall use the following additional notation:

$$\begin{aligned}\sigma^2_{\text{IC}} &= \sigma_1^2 + (n-1)[1 + 2\sigma_1\rho_1 + (n-2)\rho_0], \\ a_{\text{IC}} &= \sigma_1^2 - (n-2)(\rho_0 - \sigma_1\rho_1) - 1, \\ a_{\text{FC}} &= \sigma_1^2 - 2\sigma_1\rho_1 + \rho_0.\end{aligned}$$

Now, we can compactly express $\lambda_{\text{FC}} = \sqrt{\rho_{\text{FC}}} = {}^{\sqrt{a_{\text{FC}}}}/_{\sigma_{\text{FC}}}$ and $\lambda_{\text{IC}} = {}^{a_{\text{IC}}}/_{\sigma_{\text{IC}}\sqrt{a_{\text{FC}}}}$. To apply Equation (8.2.13), we need to verify that $\lambda^2_{\text{IC}} \leq 1$. Note, however, that λ_{IC} depends on n. Since the condition on λ_{IC} needs to be satisfied by any n, we examine its limiting behavior:

$$\lim_{n\to\infty}\lambda_{\text{IC}} = \frac{(\rho_0 - \sigma_1\rho_1)^2}{\rho_0 a_{\text{FC}}},$$

which can be shown, after little algebraic manipulation, to be less than 1 if and only if $\rho_1^2 \leq \rho_0$. Therefore, the generic expression for the probability of selecting the first stimulus

in a lineup according to the Integration model is:

$$\begin{aligned}
\Pr(\text{Selection}) &= \Pr(\sum_{j=1}^{n}\theta_j > c, \theta_1-\theta_2>0,\ldots,\theta_1-\theta_n>0)\\
&= \Pr(-\sum_{j=1}^{n}\theta_j < -c, \theta_2-\theta_1<0,\ldots,\theta_n-\theta_1<0)\\
&= \Pr(\mu_1-\sum_{j=1}^{n}\theta_j < \mu_1-c, \theta_2-\theta_1+\mu_1<\mu_1,\ldots,\theta_n-\theta_1+\mu_1<\mu_1)\\
&= \int_{-\infty}^{\infty}\varphi(x)\,\Phi\left(\frac{{}^{(\mu_1-c)}/_{\sigma_{IC}}+\lambda_{\text{IC}}x}{\sqrt{1-\lambda_{\text{IC}}^2}}\right)\left[\Phi\left(\frac{\mu_1}{\sqrt{1-\rho_0}}+\sqrt{\frac{\sigma_1^2-2\sigma_1\rho_1+\rho_0}{1-\rho_0}}x\right)\right]^{n-1}dx\\
&= \int_{-\infty}^{\infty}\varphi(x)\,\Phi\left(\sqrt{\frac{a_{\text{FC}}}{\sigma_{\text{IC}}^2a_{\text{FC}}-a_{\text{IC}}^2}}(\mu_1-c)+\sqrt{\frac{a_{\text{IC}}^2}{\sigma_{\text{IC}}^2a_{\text{FC}}-a_{\text{IC}}^2}}x\right)\times\\
&\qquad\left[\Phi\left(\frac{\mu_1}{\sqrt{1-\rho_0}}+\sqrt{\frac{a_{\text{FC}}}{1-\rho_0}}x\right)\right]^{n-1}dx.
\end{aligned} \tag{87}$$

Expression (87) particularizes the HR for the Integration model as

$$\begin{aligned}
\text{HR} = &\int_{-\infty}^{\infty}\varphi(x)\times\\
&\Phi\left(\frac{\sqrt{\sigma^2-2\sigma\rho_1+\rho_0}(\mu_1-c)+(\sigma_1^2-(n-2)(\rho_0-\sigma_1\rho_1)-1)x}{\sqrt{(\sigma_1^2+(n-1)[1+2\sigma_1\rho_1+(n-2)\rho_0])(\sigma^2-2\sigma\rho_1+\rho_0)-(\sigma_1^2-(n-2)(\rho_0-\sigma_1\rho_1)-1)^2}}\right)\times\\
&\left[\Phi\left(\frac{\mu_T}{\sqrt{1-\rho_0}}+\sqrt{\frac{\sigma^2-2\sigma\rho_1+\rho_0}{1-\rho_0}}x\right)\right]^{n-1}dx,
\end{aligned} \tag{88}$$

and the FAR as:

$$\begin{aligned}
\text{FAR} &= \int_{-\infty}^{\infty}\varphi(x)\Phi\left(\frac{-c}{\sqrt{n\,[1+(n-1)\rho_0]}}\right)[\Phi(x)]^{n-1}\,dx\\
&= \frac{1}{n}\Phi\left(\frac{-c}{\sqrt{n\,[1+(n-1)\rho_0]}}\right).
\end{aligned} \tag{35'}$$

Restricted cases

For the equal shared variance across populations, the HR simplifies to:

$$\begin{aligned}\mathrm{HR} = \int_{-\infty}^{\infty} \varphi(x)\, \Phi\left(\frac{\sqrt{\sigma^2 - \rho_0}(\mu_1 - c) + (\sigma^2 - 1)x}{\sqrt{(\sigma^2 - \rho_0)[\sigma^2 + (n-1)(1 + n\rho_0)] - (\sigma^2 - 1)^2}} \right) \times \\ \left[\Phi\left(\frac{\mu_T}{\sqrt{1 - \rho_0}} + \sqrt{\frac{\sigma^2 - \rho_0}{1 - \rho_0}} x \right) \right]^{n-1} dx.\end{aligned} \tag{89}$$

For equal variances and covariances, the HR of the Integration model is:

$$\begin{aligned}\mathrm{HR} = &\Phi\left(\frac{\mu_T - c}{\sqrt{n\,[1 + (n-1)\rho_0]}} \right) \times \\ &\int_{-\infty}^{\infty} \varphi(x) \left[\Phi\left(\frac{\mu_T}{\sqrt{1 - \rho_0}} - x \right) \right]^{n-1} dx,\end{aligned} \tag{32: revisited}$$

Appendix E

Ensemble model

According to the Ensemble model, the region where the lineup is rejected can be written as

$$n\left[\mathbf{I}_{n\times n}-\frac{1}{n}\vec{1}_n\vec{1}_n^{\intercal}\right]\vec{\theta}_n \leq nc\vec{1}_n. \qquad \text{(38: revisited)}$$

We shall see that the matrix that defines the rejection region is particularly important in computing all probabilities of this model. For $m \geq 1$, we define

$$\mathbf{K}_{m\times m}=\left[\mathbf{I}_{m\times m}-\frac{1}{m}\vec{1}_m\vec{1}_m^{\intercal}\right]. \qquad (90)$$

Thus, in particular, the rejection region will be written as $n\mathbf{K}_{n\times n}\vec{\theta}_n \leq nc\vec{1}_n$

The region where the first stimulus is selected is

$$\left(\begin{array}{c|c} n-1 & -\vec{1}_{n-1}^{\intercal} \\ \hline \vec{1}_{n-1} & -\mathbf{I}_{(n-1)\times(n-1)} \end{array}\right)\vec{\theta}_n > \left(\begin{array}{c} nc \\ \hline \vec{0}_{n-1} \end{array}\right). \qquad \text{(39: revisited)}$$

We shall denote the matrix that defines the region for selecting the first stimulus by

$$\mathbf{E}_{n\times n}=\left(\begin{array}{c|c} n-1 & -\vec{1}_{n-1}^{\intercal} \\ \hline \multicolumn{2}{c}{\mathbf{M}_{(n-1)\times n}} \end{array}\right)$$

Exchangeability

The linear combination of the left-hand side of expression (38) has a multivariate distribution with mean vector and covariance matrix:

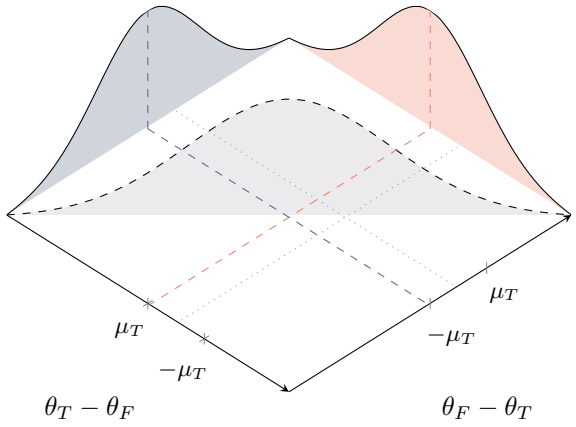
$$\mu_1\left(\begin{array}{c} n-1 \\ \hline -\vec{1}_{n-1} \end{array}\right) \qquad (91)$$

$$\begin{aligned}\mathbf{\Sigma}_{n\times n}^{\mathrm{NR}} &= n^2\mathbf{K}_{n\times n}\mathbf{\Sigma}_{n\times n}^{\mathrm{EW}}\mathbf{K}_{n\times n}^{\intercal} \\ &= \left(\begin{array}{c|c} (n-1)[(n-1)a_{\mathrm{FC}}+(1-\rho_0)] & -[(n-1)a_{\mathrm{FC}}+(1-\rho_0)]\vec{1}_{n-1}^{\intercal} \\ \hline -[(n-1)a_{\mathrm{FC}}+(1-\rho_0)]\vec{1}_{n-1} & n^2(1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)}-[(n+1)(1-\rho_0)-a_{\mathrm{FC}}]\vec{1}_{n-1}\vec{1}_{n-1}^{\intercal} \end{array}\right).\end{aligned} \qquad (92)$$

In this case, neither Equation (8.3.1) nor (8.2.13) from Tong (1990) can be applied because all covariances (and hence all correlations) are non-positive. Equation (8.3.1) requires a common non-negative correlation and Equation (8.2.13) requires a decomposition that implies that some correlations must be non-negative when $n \geq 3$. Moreover, note that the first row of $\mathbf{\Sigma}^{\text{EW}}_{n\times n}$ minus the sum of all other rows equals zero. Therefore, this covariance matrix is singular, and we will need to compute the relevant probabilities in a suitable subspace. Figure E1 illustrates the subspace for the case of one target and one filler together with the density over that subspace.

Figure E1

Illustration of the joint distribution of $2\mathbf{K}_{2\times 2}\vec{\theta}_2$



Note. Joint distribution of the transformation of two internal responses for the rejection of the lineup in the Ensemble model where one stimulus from the noise population (θ_F), and one stimulus from the signal population (θ_T) are presented. Dashed lines that run parallel to the axes indicate the locations of the means of $\theta_T - \theta_F$, at μ_T, and of $\theta_F - \theta_T$, at $-\mu_T$. The dashed lines show the location of 0 which indicates the region where one response is larger than the other. The joint density clearly shows the loss of dimensionality of the transformation by $2\mathbf{K}_{2\times 2}$.

In general, the appropriate subspace of $n\mathbf{K}_{n\times n}$ is easier to identify by using the law

of total probability for continuous variables to reexpress the probability that the evidence falls in the rejection area defined in expression (38). We will denote

$$\sigma_{\mathrm{E1}} = (n-1)[(n-1)a_{\mathrm{FC}} + (1-\rho_0)] \tag{93}$$

and

$$\mathbf{K}^{-}_{(n-1)\times n} = \left(\vec{0}_{n-1} \vdots \mathbf{I}_{(n-1)\times(n-1)}\right)\mathbf{K}_{n\times n},$$

to write the probability of rejection as:

$$\begin{aligned}\Pr(n\mathbf{K}_{n\times n}\vec{\theta}_n \leq nc\vec{1}_n) = \int_{-\infty}^{nc} \frac{1}{\sigma_{\mathrm{E1}}}\varphi\left(\frac{x-(n-1)\mu_1)}{\sigma_{\mathrm{E1}}}\right) \times \\ \Pr\left(n\mathbf{K}^{-}_{(n-1)\times n}\vec{\theta}_n \leq nc\vec{1}_{n-1}\Big| n\theta_1 - \sum_{k=1}^{n}\theta_j = x\right)dx.\end{aligned} \tag{94}$$

That is, we condition on the value of the first evidence variable and based on the law of total probability recover the complete joint probability by integrating over its values within the rejection area.

Next, note that $n\mathbf{K}^{-}_{(n-1)\times n}\vec{\theta}_n$ given $n\theta_1 - \sum_{k=1}^{n}\theta_j = x$ has a multivariate normal distribution with mean

$$-\frac{x}{n-1}\vec{1}_{n-1}$$

and with covariance matrix

$$n^2(1-\rho_0)\mathbf{K}_{(n-1)\times(n-1)}.$$

We use the generic matrix $\mathbf{K}_{m\times m}$ to identify the dimension of the subspace and to find expressions for the probability of rejection in the Ensemble model. Later, we will also use it to compute the HR and FAR for this model. Note that the trace of $\mathbf{K}_{m\times m}$ is $m-1$ and that the matrix is idempotent—that is, $\mathbf{K}_{m\times m}\mathbf{K}_{m\times m} = \mathbf{K}_{m\times m}$. Since the matrix is idempotent, all of its eigenvalues are either 0 or 1, and since its trace is $m-1$, precisely $m-1$ of the eigenvalues equal 1. Consequently, the rank of $\mathbf{K}_{m\times m}$ is $m-1$ and a vector of variables whose covariance matrix is a (scalar) multiple of $\mathbf{K}_{m\times m}$ is completely described in a subspace with dimension $m-1$. For a multivariate normal distribution whose mean is μ_m

and covariance matrix is $k^2\mathbf{K}_{m\times m}$, for some constant k, the distribution can be written as

$$k\mathbf{C}_{m\times(m-1)}\vec{Z}_{m-1} + \vec{\mu}_m,$$

where $\vec{Z}_{m-1}$ is a vector of $m-1$ independent standard normally distributed random variables, and $\mathbf{C}_{m\times(m-1)}$ is some matrix of rank $m-1$ such that $\mathbf{C}_{m\times(m-1)}\mathbf{C}^{\intercal}_{m\times(m-1)} = \mathbf{K}_{m\times m}$. For $m = 1$, $\mathbf{K}_{1\times 1} = 0$ and it is not necessary to find any matrix to decompose it into, and $\mathbf{C}_{1\times 0}$ can be left undefined. The Cholesky decomposition of $\mathbf{K}_{m\times m}$ allows us to find one such matrix by taking the nonzero columns of the decomposition; that is

$$\mathbf{L} = \left(\mathbf{C}_{m\times(m-1)} \vdots \vec{0}_m\right)$$

is lower triangular and $\mathbf{L}\mathbf{L}^{\intercal} = \mathbf{C}_{m\times(m-1)}\mathbf{C}^{\intercal}_{(m-1)\times m} = \mathbf{K}_{m\times m}$. Now, the Cholesky algorithm (in which one column is computed at each step) shows that the first column of the decomposition is

$$\begin{pmatrix} \sqrt{\frac{m-1}{m}} \\ \cdots\cdots\cdots\cdots \\ -\sqrt{\frac{1}{m(m-1)}}\vec{1}_{m-1} \end{pmatrix}$$

and that the input matrix for the second step of the algorithm[8] is

$$\begin{pmatrix} 1 & \vdots & \vec{0}^{\intercal}_{m-1} \\ \cdots & \vdots & \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \vec{0}_{m-1} & \vdots & \mathbf{I}_{(m-1)\times(m-1)} - \frac{1}{m-1}\vec{1}_{m-1}\vec{1}^{\intercal}_{m-1} \end{pmatrix}$$

Thus, we see that the matrix $\mathbf{C}_{m\times(m-1)}$ can be written as:

$$\mathbf{C}_{m\times(m-1)} = \begin{pmatrix} \sqrt{\frac{m-1}{m}} & \vdots & \vec{0}^{\intercal}_{m-2} \\ \cdots\cdots\cdots\cdots & \vdots & \cdots\cdots\cdots\cdots \\ -\sqrt{\frac{1}{m(m-1)}}\vec{1}_{m-1} & \vdots & \mathbf{C}_{(m-1)\times(m-2)} \end{pmatrix} \tag{95}$$

[8] It may be noted that the algorithm that produces the matrix in the lower right block is the same as the one used to compute the covariance matrix of a vector of random variables conditioned of the first component.

and, in extenso, it is

$$\begin{pmatrix} \sqrt{\frac{m-1}{m}} & 0 & \cdots & \cdots & \cdots & 0 \\ -\sqrt{\frac{1}{m(m-1)}} & \ddots & & & & \vdots \\ \vdots & \ddots & \sqrt{\frac{m-k}{m-k+1}} & & & \\ & & -\sqrt{\frac{1}{(m-k)(m-k+1)}} & \ddots & & 0 \\ & & \vdots & & \ddots & \sqrt{\frac{1}{2}} \\ -\sqrt{\frac{1}{m(m-1)}} & & -\sqrt{\frac{1}{(m-k)(m-k+1)}} & & & -\sqrt{\frac{1}{2}} \end{pmatrix} \tag{96}$$

Therefore, we can construct $\mathbf{C}_{(n-1)\times(n-2)}$ as in expression (96), taking $m = n-1$, such that

$$n\sqrt{1-\rho_0}\,\mathbf{C}_{(n-1)\times(n-2)}\vec{Z}_{n-1} - \frac{x}{n-1}\vec{1}_{n-1}, \tag{97}$$

where $\vec{Z}_{n-1}$ are independently distributed standard normal variables, have the same distribution as $n\mathbf{K}^{-}_{(n-1)\times n}\vec{\theta}_n$ conditioned on $n\theta_1 - \sum_{k=1}^{n}\theta_j = x$. Based on expression (97), we can compute

$$\begin{aligned} \Pr\left(n\mathbf{K}^{-}_{(n-1)\times n}\vec{\theta}_n \le nc\vec{1}_{n-1} \middle| n\theta_1 - \sum_{k=1}^{n}\theta_j = x \right) = \\ \int_{a_1}^{b_1}\cdots\int_{a_k}^{b_k}\cdots\int_{a_{n-1}}^{b_{n-1}} \prod_{j=1}^{n-2}\varphi(z_j)\,d\vec{z}_{n-2}, \end{aligned} \tag{98}$$

where

$$\begin{aligned} b_1 &= \sqrt{\frac{n-1}{n-2}}\left[\frac{c}{\sqrt{1-\rho_0}} + \frac{x}{n(n-1)\sqrt{1-\rho_0}}\right] \\ b_k &= \sqrt{\frac{n-k}{n-k-1}}\left[\frac{c}{\sqrt{1-\rho_0}} + \frac{x}{n(n-1)\sqrt{1-\rho_0}} + \sum_{j=1}^{k-1}\sqrt{\frac{1}{(n-j)(n-j-1)}}z_j\right], \\ & k = 2,\ldots,(n-2) \end{aligned} \tag{99}$$

and

$$a_k = -(n-k-1)b_k, \qquad k = 1,\ldots,(n-2). \tag{100}$$

By replacing expression (98) into (94), and and by a change of variables, we obtain

that the generic expression for the Rejection rates is:

$$\Pr(\text{Reject}) = \int_{a_0}^{b_0}\int_{a_1}^{b_1}\cdots\int_{a_k}^{b_k}\cdots\int_{a_{n-2}}^{b_{n-2}} \varphi(x)\prod_{j=1}^{n-2}\varphi(z_j)\,d\vec{z}_{n-2}dx, \tag{101}$$

where the limits of integration are:

$$\begin{aligned}
b_0 &= \frac{nc-(n-1)\mu_1}{\sigma_{\text{E1}}}\\
b_1 &= \sqrt{\frac{n-1}{n-2}}\left[\frac{nc+\mu_1}{n\sqrt{1-\rho_0}}+\frac{\sigma_{\text{E1}}x}{n(n-1)\sqrt{1-\rho_0}}\right]\\
b_k &= \sqrt{\frac{n-k}{n-k-1}}\left[\frac{nc+\mu_1}{n\sqrt{1-\rho_0}}+\frac{\sigma_{\text{E1}}x}{n(n-1)\sqrt{1-\rho_0}}+\sum_{j=1}^{k-1}\sqrt{\frac{1}{(n-j)(n-j-1)}}z_j\right],\\
&\quad k=2,\ldots,(n-2),
\end{aligned} \tag{102}$$

and

$$\begin{aligned}
a_0 &= -(n-1)\frac{nc+\mu_1}{\sigma_{\text{E1}}}\\
a_k &= -(n-k-1)b_k, \quad k=1,\ldots,(n-2).
\end{aligned} \tag{103}$$

To compute RR_{TP}, σ^2_{E1} becomes $(n-1)[(n-1)(\sigma^2-2\sigma\rho_1+\rho_0)+(1-\rho_0)]$ for TP arrays, and to find RR_{TA}, it simplifies to $n(n-1)(1-\rho_0)$.

To derive the Hit and the False Alarm Rates, we proceed similarly. First, we note that $\mathbf{E}_{n\times n}\vec{\theta}_n$ has a multivariate normal distribution with mean vector (104) and covariance matrix (105) given by

$$\mu_1\begin{pmatrix} n-1 \\ \cdots\cdots \\ \vec{1}_{n-1}\end{pmatrix} \tag{104}$$

$$\mathbf{\Sigma}^{NC}_{n\times n} = \left(\begin{array}{c:c} \sigma^2_{\text{E1}} & \frac{\sigma^2_{\text{E1}}}{n-1}\vec{1}^{\intercal}_{n-1} \\ \hdashline \frac{\sigma^2_{\text{E1}}}{n-1}\vec{1}_{n-1} & (1-\rho_0)\mathbf{I}_{(n-1)\times(n-1)}+a_{\text{FC}}\vec{1}_{n-1}\vec{1}^{\intercal}_{n-1}\end{array}\right) \tag{105}$$

Next, we see that $\mathbf{M}_{(n-1)\times n}\vec{\theta}_n$ conditioned on $n\theta_1-\sum_{j=1}^{n}\theta_j = x$ has a multivariate distribution with mean

$$\frac{x}{n-1}\vec{1}_{n-1}$$

and covariance matrix

$$(1-\rho_0)\mathbf{K}_{(n-1)\times(n-1)}.$$

From the computations for the Rejection Rate above, it is apparent that $\mathbf{M}_{(n-1)\times n}\vec{\theta}_n$ conditioned on $n\theta_1 - \sum_{j=1}^{n}\theta_j = x$ can be expressed as

$$\sqrt{1-\rho}\,\mathbf{C}_{(n-1)\times(n-2)}\vec{Z}_{n-2} + \frac{x}{n-1}\vec{1}_{n-1} \tag{106}$$

where $\vec{Z}_{n-2}$ are independently distributed standard normal variables. Similarly to the Rejection Rate, we arrive at expression

$$\Pr\left(\mathbf{M}_{(n-1)\times n}\vec{\theta}_n > \vec{0}_{n-1}\middle| n\theta_1 - \sum_{j=1}^{n}\theta_j = x\right) = \int_{a_1}^{b_1}\cdots\int_{a_{n-2}}^{b_{n-2}}\prod_{j=1}^{n-2}\varphi(z_j)\,d\vec{z}_{n-2} \tag{107}$$

where

$$\begin{aligned} a_1 &= -\frac{x}{\sqrt{(n-1)(n-2)(1-\rho_0)}} \\ a_k &= -\sqrt{\frac{n-k}{n-k-1}}\left[\frac{x}{(n-1)\sqrt{1-\rho_0}} - \sum_{j=1}^{k-1}\sqrt{\frac{1}{(n-j)(n-j-1)}}z_j\right], \\ & \quad k = 2,\ldots,(n-2) \end{aligned} \tag{108}$$

and

$$\begin{aligned} b_1 &= (n-2)\frac{x}{\sqrt{(n-1)(n-2)(1-\rho_0)}} \\ b_k &= -(n-k-1)a_k, \qquad k = 1,\ldots,(n-2). \end{aligned} \tag{109}$$

Thus, after a change of variables, the Hit rate for the Ensemble model is computed as:

$$\mathrm{HR} = \int_{a_0}^{\infty}\int_{a_1}^{b_1}\cdots\int_{a_{n-2}}^{b_{n-2}}\varphi(x)\prod_{j=1}^{n-2}\varphi(z_j)\,d\vec{z}_{n-2}dx, \tag{110}$$

where

$$\begin{aligned} a_0 &= \frac{nc-(n-1)\mu_T}{\sigma_{E1}} \\ a_1 &= -\frac{\sigma_{\mathrm{E1}}x + (n-1)\mu_T}{\sqrt{(n-1)(n-2)(1-\rho_0)}} \\ a_k &= -\sqrt{\frac{n-k}{n-k-1}}\left[\frac{\sigma_{\mathrm{E1}}x}{(n-1)\sqrt{1-\rho_0}} + \frac{\mu_T}{\sqrt{1-\rho_0}} - \sum_{j=1}^{k-1}\sqrt{\frac{1}{(n-j)(n-j-1)}}z_j\right], \\ & \quad k = 2,\ldots,(n-2) \end{aligned} \tag{111}$$

and

$$b_k = -(n-k-1)a_k, \qquad k = 1,\ldots,(n-2), \tag{112}$$

with σ^2_{E1} equal to $(n-1)[(n-1)(\sigma^2 - 2\sigma\rho_1 + \rho_0) + (1-\rho_0)]$ since HR is computed for TP arrays. The False alarm rate is similarly computed, but with simplified limits of integration:

$$\begin{aligned} a_0 &= \frac{nc}{\sqrt{n(n-1)(1-\rho_0)}} \\ a_1 &= -\sqrt{\frac{n}{n-2}}x \\ a_k &= -\sqrt{\frac{n-k}{n-k-1}}\left[\sqrt{\frac{n}{n-1}}x - \sum_{j=1}^{k-1}\sqrt{\frac{1}{(n-j)(n-j-1)}}z_j\right], \\ & \quad k = 2, \ldots, (n-2) \end{aligned} \tag{113}$$

and

$$b_k = -(n-k-1)a_k, \qquad k = 1, \ldots, (n-2). \tag{114}$$

Restricted cases

For the restricted cases, the expressions for the Ensemble model change only the form that the term σ_{E1} takes for the integration limits. For the case of unequal variance with equal covariance, σ^2_{E1} reduces to $(n-1)[(n-1)(\sigma^2 - \rho_0) + (1-\rho_0)$, and for equal variances it further simplifies to $n(n-1)(1-\rho_0)$.