

Indecision on decisional separability

Michael L. Mack · Jennifer J. Richler ·
Isabel Gauthier · Thomas J. Palmeri

© Psychonomic Society, Inc. 2010

Abstract The theoretical framework of General Recognition Theory (GRT; Ashby & Townsend, *Psychological Review*, 93, 154–179, 1986) coupled with the empirical analysis tools of Multidimensional Signal Detection Analysis (MSDA; Kadlec & Townsend, *Multidimensional models of perception and recognition*, pp. 181–228, 1992) have become one important method for assessing dimensional interactions in perceptual decision-making. In this article, we critically examine MSDA and characterize cases where it is unable to discriminate two kinds of dimensional interactions: perceptual separability and decisional separability. We performed simulations with known instances of violations of perceptual or decisional separability, applied MSDA to the data generated by these simulations, and evaluated MSDA on its ability to accurately characterize the perceptual versus decisional source of these simulated dimensional interactions. Critical cases of violations of perceptual separability are often mischaracterized by MSDA as violations of decisional separability.

Keywords GRT · MSDA · Decisional · Perceptual

How are dimensions of a stimulus combined and used to make a perceptual decision? Are dimensions processed independently or do they interact, and if so, how? This fundamental question has been asked for a broad range of domains, including simple perceptual stimuli (e.g., Shepard, 1964), faces (e.g., Richler et al., 2008; Thomas, 2001; Wenger & Ingvalson, 2002), multimodal perception-action (e.g., Amazeen & DaSilva, 2005) and visual-haptic stimuli

(e.g., Oberle & Amazeen, 2003), and social perception (e.g., Farris, Viken, & Treat, 2010).

A central issue of characterizing dimensional interactions is distinguishing between interactions at perceptual or decisional levels. For example, faces are widely believed to be processed holistically, such that a whole face is recognized without explicit recognition of face parts. Holistic processing effects suggest that the different dimensions of a face (nose, eyes, mouth, etc.) are combined, but at what level does this interaction occur? Are the face dimensions encoded into a holistic perceptual representation (e.g., Hole, 1994; Young, Hellawell, & Hay, 1987) or are the face dimensions encoded independently at the perceptual level but interact at a later decisional stage (e.g., Wenger & Ingvalson, 2002, 2003)? Understanding not only *that* stimulus dimensions interact but also *how* they interact provides insight into the processes underlying perceptual decision-making.

The theoretical framework of General Recognition Theory (GRT; Ashby & Townsend, 1986) coupled with the empirical analysis tools of Multidimensional Signal Detection Analysis (MSDA; Kadlec & Townsend, 1992) has become one important method for assessing dimensional interactions. In this article, we critically examine MSDA and characterize cases where it is unable to determine the nature of certain kinds of dimensional interactions. After briefly reviewing GRT and MSDA, we investigate the application of MSDA through a series of simulations. Known perceptual or decisional dimensional interactions are embedded in these simulations, MSDA is then applied to the data generated by these simulations, and MSDA is evaluated on its ability to accurately characterize the perceptual versus decisional source of the simulated dimensional interactions. We observed that perceptual interactions are often mischaracterized by MSDA as decisional interactions.

M. L. Mack · J. J. Richler · I. Gauthier · T. J. Palmeri (✉)
Psychology Department, Vanderbilt University,
PMB 407817, 2301 Vanderbilt Place,
Nashville, TN 37240-7817, USA
e-mail: thomas.j.palmeri@vanderbilt.edu

Dimensional interactions in general recognition theory

GRT (Ashby & Townsend, 1986) is a multidimensional generalization of classic signal detection theory (SDT; Green & Swets, 1966), offering a rigorous theoretical framework for investigating dimensional interactions. Like SDT, GRT assumes that perception is inherently noisy. In SDT, perceptual effects are represented by univariate normal distributions of percepts. GRT extends perceptual effects to a multidimensional perceptual space, with stimuli represented by multivariate probability distributions.

Figure 1a illustrates the distributions for four stimuli defined by two dimensions, A and B, which can each take on one of two possible levels, 1 and 2; for purposes of notation, as an example, level 1 along dimension B will be denoted B_1 , and a stimulus that has level 2 along dimension A and level 1 along dimension B will be denoted A_2B_1 . The vertical dimension reflects the likelihood that a physical stimulus will be perceived as some combination of the two perceptual dimensions. Decision bounds, represented by dotted lines in Fig. 1, parse the space into different response regions. These boundaries can be linear or nonlinear. They can be orthogonal or nonorthogonal to the axes of the representational dimensions. To simplify the visual representation of these multidimensional distributions, we can draw contours of equal likelihood, which are cross sections of the distributions at some particular likelihood (Fig. 1b), thereby illustrating variance along each individual dimension and covariance between dimensions. In Fig. 1b, decision boundaries are represented by dotted lines that define four response regions.

Within GRT, dimensional interactions in multidimensional stimuli can be characterized by either perceptual or

decisional factors (Ashby & Townsend, 1986; Kadlec & Townsend, 1992). Specifically, dimensional interactions can have a perceptual locus as violations of *Perceptual Independence* (PI) or violations of *Perceptual Separability* (PS), or a decisional locus as violations of *Decisional Separability* (DS). Examples of how these GRT constructs can be violated are illustrated in Fig. 2, and we will now discuss each in turn.

Stimulus dimensions are perceptually independent when the perceptual effect of one dimension is statistically independent of the perceptual effect of another dimension. When PI is satisfied, variability in the perception of dimension A is uncorrelated with variability in the perception of dimension B, as illustrated by the circular equal likelihood contours in Fig. 1b. PI is violated when the two perceptual dimensions of a stimulus are correlated, as reflected by the diagonal ellipses in Fig. 2a. In this case, some intrinsic property of perceptual processing gives rise to correlated noise across the two dimensions. Unlike violations of PS and DS, PI is a within-stimulus effect, in that it can be observed in a single stimulus.

Stimulus dimensions are perceptually separable when the distribution of perceptual effects for one dimension does not vary across levels of the other dimension. If PS holds, the distribution of the perceived A dimension is unaffected by the level along dimension B, as illustrated by the perceptual distributions forming a rectangle in Fig. 1b. PS is violated when the perception of one dimension depends on the level of the other dimension, which could be reflected in the mean or the variance or both, as illustrated by a non-rectangular arrangement of perceptual distributions in Fig. 2b. In this case, the perception of A_2 depends on whether the stimulus has value B_1 or B_2 along dimension B.

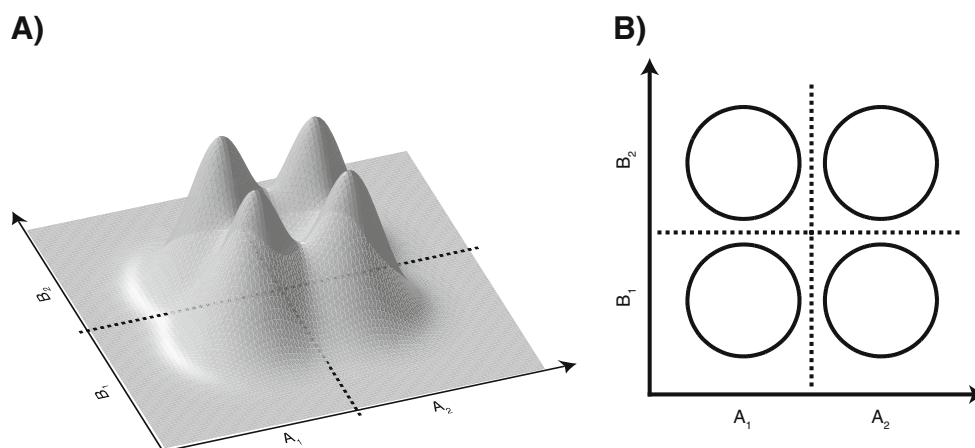


Fig. 1 Panel a illustrates the distribution for four stimuli in two dimensions. The third dimension reflects the likelihood that a physical stimulus will be perceived as some combination of the two dimensions. The two dotted lines parallel to the dimensions represent

decision boundaries. Panel b illustrates a simplified representation of these multidimensional distributions as contours of equal likelihood and decision boundaries that carve the space into different response regions

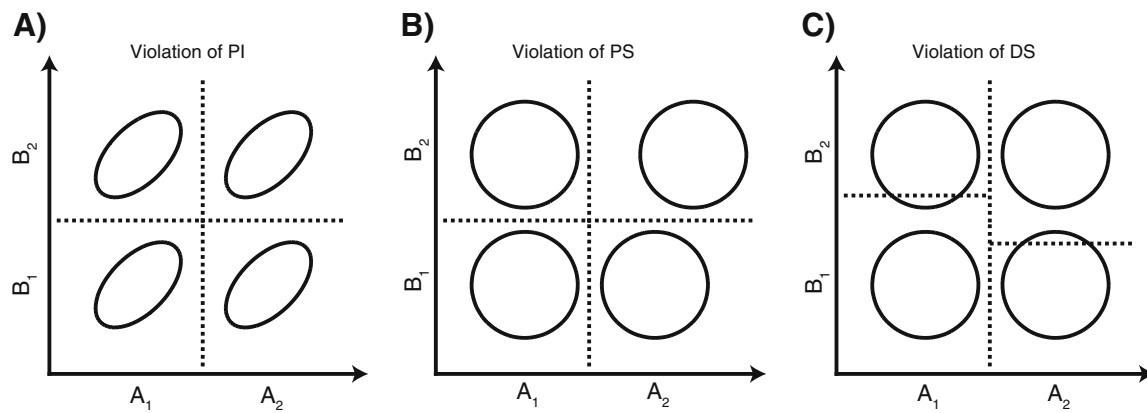


Fig. 2 Schematic of violations of GRT constructs (PI (a), PS (b), and DS (c)). Perceptual distributions are represented by equal likelihood contours and decision boundaries by *dashed lines*

Finally, responses to each dimension of a stimulus are decisionally separable when the location of the boundary for making a decision about one dimension does not depend on the level of the other dimension. For example, if DS holds, the boundary used for decisions about dimension A is in the same location irrespective of the level of dimension B, as illustrated in Fig. 1b. When DS is violated, the location of the decision boundary for one dimension depends on the level of the other dimension, as illustrated in Fig. 2c. For example, if DS is violated, participants might be biased to respond that dimension B has one level versus another level depending on the level of dimension A of the stimulus.

The GRT framework offers a fine-grained approach to considering qualitatively different kinds of dimensional interactions. Of particular interest is the insight from GRT that dimensional interactions that are observed during what is ostensibly a perceptual task could reflect interactions that are taking place at a perceptual level, decisional level, or both. Applying GRT to empirical data to uncover perceptual and decisional loci of dimensional interactions has been performed using two main approaches. One approach involves fitting models to observed data that impose parameter constraints that implement particular violations of GRT constructs (e.g., Ashby & Lee, 1991; Macho, 2007; Maddox, 2001; Maddox & Bogdanov, 2000; Thomas, 2001; Wickens, 1992). Analysis and comparison of these models permits inferences about which GRT constructs hold and which are violated for a given task and stimulus set.

Here we focus our analysis on the second approach, called Multidimensional Signal Detection Analysis (MSDA; Kadlec & Townsend, 1992). MSDA is a statistical toolbox that implements a series of theorems that can be used to make inferences about violations of GRT constructs. While developed over a decade ago, this toolbox has gradually been gathering users doing

research across a wide range of domains. What follows is a summary of MSDA, followed by a series of simulations to test the inferential validity of MSDA. Our focus is on a key inferential limitation and its impacts on distinguishing perceptual versus decisional loci of dimensional interactions.

Multidimensional signal detection analysis

MSDA consists of a set of theorems about the relationship between observed response probabilities and the latent perceptual representations and decisional processes embodied in GRT (Kadlec & Townsend, 1992). An array of statistical tests determines whether empirical data satisfy these theorems, thereby allowing inferences about violations of PI, PS, and DS. MSDA was originally developed in the context of experimental paradigms using simple feature-present/feature-absent stimulus dimensions (Kadlec & Townsend, 1992; Kadlec & Hicks, 1998). However, MSDA has since been applied to a far wider range of paradigms to understand face recognition (Richler et al., 2008; Wenger & Ingvalson, 2002, 2003), perception-action coupling (Amazeen & DaSilva, 2005), visual-haptic interactions (Oberle & Amazeen, 2003), and social perception (Farris, Viken, & Treat, 2010). Furthermore, MSDA is the method of analysis advocated by Macmillan and Creelman (2005) for multidimensional experimental designs.

The statistical tests in MSDA are conducted at two levels of analysis: marginal and conditional. Here we focus on inferences about violations of PS and DS that are assessed with marginal analyses. These include (a) a test of marginal response invariance and (b) tests of equivalence of marginal d' and marginal beta values. The test of marginal response invariance evaluates whether the probability of correctly reporting the level of one dimension is independent of the level of the other dimension; for example, is the probability

of correctly reporting that dimension A has level A_1 independent of whether dimension B has level B_1 or B_2 ? The tests of marginal equivalence compare differences between signal detection parameters d' or β for each level of one dimension collapsed across both levels of the other dimension; for example, one marginal test compares d' when dimension A has level A_1 versus level A_2 collapsed across both levels of dimension B.

The statistical tests of MSDA are related to GRT constructs through a set of theorems and propositions outlined by Kadlec and Townsend (1992). We will briefly review the relevant propositions regarding PS and DS.

According to Proposition 1a, PS holds for a dimension if marginal d' values are equal across the levels of the other dimension. However, as described in Proposition 1b, equivalent marginal d' values does not imply PS since d' is a standardized difference in distribution means. Proposition 1c summarizes the necessary conditions for concluding that PS holds for both dimensions: (i) equal variances of the marginal densities for one dimension across the levels of the other dimension, (ii) equivalence of marginal d' for both dimensions across levels of the other dimensions, and (iii) the means of the perceptual distributions satisfy a Euclidean diagonal relationship. Figure 2b offers a simple illustration of a violation of this proposition: in this case, marginal d' for dimension A when dimension B has level B_1 is not equal to marginal d' for dimension A when dimension B has level B_2 , thus condition (ii) is not satisfied, thereby indicating a violation of PS.

There are three important points to highlight about this proposition: (1) PS holds for both dimensions only when all three of these conditions are satisfied, (2) PS is assessed independently of DS, and (3) satisfying both conditions (ii) and (iii) requires a rectangular configuration of the perceptual distributions. One test of the rectangularity of a perceptual space, known as a *diagonal d' test*, was initially suggested by Kadlec and Townsend (1992) and fully described by Kadlec and Hicks (1998). The test involves assessing the distances between the diagonally separated distributions in separate blocks (i.e., the distance between A_1B_1 and A_2B_2 versus the distance between A_1B_2 and A_2B_1); rectangular configurations will have equal diagonal distances. However, the diagonal d' test is known to be inappropriate when PI is violated or when perceptual distributions have unequal variances (Thomas, 1995, 1999, 2003).

The necessary conditions for DS are described in proposition 2a and 2b of Kadlec and Townsend (1992). Proposition 2a states that if DS and PS hold for a dimension, the marginal betas for that dimension are equal across the levels of the other dimension. Figure 2c illustrates a violation of this proposition: The criterion value for dimension B depends on the level of dimension

A, resulting in a difference in marginal betas, thereby indicating a violation of DS. Unlike the direct test of PS, the test of DS is indirect in that it depends on the status of PS. This relationship is further clarified in proposition 2b (i): If DS holds but PS fails, then it is not necessarily true that marginal beta values for one dimension across the levels of the other dimensions will be equal. In other words, a difference in criterion values is consistent with a violation of DS, but it does not logically follow that DS is actually violated.

Two implications fall out of these propositions. The more general implication is that applying MSDA's inferential logic to empirical data is governed by the relationship between DS and PS: Inferences about the status of DS depend on whether PS is supported or rejected.

The second implication is that a violation of PS may influence estimates of the decision criteria used to make inferences about DS. At first blush, this seems to mean only that the inference for PS must be considered before assessing DS. Indeed, following the inferential logic proposed by Kadlec and Townsend (1992, their Fig. 8, p. 352), when PS and marginal response invariance are rejected, no inferences can be drawn about DS based on marginal tests. This speaks to the asymmetry in MSDA's inferential logic; if PS and DS hold, marginal estimates will be equivalent, but equivalent marginal estimates do not necessarily indicate that PS and DS hold. Beyond this general limitation of MSDA's logic, another aspect of this implication that is not universally recognized is that the estimation of critical measures for assessing DS may be influenced by any deviation in marginal d' values, regardless of whether statistical tests suggest that PS is supported or PS is rejected. This may lead to erroneous inferences about DS. Here we investigate in a series of simulations whether violations of PS have a systematic influence on the estimation of the decision criteria, thereby influencing how MSDA draws inferences regarding DS.

Simulations

Our tests of MSDA follow a straightforward logic: A simulated space of distributions and decision boundaries are created in a way that violates one specific GRT construct in some qualitative way and by some quantitative degree. If MSDA successfully uncovers that violation, and does not erroneously uncover a violation that is not present, then MSDA has made a successful inference; otherwise, it has not.

Each simulation included four stimulus conditions. Each stimulus condition was associated with a multivariate normal distribution in two-dimensional space, as illustrated

earlier. Each simulation used a total of 2,000 trials.¹ On each simulated trial, a random sample stimulus was drawn from one of the four distributions. Because normal distributions are used, a sample stimulus from any of the four stimuli distributions could be located in any of the four response regions defined by the decision boundaries. This results in a 4×4 confusion matrix, with each row a stimulus and each column a response. The resulting confusion matrix was analyzed with MSDA, as described below. We repeated the simulation and MSDA analyses 5,000 times for each space of distributions and decision boundaries.

We conducted two versions of MSDA marginal statistical tests on the response probabilities in the simulated confusion matrix. The first followed the standard methods of estimating and comparing signal detection parameters and variances, like that outlined in Macmillan and Creelman (2005). Marginal response invariance was assessed by an equivalence test of probabilities of responding to a dimension across the levels of the other dimension. Violations of PS were assessed by differences in marginal d' values for the relevant dimension. Violations of DS were assessed by differences in marginal $c = -0.5[\Phi^{-1}(\textit{hit rate}) + \Phi^{-1}(\textit{false alarm rate})]$, where Φ is the standard normal distribution function. Marginal c is used instead of marginal beta as an estimate of decision criteria because statistical equivalence tests exist for marginal c , but not for marginal beta.

The second version of MSDA tests followed the methods of Kadlec (1995) using a Matlab implementation of the *MSDA_2* software (Kadlec, 1995, 1999); our Matlab implementation produces identical results to the original Pascal implementation of *MSDA_2*. We used *MSDA_2* because it has become a common off-the-shelf tool for conducting MSDA analyses (Amazeen & DaSilva, 2005; Copeland & Wenger, 2006; Farris et al., 2010; Oberle & Amazeen, 2003; Richler et al., 2008; Wenger & Ingvalson, 2002, 2003). Unlike the standard method, *MSDA_2* decision bounds are estimated by marginal $\textit{crit} = -\Phi^{-1}(\textit{false alarm rate})$, denoted henceforth by $z(\textit{FAR})$.

Each simulation began with a configuration representing no violations of PS or DS: the means of the multivariate normal distributions were equally spaced from neighboring distributions by 1 unit, the distributions were organized in a

square configuration, all distributions had equal standard deviation of 1 unit along each dimension and no covariance, and decision boundaries equally separated the distributions at the midpoint between the means along a dimension, as illustrated in Fig. 1. Systematic violations of GRT constructs were then introduced by systematically varying parameters of the multivariate normal distributions or decision boundaries (see Fig. 3). DS was violated by shifting the decision boundary for one dimension depending on the level of the other dimension (Δc ; Fig. 3a). PS was violated in two different ways: by shifting the location of one distribution along one dimension to increase the marginal d' for one value of the second dimension ($\Delta d'$; Fig. 3b) and by shifting the location of two distributions along one dimension to introduce a nonrectangularity in the configuration of the distributions ($\Delta d'$; Fig. 3c). Each construct (PS and DS) was investigated independently; for example, when examining a violation of PS ($\Delta d' > 0$), we assumed no violation of DS ($\Delta c = 0$).

Simulation results are summarized in Fig. 3. Each of the three simulated violations, depicted in the left column, is a row in the figure. The results of the marginal tests, using the standard method of estimating signal detection parameters (d' , c , MRI) as well as the criterion value from *MSDA_2* ($z(\textit{FAR})$), are shown in the middle column as the proportion of simulations that resulted in a significant difference on the test. The right column summarizes the inferential conclusions of MSDA both with the standard method (black bars) and *MSDA_2* (white bars). The plots in the right column show the proportion of simulations with the various combinations of PS and DS inferences for the highest degree of the simulated violation used in the middle column (e.g., in Fig. 3a, the right column panel corresponds to the MSDA inferences when $\Delta c = 0.4$). Notationally, the x-axis of the right column panels signify all six possible combinations of inferences, with DS or PS denoting no violation, \sim DS or \sim PS denoting a violation, and ?DS or ?PS denoting cases where inferences cannot be made.

We first present simulations of violations that serve as a simple test of MSDA and allow us to validate our simulation methods. For a violation of DS (Fig. 3a), the marginal tests of MSDA correctly inferred the nature of the violations that produced the data. The marginal c and $z(\textit{FAR})$ tests showed significant differences that increased in proportion with larger violations while marginal d' tests were unaffected. Following the inferential logic of MSDA, the constant marginal d' values infer support for PS and the significant difference in marginal c and $z(\textit{FAR})$ values infer a violation of DS. This is reflected in the relatively large proportion of correct “PS, \sim DS” inferences in the plot in the right column. Since the $z(\textit{FAR})$ measure of *MSDA_2* is dependent only on the lower marginal distribution, it is less

¹ The number of trials per simulation (2000) is similar to the number of trials used in a recent study that employed MSDA in the context of face recognition (Richler et al., 2008). We also conducted simulations with fewer trials (200, 400, and 1,000). In general, fewer trials per simulation led to fewer significant differences detected in all of the statistical tests, as would be expected by the lower power of these tests. Importantly, the relative proportion of significant differences between the marginal tests and the inferences with regard to PS and DS were qualitatively similar to the results reported here for simulations with 2,000 trials.

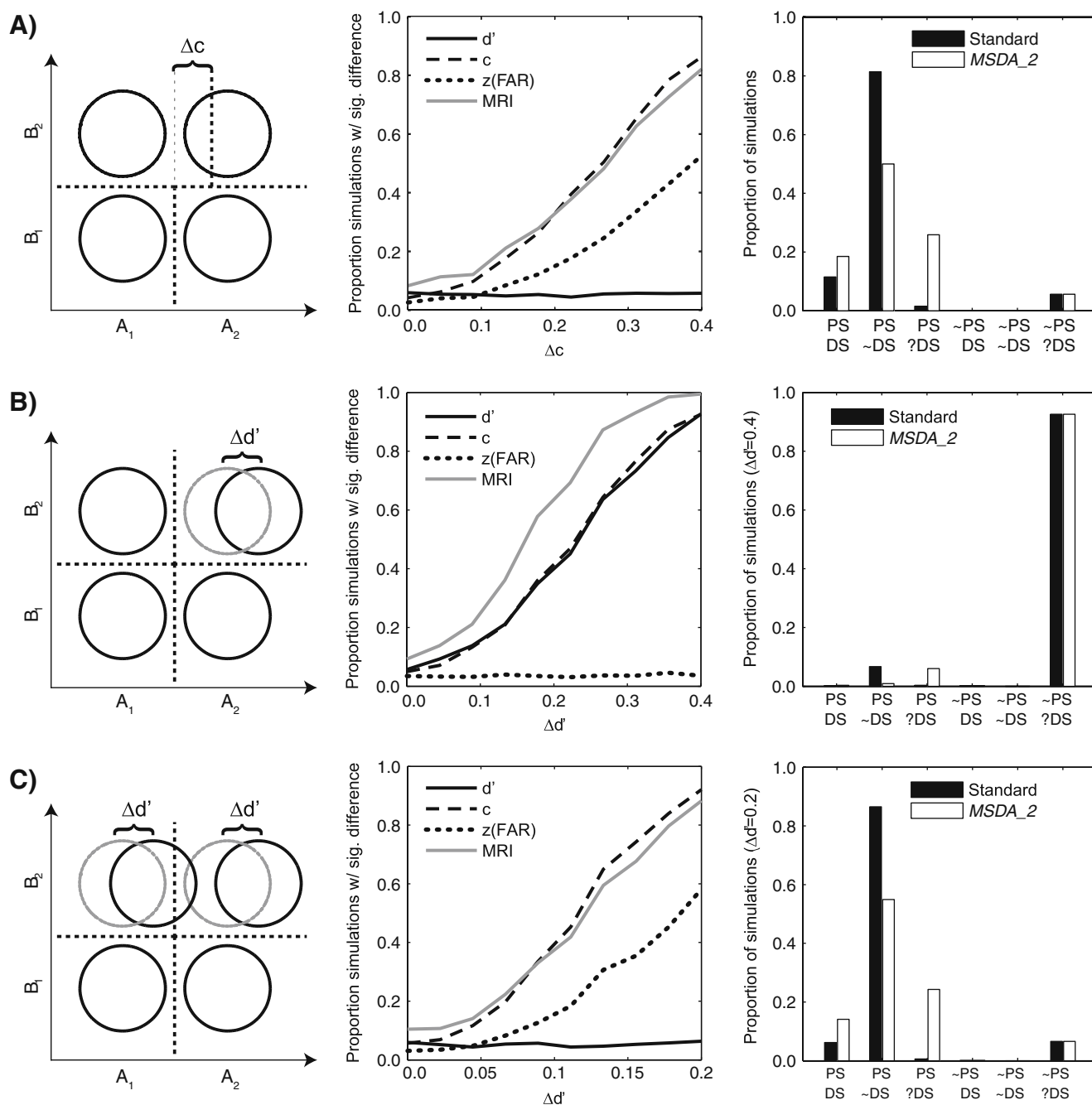


Fig. 3 Results for simulated violations of DS and PS. The *left column* illustrates the type of violation in each simulation in terms of the change in perceptual distributions or decision criteria. The *middle column* shows the results of the MSDA marginal tests. The proportion of iterations with a significant difference in the marginal response invariance test (gray line), marginal d' test (black line), marginal c test (dashed line), and marginal $z(\text{FAR})$ test (dotted line) are plotted as

function of the size of the simulated violation (Δc or $\Delta d'$). The *right column* summarizes the inferential conclusions of both the standard method of MSDA (black bars) and MSDA_2 (white bars; Kadlec, 1995, 1999). The plots show the proportion of simulations with the various combinations of PS and DS inferences (“~” indicates a violation, “?” indicates that an inference cannot be made) for the highest degree of the simulated violation

sensitive to the shift in criteria than marginal c . This results in fewer inferences of a violation of DS and more inferences of “PS, ?DS” (the status of DS cannot be inferred if PS holds, marginal c values are equivalent, and MRI does not hold [Kadlec & Townsend, 1992]); even so,

MSDA_2 makes the correct inference in the largest proportion of simulations.

For the first simulated violation of PS (Fig. 3b), the proportion of significant differences in marginal d' increased with a larger violation, suggesting a violation of PS.

Note that the proportion of significantly different marginal c values matched that of the measures for detecting PS violations; this is consistent with the known relationship between violations of PS and certain estimates of decision criteria (Kadlec & Townsend, 1992; proposition 2b). MSDA includes the necessary logic to manage this relationship; PS is violated and marginal response invariance does not hold, so no inferences about DS can be drawn. The marginal $z(\text{FAR})$ measure in *MSDA_2* is not affected by the violation of PS. In this simulation, both versions of MSDA correctly report that PS is violated ($\sim\text{PS}$) and the status of DS is unknown (?DS).

The simulation above shows that differences in marginal d' values can introduce an artifact in the estimates of marginal c , such that when PS is violated, DS cannot be assessed. We next show that this same artifact in estimating decision criteria can occur when PS is violated but MSDA fails to detect that violation. This leads to an erroneous inference that a violation of DS is present, when it is not.

Figure 3c illustrates the other simulated violation of PS, which is a version of mean-shift integrality (Maddox, 1992).² The relative distance between the perceptual distributions along a dimension at the two levels of the other dimension are equivalent, but there exists a (mean) shift in the representations depending on the level of a dimension: the representation of one dimension depends on the level of the other dimension. Note that the decision boundaries used in this set of simulations remain constant across the levels of the two dimensions. So, in these simulations, we know that PS is violated and DS holds. It has been well documented that standard application of MSDA as originally proposed by Kadlec and Townsend (1992) is incapable of dealing with mean-shift integrality. Without a test of the rectangular configuration of the perceptual distributions, mean-shift integrality goes undetected. To be clear, the propositions of MSDA clearly define mean-shift integrality as a violation of PS (Kadlec & Townsend, 1992). The limitation is in detecting this violation when applying MSDA to empirical data. As expected, when the simulated data were analyzed using both methods of MSDA, marginal d' values were constant across the magnitude of the simulated violation and PS is erroneously inferred, as reflected in the right panel of Fig. 3c.

² We also conducted simulated violations of PS caused by differences in variance along particular values of dimensions, keeping means constant. While the MSDA propositions have conditions that mandate equal variances and covariances, testing these conditions are not part of standard MSDA analyses and are rarely tested in practice. About 60% of simulations correctly inferred violations of PS, despite the fact that the MSDA analyses are not designed specifically to pick up violations that might be caused by differences in variance. The remaining simulations inferred no violation of PS, with about 20% inferring $\sim\text{DS}$ and 10% each inferring DS or ?DS.

What about DS? As the size of the simulated mean shift increases, the proportion of significant differences in marginal c and $z(\text{FAR})$ values also increases, as illustrated in the figure. According to the propositions underlying MSDA, we expect violations of PS to create significant differences in the test of marginal c values. Similarly, a shift in the mean of the lower marginal distribution will create a significant differences in the test of marginal $z(\text{FAR})$. These simulations emphasize that a mindset that might be adopted using unidimensional signal detection theory should not be applied to the multidimensional case. Here, the significant changes in criterion, marginal c and $z(\text{FAR})$, do not reflect a true decisional effect, but are artifacts caused by an underlying violation of PS.

Moreover, with a straight application of MSDA, without a test for mean shift integrality, finding that PS holds (in this case erroneously) and that there is a significant difference in marginal criterion values, implies that there is a violation of DS (also erroneous). With the standard method version of MSDA, the nonrectangular configuration of perceptual distributions goes undetected and an incorrect inference about a violation of DS occurs on approximately 90% of the simulations, as show in the right panel of Fig. 3c. *MSDA_2* makes the incorrect inference about a violation of DS on approximately 55% of the simulations, again because of its less sensitive criterion measure. For both versions of MSDA, the appropriate inference (“ $\sim\text{PS}$, ?DS”) occurred in only 5% of simulations.

General discussion

There has been growing interest in characterizing perceptual versus decisional components of dimensional interactions in a wide variety of domains, ranging from multimodal interactions to face recognition to social perception (e.g., Amazeen & DaSilva, 2005; Copeland & Wenger, 2006; Farris et al., 2010; Oberle & Amazeen, 2003; Richler et al., 2008; Wenger & Ingvalson, 2002, 2003). This work has used a statistical technique called Multidimensional Signal Detection Analysis (MSDA; Kadlec & Townsend, 1992) to characterize perceptual versus decisional loci using constructs from General Recognition Theory (GRT; Ashby & Townsend, 1986). We reported simulations that highlight a significant inferential limitation of MSDA that has been underappreciated in its application to distinguishing perceptual versus decisional sources of dimensional interactions.

The key focus of our critique was a form of dimensional interaction called mean shift integrality (e.g., see Maddox, 2001), a violation of perceptual separability in the language of GRT. It has been long acknowledged that the standard application of MSDA, including the widely used *MSDA_2*

toolkit (Kadlec, 1999), does not include tests for mean shift integrality in its inferential logic. On its own, this could simply mean that some violations of PS might go undetected if only MSDA were used. However, inferences about DS depend entirely upon whether valid inferences about PS are made. According to the propositions underlying MSDA, if PS is violated, then no valid inferences about DS can be made. Therefore, if violations of PS go undetected, erroneous inferences about violations of DS can be the result.

This is what happens in simulated cases of mean shift integrality. Differences in the location of the perceptual distributions introduce an artifact in estimates of decision criteria. This mean shift goes undetected by tests of marginal d' values but leads to a significant difference in marginal c and $z(\text{FAR})$ values. PS is violated but goes undetected; DS is not violated, but an erroneous violation of DS is inferred because of the significant difference in criterion. Failing to detect mean shift integrality that is present is not simply a matter of failing to characterize a potentially important perceptual locus of dimensional interactions. Failing to detect mean shift integrality that is present can lead to erroneous inferences that a decisional locus of dimensional interactions exists when it does not.

We have concentrated on a somewhat idealized version of mean-shift integrality where the mean shift for both distributions along one value of a dimension is equivalent. However, the problem we are describing is not limited to this special case. Any shift in the means of the marginal distributions, equivalent across distributions or not, can introduce an artifact in the estimation of decision criteria. When the mean difference goes undetected (e.g., underpowered analyses, small effect, high variability), the inference for DS will be confounded.

Both of the MSDA methods we tested assessed PS with tests of marginal d' and MRI without any test of the rectangularity of the perceptual distributions, so the strength of the inferences that can be made about PS, and hence DS as well, are limited (Kadlec & Townsend, 1992). As noted earlier, a diagonal d' test has been proposed as an additional constraint on assessing PS (Kadlec & Hicks, 1998). However, this test requires the assumption of a distance classifier and has been shown to be invalid when perceptual distributions exhibit unequal or correlated variances across stimulus dimensions (Thomas, 1995, 1999, 2003). These assumptions are clearly inappropriate for any experimental setting. New tests are needed, not only to correctly characterize the full spectrum of violations of PS, but to allow valid inferences regarding DS as well.

It is important to place our criticism in its appropriate context. We are not rejecting the theoretical framework of GRT (Ashby & Townsend, 1986) or MSDA (Kadlec & Townsend, 1992). GRT and the theoretical underpinnings of

MSDA are sound. The main issue we have highlighted is a breakdown in applying the propositions of MSDA. From the theoretical perspective of MSDA, any violation of PS prevents any inferences to be drawn about DS. Often, differences in the location (or variances) of marginal densities are not rigorously tested. Even if tested, care must be taken to avoid the possibility that these differences go undetected due to a variety of factors (e.g., small effect size, too few data points, high variability), which could lead to an erroneous inference that PS holds.

To our knowledge, this work is the first to document problems with MSDA related to incorrect inferences regarding DS driven by violations of PS. The limitations of MSDA in inferring certain violations of PS per se have been long known and acknowledged (e.g., Kadlec & Townsend, 1992). However, when PS is violated but remains undetected, following the propositional logic of MSDA can lead to erroneous conclusions about DS. Illustrating this problem seems particularly important considering that the vast majority of studies that apply the MSDA framework find evidence for violations of DS, sometimes in cases where such violations seem counterintuitive (Amazeen & DaSilva, 2005; Farris et al., 2010; Oberle & Amazeen, 2003; Valdez & Amazeen, 2008; Wenger & Ingvalson, 2002, 2003) including some of our own work (Richler et al., 2008). All of these studies employed MSDA methods similar to the approaches we used here (Kadlec & Townsend, 1992; Kadlec, 1995, 1999), one of these studies included additional tests of diagonal d' (Wenger & Ingvalson, 2003), and a few included converging model-fitting methods (Copeland & Wenger, 2006; Cornes, Donnelly, Godwin, & Wenger, 2010; Valdez & Amazeen, 2008).

It is quite possible that many of these cases reflect true violations of DS. There is converging evidence that certain kinds of dimensional interactions that seem perceptual may be caused by decisional factors (e.g., Cheung, Richler, Palmeri, & Gauthier, 2008). Our research reported in this paper does not discount the decisional results found using MSDA. Instead, those inferences remain equivocal. The critical problem is distinguishing true violations of DS from violations of DS produced by artifacts. One alternative direction is found in the method of fitting GRT models to empirical data (e.g., Ashby & Lee, 1991; Macho, 2007; Wickens, 1992) alongside drawing inferences with MSDA (e.g., Thomas, 2001) to find converging evidence for the status of GRT constructs. Unfortunately, such techniques often require paradigms that demand significantly more data points than those that have been typically analyzed using MSDA.

The recent use of MSDA in new domains (e.g., Farris et al., 2010) and its recommendation in the latest edition of Macmillan and Creelman's *Detection Theory: A User's Guide* (2005) for designs that are aimed at assessing SDT in

multidimensional spaces, press for even greater awareness of the current limitations in applying MSDA in practice. We hope that this article will prompt further research into developing new inferential tools that will allow researchers to feel confident about making inferences regarding perceptual versus decisional loci of dimensional interactions using the language of GRT.

Acknowledgements This work was supported by the Temporal Dynamics of Learning Center (SBE-0542013), an NSF Science of Learning Center, and a grant from the James S. McDonnell Foundation.

References

- Amazeen, E. L., & DaSilva, F. (2005). Psychophysical test for the independence of perception and action. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 170–182.
- Ashby, F. G., & Lee, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General*, *120*, 150–172.
- Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154–179.
- Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2008). Revisiting the role of spatial frequencies in the holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1327–1336.
- Copeland, A. M., & Wenger, M. J. (2006). An investigation of perceptual and decisional influences on the perception of hierarchical forms. *Perception*, *35*, 511–529.
- Cornes, K., Donnelly, N., Godwin, H., & Wenger, M. J. (2010). Perceptual and decisional factors influencing the discrimination of inversion in the thatcher illusion. *Journal of Experimental Psychology: Human Perception and Performance*.
- Farris, C., Viken, R. J., & Treat, T. A. (2010). Perceived association between diagnostic and non-diagnostic cues of women's sexual interest: General Recognition Theory predictors of risk for sexual coercion. *Journal of Mathematical Psychology*, *54*(1), 137–149.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception*, *23*, 65–74.
- Kadlec, H. (1995). Multidimensional signal detection analyses (MSDA) for testing separability and independence: A Pascal program. *Behavior Research Methods, Instruments, & Computers*, *27*, 442–458.
- Kadlec, H. (1999). MSDA2: Updated version of software for multidimensional signal detection analyses. *Behavior Research Methods, Instruments, & Computers*, *31*, 384–385.
- Kadlec, H., & Hicks, C. L. (1998). Invariance of perceptual spaces and perceptual separability of stimulus dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 80–104.
- Kadlec, H., & Townsend, J. T. (1992). Signal detection analysis of dimensional interactions. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 181–228). Hillsdale: Lawrence Erlbaum Associates.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah: Lawrence Erlbaum Associates.
- Macho, S. (2007). Feature sampling in detection: Implications for the measurement of perceptual independence. *Journal of Experimental Psychology: General*, *136*(1), 133–153.
- Maddox, W. T. (1992). Perceptual and decisional separability. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 147–180). Hillsdale: Lawrence Erlbaum Associates.
- Maddox, W. T. (2001). Separating perceptual processes from decisional processes in identification and categorization. *Perception & Psychophysics*, *63*, 1183–1200.
- Maddox, W. T., & Bogdanov, S. V. (2000). On the relation between decision rules and perceptual representation in multidimensional perceptual categorization. *Perception & Psychophysics*, *62*, 984–997.
- Oberle, C. D., & Amazeen, E. L. (2003). Independence and separability of volume and mass in the size-weight illusion. *Perception & Psychophysics*, *65*, 831–843.
- Richler, J. J., Gauthier, I., Wenger, M. J., & Palmeri, T. J. (2008). Holistic processing of faces: Perceptual and decisional components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 328–342.
- Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54–87.
- Thomas, R. D. (1995). Gaussian general recognition theory and perceptual independence. *Psychological Review*, *102*, 192–200.
- Thomas, R. D. (1999). Assessing sensitivity in a multidimensional space: Some problems and a definition of a general d' . *Psychonomic Bulletin & Review*, *6*, 224–238.
- Thomas, R. D. (2001). Perceptual interactions of facial dimensions in speeded classification and identification. *Perception & Psychophysics*, *63*, 625–650.
- Thomas, R. D. (2003). Further considerations of a general d' in multidimensional space. *Journal of Mathematical Psychology*, *47*, 220–224.
- Valdez, A., & Amazeen, E. (2008). Sensory and perceptual interactions in weight perception. *Perception & Psychophysics*, *70*, 647–657.
- Wenger, M. J., & Ingvalson, E. M. (2002). A decisional component of holistic encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 872–892.
- Wenger, M. J., & Ingvalson, E. M. (2003). Preserving informational separability and violating decisional separability in facial perception and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1106–1118.
- Wickens T. D. (1992). Maximum likelihood estimation of a multivariate Gaussian rating model with excluded data. *Journal of Mathematical Psychology*, *36*, 213–234.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747–759.