Journal of Experimental Psychology:
Learning, Memory, and Cognition
2008, Vol. 34, No. 2, 328–342

0278-7393/08/$12.00 DOI: 10.1037/0278-7393.34.2.328

Holistic Processing of Faces: Perceptual and Decisional Components

Jennifer J. Richler and Isabel Gauthier
Vanderbilt University

Michael J. Wenger
Pennsylvania State University

Thomas J. Palmeri
Vanderbilt University

Researchers have used several composite face paradigms to assess holistic processing of faces. In the selective attention paradigm, participants decide whether one face part (e.g., top) is the same as a previously seen face part. Their judgment is affected by whether the irrelevant part of the test face is the same as or different than the relevant part of the study face. This failure of selective attention implies holistic processing. However, the authors show that this task alone cannot distinguish between perceptual and decisional sources of holism. The distinction can be addressed by the complete identification paradigm, in which both face parts are judged to be same or different, combined with analyses based on general recognition theory (F. G. Ashby & J. T. Townsend, 1986). The authors used a different paradigm, sequential responses, to relate these 2 paradigms empirically and theoretically. Sequential responses produced the same results as did selective attention and complete identification. Moreover, disruptions of holistic processing by systematic misalignment of the faces corresponded with systematic and significant changes in the decisional components, but not in the perceptual components, that were extracted using general recognition theory measures. This finding suggests a significant decisional component of holistic face processing in the composite face task.

Keywords: face perception, holistic processing, decisional factors

It is widely believed that people recognize faces by using some form of "holistic" processing, in which the whole face may be recognized without any explicit recognition of face parts. For example, facial features, such as a nose or a mouth, are recognized better in the context of a whole face than in isolation (Farah & Tanaka, 1993). Furthermore, recognition of individual facial features is impaired when the configural information (spatial relations between features) in the test face differs from that in the study face (Tanaka & Sengco, 1997); this effect is not found with scrambled faces, inverted faces, or common objects.

The composite effect also provides evidence for holistic face recognition. Composite faces combine the top half of one face with the bottom half of another face. When aligned, the face halves appear to fuse together to produce a novel face; this effect makes it difficult for one to selectively recognize either half of the composite. Given a composite of the top of George Bush's face and the bottom of Tony Blair's face, participants are slow and make errors when asked to recognize one half of an aligned, compared with a misaligned, composite (Young, Hellawell, & Hay, 1987).

A similar cost is observed in a sequential matching task that uses unfamiliar composite faces (Hole, 1994). In this task, participants study one face composite and judge whether the top or bottom of a second face is the same as or different than the relevant part of the composite, after a brief delay. Because participants must ignore the irrelevant half of the composite, we refer to this task as the selective attention task. Accuracy in judging the same–different status of the relevant part is significantly affected by the same–different status of the irrelevant part. However, if the composite faces are inverted, or if the top and bottom halves of the test face are misaligned, the same–different status of the irrelevant part has little impact on performance (Diamond & Carey, 1986; Hole, 1994; Young et al., 1987). When the meaningful configuration of facial features is disrupted by inversion or misalignment, holistic processing may be disrupted as well (see also Tanaka & Sengco, 1997).

Although such holistic effects are usually described as true perceptual effects (e.g., Young et al., 1987; Hole, 1994), these effects could arise from decisional factors, such as differences in response bias (Wenger & Ingvalson, 2002, 2003; Wenger & Townsend, 2006), or from differences in performance that have little to do with whether or not a stimulus is being processed holistically (e.g., Ingvalson & Wenger, 2005; Loftus, Oberg, & Dillon, 2004; O'Toole, Wenger, & Townsend, 2001; Sekuler, Gaspar, Gold, & Bennett, 2004). The relationship between percep-

Jennifer J. Richler, Isabel Gauthier, and Thomas J. Palmeri, Department of Psychology, Vanderbilt University; Michael J. Wenger, Department of Psychology, Pennsylvania State University.

This work was supported by a grant from the James S. McDonnell Foundation. It was also supported by the Temporal Dynamics of Learning Center (National Science Foundation Science of Learning Center Grant SBE-0542013). Supplementary information related to the GRT analyses reported in this article is available online at http://www.psy.vanderbilt.edu/faculty/palmeri/holistic2007/. We would like to thank Michael Mack for helpful discussion and for his assistance in implementing and running the Monte Carlo simulations.

Correspondence concerning this article should be addressed to Thomas J. Palmeri, Department of Psychology, Vanderbilt University, 301 Wilson Hall, Nashville, TN 37203. E-mail: thomas.j.palmeri@Vanderbilt.edu

tual and decisional factors is articulated in classic signal detection theory (Green & Swets, 1966) and was generalized to multiple dimensions within general recognition theory (GRT; Ashby & Townsend, 1986).

Researchers who have adopted a GRT approach have concluded that there is a significant decisional component to holistic processing of faces (Wenger & Ingvalson, 2002, 2003). However, these studies used a complete identification task that requires divided attention across both halves of a composite face, whereas the vast majority of face recognition studies have used a selective attention task. It is unknown whether evidence for a decisional component will be found in a task that requires selective rather than divided attention. If these tasks produced different results due to their differing task demands, the generality of decisional influences in face processing would be brought into question. Furthermore, prior work has not examined whether decisional measures of holism within GRT are systematically affected by experimental manipulations (e.g., misalignment) that are known to systematically disrupt holistic processing.

At this juncture, let us provide a road map for the rest of this article. In the next section, we discuss how performance in the composite face paradigm can be described within GRT. We next describe the complete identification task and relate it to the selective attention task. We then propose using a new task, the sequential responses task, which combines key elements of the complete identification task and the selective attention task. The sequential responses task forges an empirical bridge between the literatures on the other two tasks. As we show, it allows us to measure selective attention to a single part of a face, as does the selective attention task, and it permits GRT analyses of perceptual and decisional components of holistic processing, as does the complete identification task. We close the introductory section with a demonstration through Monte Carlo simulations of how GRT analyses can reveal more than can some standard behavioral measures of holistic face recognition. We then describe a study that experimentally compares the sequential responses task with both the complete identification task and the selective attention task. Finally, we describe an experiment that measures how perceptual and decisional components of holistic processing within a GRT framework change when we systematically manipulate the alignment of the test composite.

GRT

The paradigms we discuss in this article are same–different tasks. When signal detection theory is applied to a standard same–different task, there is a percept of the first stimulus, a percept of the second stimulus, and a comparison process that computes similarity between the stimuli. For example, in a simple face matching task, in which participants judge whether two sequentially presented faces are the same or not, we have two distributions of perceptual similarity, one for same trials and one for different trials. A response criterion determines whether the participant responds "same" or "different." Discriminability and response bias can be calculated on the basis of hits and false alarms (Green & Swets, 1966).

GRT (Ashby & Townsend, 1986) is a multidimensional generalization of signal detection theory that can be used to distinguish between perceptual and decisional loci of holistic effects. When stimuli are multidimensional, the perceptual effect of a combination of components is represented by a multidimensional probability distribution. Figure 1A illustrates this distribution for four stimuli defined by two dimensions. Perceptual effects are noisy, so the third dimension reflects the likelihood that a physical stimulus will be perceived as some combination of the two dimensions. To simplify the visual representation of these multidimensional distributions, we draw contours of equal likelihood, which are cross sections of the distributions, as in Figure 1B. These cross sections can readily illustrate variance along each individual dimension and any covariance among dimensions. With a multidimensional space, a decision process employs decision boundaries to parse that space into different response regions; these boundaries can be linear or nonlinear, and they can be orthogonal or nonorthogonal to the axes of the perceptual dimensions. For the composite face

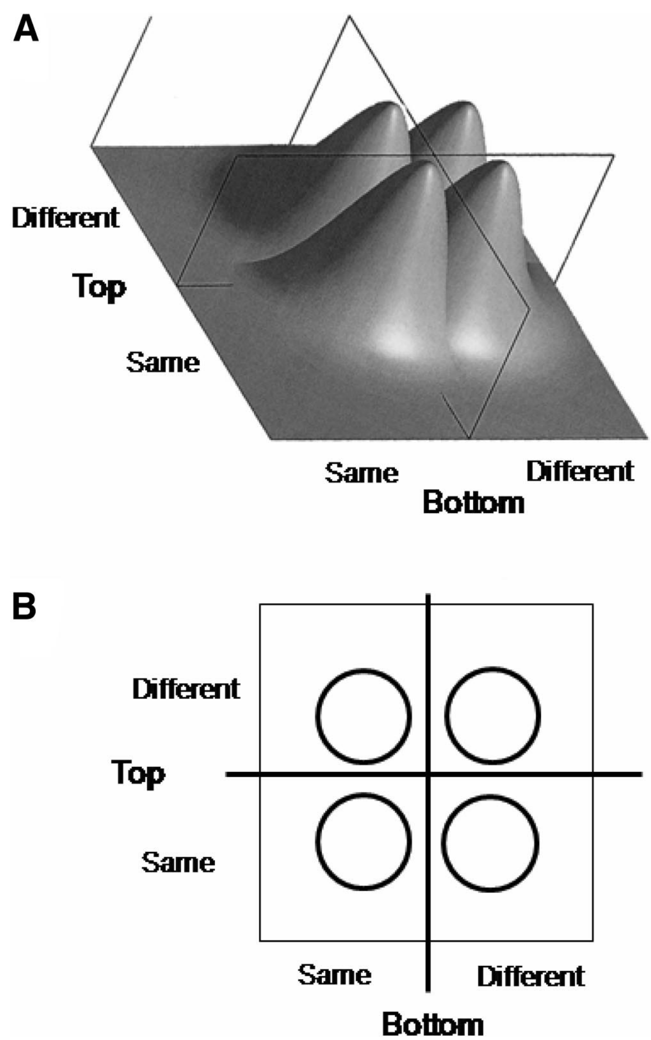

Figure 1. Example of a representation of the four stimuli in a sequential matching task in three-dimensional perceptual space. A: Bivariate normal probability distributions of perceptual effects for each of the four stimuli and decision boundaries. B: The contours of equal likelihood for these distributions, which are created by slicing each distribution along an equal plane and then viewing the sliced distributions from above.

paradigm, the dimensions reflect the same–different status of the top and bottom parts of the test face, and the decision boundaries reflect the criteria for choosing a same or different response for the top or bottom part.

The theoretical power of GRT with respect to questions about face perception and memory comes from defining holism in terms of perceptual and decisional factors (O'Toole et al., 2001; Thomas, 2001a, 2001b; Wenger & Ingvalson, 2002, 2003). According to GRT, holism can emerge from violations of perceptual independence (PI), perceptual separability (PS), or decisional separability (DS), in any combination. These constructs are described in turn.

Stimulus dimensions are perceptually independent when the perceptual effect of one part is statistically independent of the perceptual effect of another part. If faces exhibit PI and if the distributions of perceptual evidence are modeled as being multivariate normal, variability in the perceived sameness of the top part would be uncorrelated with variability in the perceived sameness of the bottom part. This configuration is illustrated by the circular equal likelihood contours in Figure 2A(i). PI is violated when the two perceptual dimensions of some stimulus are correlated, resulting in the elongated ellipse seen in Figure 2A(ii). For instance, some intrinsic property of perceptual processing could give rise to correlated noise across parts. Unlike the other violations we discuss, PI is deemed a within-stimulus effect, and it is possible to observe PI within a single stimulus. Because of this possibility, a violation of PI has been described as the strongest form of holism within GRT (O'Toole et al., 2001; Wenger & Ingvalson, 2002, 2003).

Stimulus dimensions are perceptually separable when the distribution of perceptual effects for one dimension does not vary across levels of the other dimension. If composite faces exhibit PS, the distribution of the perceived sameness of the top part would be unaffected by whether the bottom part is same or different. As shown in Figure 2B(i), this configuration can be illustrated by connecting the centers of the four perceptual distributions into a rectangle. PS is violated when the distribution of perceived sameness for one part depends on the level of perceived sameness of the other part. As shown in Figure 2B(ii), this violation occurs when the connected perceptual distributions form a nonrectangular quadrilateral. For example, if faces violate PS, face bottoms may appear more different from one another when the tops are different than when the tops are the same.

Finally, responses to each part of the stimulus are decisionally separable when the location of the decision boundary about one dimension is unaffected by the level of the other dimension. If DS applies to faces, the boundary established for decisions about the bottom part is in the same location irrespective of whether the top part is same or different. As shown in Figure 2C(i), this configuration can be illustrated by linear decision boundaries that are parallel to dimensional axes. As shown in Figure 2C(ii), when DS is violated, the location of the decision boundary for one part depends on the other part. For example, different decision boundaries for whether the bottom part is same or different are established, depending on whether the top part is same or different.

Tasks Used to Assess Holistic Processing

Although the GRT framework has the power to distinguish between perceptual and decisional loci of holistic effects, only a few studies have applied this framework to face recognition (e.g., Thomas, 2001a, 2001b; Wenger & Ingvalson, 2002, 2003). Unfortunately, data from the standard selective attention task used throughout most of the face literature (e.g., Boutet, Gentes-Hawn, Chaudhuri, 2002; Farah, Wilson, Drain, & Tanaka, 1998; Gauthier, Curran, Curby, & Collins, 2003; Gauthier & Tarr, 2002; Hole, 1994; Hole, George, & Dunsmore, 1999; Le Grand, Mondloch, Maurer, & Brent, 2004; Robbins & McKone, 2007; Young et al., 1987) cannot be analyzed from the perspective of GRT. Specifically, GRT analyses require judgments about all parts of the stimulus on every trial.[1] Researchers use a different paradigm, the complete identification task, to draw conclusions about perceptual or decisional loci of holistic effects.

One goal in this work is to empirically and theoretically relate performance on the complete identification task, which is required for GRT analyses, to performance on the selective attention task, which is used throughout the face literature. But these two tasks differ fundamentally in their attentional demands, as well as in how they are typically analyzed. Both tasks (see Figure 3) are variations on the composite face paradigm (Young et al., 1987). Participants study a face composite made of the top of one face and the bottom of another face. After a brief delay, they see a second composite face, in which the top and bottom can each be the same as or different from the relevant half of the study face. Participants are asked to make same–different judgments on the top or bottom half or on both halves. It is the nature of this testing that varies across tasks.

In the complete identification task, participants judge the same–different status of both parts on every trial (e.g., Kadlec & Hicks, 1998; Thomas, 2001a, 2001b; Townsend, Hu, & Evans, 1984; Townsend, Hu, & Kadlec, 1988; Wenger & Ingvalson, 2002, 2003). There are four possible responses: (a) same top and bottom; (b) different top, same bottom; (c) same top, different bottom; and (d) different top and bottom. It is only with this full factorial combination of same–different data that GRT violations of PI, PS, or DS can be detected. In the selective attention task, participants judge whether one cued part (e.g., top) is same or different while they ignore the irrelevant part (e.g., bottom; see, e.g., Farah et al., 1998; Gauthier et al., 2003; Gauthier & Tarr, 2002).[2] Holistic

[1] Ashby and Maddox (1994) proposed that one could use GRT to analyze response times from tasks where only one part is responded to on every trial in order to assess perceptual integrality (but see Nosofsky & Palmeri, 1997). However, these analyses rely on the assumption that DS holds. Given that previous work with faces has shown violations of DS (Wenger & Ingvalson, 2002, 2003) and given our interest in measuring whether both perceptual and decisional effects are occurring, we do not consider this approach here.

[2] At least two versions of the selective attention task have been used in the face processing literature (see Gauthier & Bukach, 2007). In the partial design, originally used by Young et al. (1987) and Hole (1994) and used in some more recent work (e.g., Hole et al., 1999; Le Grand et al., 2004; Robbins & McKone, 2003, 2007), the irrelevant part is always different, whereas the cued part may be same or different. By contrast, in the complete design, the cued part can be same or different, as in the partial design, but the irrelevant part can also be same or different (Farah et al., 1998; Gauthier et al., 2003; Gauthier & Tarr, 2002). Because the complete design contains a full factorial combination of trial types, it is more similar to the complete identification task than is the partial design. For this reason, we consider only the complete factorial version of the selective attention task throughout this article.

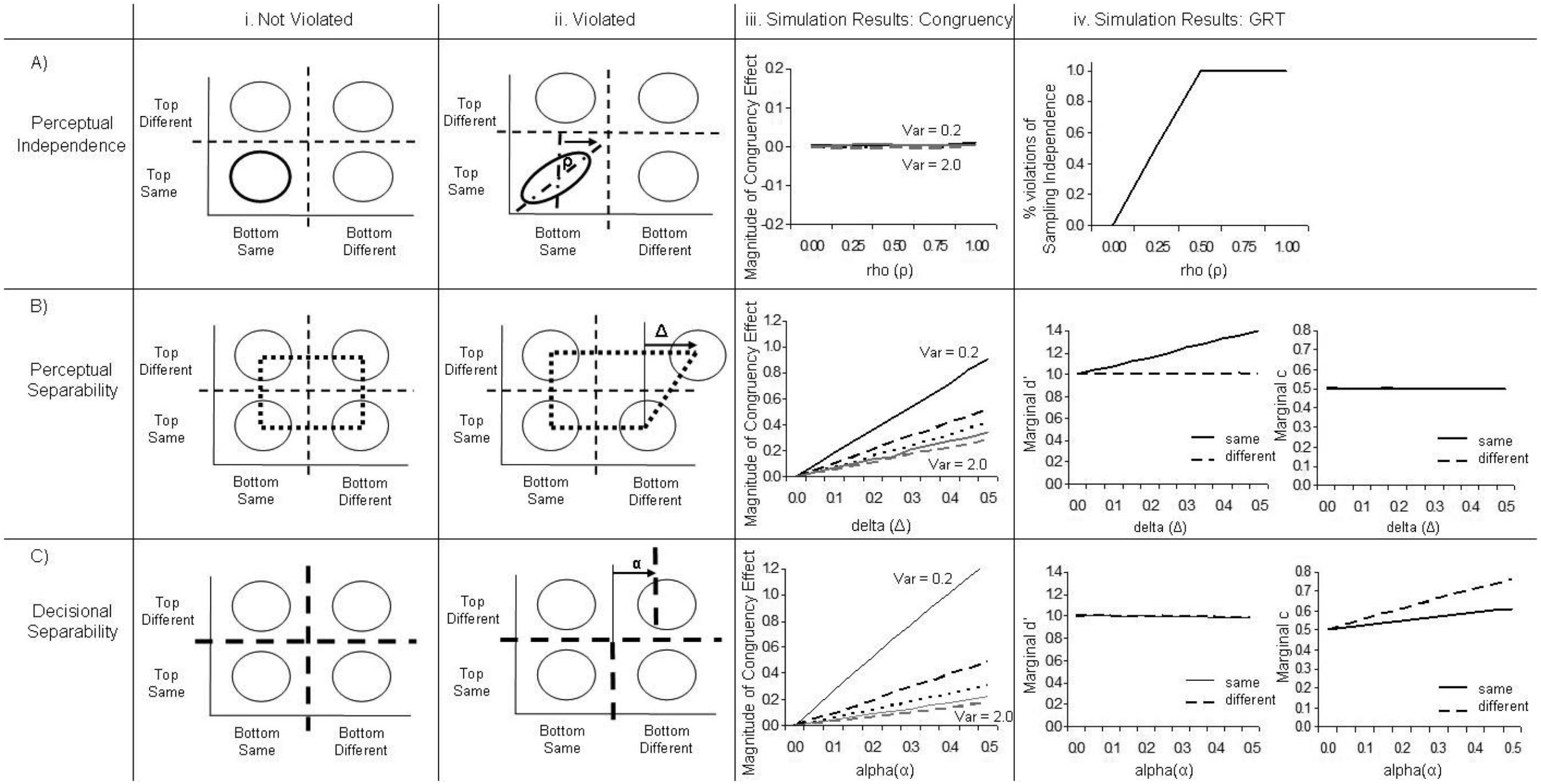

Figure 2. Schematics of the GRT constructs in their nonviolated and violated configurations, how these violations were implemented in the Monte Carlo simulations, and the results of these simulations in terms of both the magnitude of the congruency effect and the GRT analyses. Row A: (i) Perceptual independence (PI) and (ii) a violation of PI; PI is violated when there is systematic noise that is correlated between the top and bottom parts of the stimulus. PI was violated in the simulations by varying the degree of correlation (rho). In (iii), no violation of PI results in a congruency effect. In (iv), as PI is increasingly violated, the GRT analyses detect more violations of sampling independence. Row B: (i) Perceptual separability (PS) and (ii) a violation of PS; PS is violated when the location of the perceptual distribution for one part is based on the same–different status of the other part. PS was violated in the simulations by moving one distribution away from the nonviolated location (delta). In (iii), a congruency effect is observed as the distribution is moved further away. The different lines reflect different values of dimensional variance. In (iv), increasing the magnitude of the violation of PS results in larger differences in marginal d' values but no differences between marginal c values. Row C: (i) Decisional separability (DS) and (ii) a violation of DS; DS is violated when the location of the decision boundaries for one part depends on whether the other part is same or different. DS was violated in the simulations by increasing the distance between the decision boundaries for bottom decisions when the top was same or the top was different (alpha). In (iii), a congruency effect is observed as the decision boundaries are separated from each other. The different lines reflect different values of dimensional variance. In (iv), increasing the magnitude of the violation of DS results in a larger difference in marginal c values but no differences between marginal d' values. GRT = general recognition theory; Var = variance.

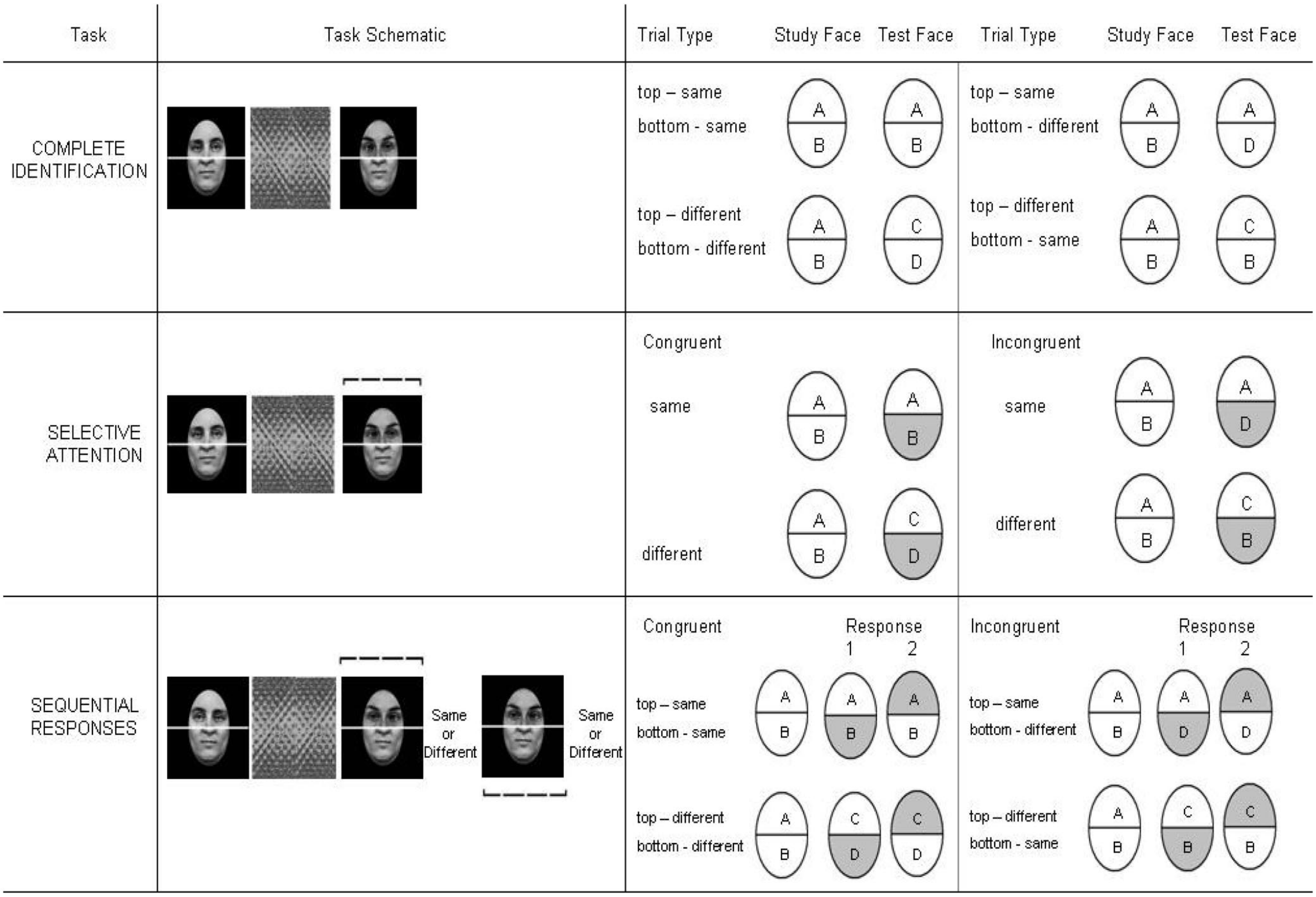

Figure 3. Schematic diagrams and trial types in the complete identification task, the selective attention task, and the sequential responses task. The relevant part is shown in white, and the irrelevant part is shown in gray. Because both parts are responded to in the complete identification task, both parts are relevant. In the selective attention task, the top part is being cued by the square bracket. Because participants do not know which part they will need to respond to until the test face appears, both parts of the study face are relevant. In the sequential responses task, first the top part is cued and then the bottom part is cued by square brackets. The result is two responses, in which first the top part is relevant and then the bottom part is relevant. Because both parts must be responded to, both parts of the study face are relevant.

processing is defined in terms of a failure of selective attention—a congruency effect—as measured by the difference in d' between congruent trials (both the top and the bottom are same or different) and incongruent trials (one part is same and one part is different).

How critical are these key differences in method and analyses? In the complete identification task, participants must divide attention across both face parts on each trial to judge the same–different status of the top and bottom parts. Observing evidence for holism is perhaps not all that surprising, as the task requires that attention be distributed across the entire face. Furthermore, evidence for decisional holism in GRT could be driven in part by the unitary response required by the task (see Wenger & Ingvalson, 2002, 2003, for evidence suggesting otherwise). By contrast, the selective attention task requires a response about only one cued part and explicitly instructs participants to ignore the uncued part. This key contrast between distributed versus focused attention motivated us to use another task to bridge these two approaches.

In this study, we used a sequential responses task (see Figure 3), in which participants made two same–different responses on every trial; for example, they might first be cued to the top part and then be cued to the bottom part. Note that the first cued judgment is identical to the single response recorded in the selective attention task. This task allows us to compare the first response in the sequential responses task with the single response in the selective attention task. Although the nature of the sequential responses is different from that of the unitary response made in the complete identification task, by the end of every trial, we have sufficient data to analyze the results in terms of GRT constructs. By using a task that has the selective attention aspects of the selective attention task but that solicits all the responses necessary for GRT analyses, as in the complete identification task, we are able to distinguish between perceptual and decisional holism and to provide an empirical link to the larger face literature. There are precedents for the sequential responses design in previous work

(Kadlec & Hicks, 1998; Townsend et al., 1984, 1988), but, to our knowledge, none of these studies has addressed questions of holism in face processing.

Monte Carlo Simulations

One natural question is whether the additional data are actually necessary. For example, if congruency effects emerged only when there were violations of DS and not when there were violations of PS or PI, GRT analyses would reveal no more than congruency effects. However, if congruency effects arose due to violations of any of these constructs, GRT analyses would be a far more powerful analytic tool. As the following simulations show, GRT analyses, but not congruency effects, are capable of distinguishing between perceptual and decisional holism.

We performed Monte Carlo simulations in which PI, PS, and DS were systematically violated and examined the congruency effect that emerged. We assumed four multivariate normal distributions, one for each of the combinations of same or different top with same or different bottom. The second column of Figure 2 illustrates the violations. As shown in Panel A(ii), we simulated violations of PI by systematically varying the correlation (rho) in the covariance matrix for one distribution; a zero correlation represents no violation of PI. As shown in Panel B(ii), we simulated violations of PS by systematically moving the location of one of the four distributions (delta), thereby changing the configuration from a square with delta equal to 0, representing no violation of PS, through progressively more trapezoidal configurations for larger values of delta. As shown in Panel C(ii), we simulated violations of DS by systematically changing the decision boundary for bottom decisions on the basis of whether the top was the same or different and by using greater disparity to simulate increasing magnitudes of alpha. We investigated violations of each GRT construct independently, assuming no violation of the other two constructs. We also systematically varied the variance along the two dimensions; as the results demonstrate, this variance rescales the magnitude of the effects but has no effect on their qualitative nature.

For each set of parameters defining the multivariate distributions (rho and delta) and decision boundaries (alpha), we ran 4,000 simulated trials. On each trial, we randomly selected one of the four distributions and then randomly drew a sample from that selected distribution. The response on that trial was determined by where that randomly selected sample was located with respect to the decision boundaries. Each response could then be characterized in terms of hits (saying "same" when the part was same) or false alarms (saying "same" when the part was different). From the hits and false alarms, we calculated a d' across congruent trials (when both the top and the bottom were same or different) and a d' across incongruent trials (when one part was same and the other was different). The difference in d' for congruent and incongruent trials is defined as the magnitude of the simulated congruency effect and is shown in Figure 2(iii). Each graph plots congruency effect as a function of the parameter manipulated in that simulation. Each line represents a different value of variance.

To begin with, violations of PI did not produce a congruency effect, as shown by the overlapping flat lines in the top graph in the third column. This is surprising, given that a violation of PI, as a within-stimulus effect, is often considered to be the strongest form of holism (O'Toole et al., 2001; Wenger & Ingvalson, 2002). Changes in the magnitude of the congruency effect caused by manipulations such as alignment or inversion cannot be caused by violations of PI, as were those instantiated in these simulations. If there are violations of PI produced by experimental manipulations, these could, however, be detected with GRT analyses, as we show later.

Both violations of PS and violations of DS produce significant congruency effects, as shown by the linearly increasing functions in the middle and bottom graphs of Figure 2(iii). Congruency effects are typically interpreted as evidence for perceptual holism. However, because both perceptual and decisional loci produce significant congruency effects, one cannot use a congruency effect by itself, or the fact that congruency is influenced by manipulations such as misalignment, to distinguish between a perceptual and a decisional basis for holistic processing.

Our simulations reveal that congruency effects can be produced in several ways and so are not diagnostic with respect to perceptual versus decisional sources of holism. Can analyses from GRT more accurately recover these sources? To test this, we analyzed these simulated data using multidimensional signal detection analysis (MSDA), which derives inferences about violations of PI, PS, and DS from estimates of signal detection parameters (Kadlec & Hicks, 1998; Kadlec & Townsend, 1992).

MSDA

In the case of unidimensional SDT, there is one discriminability (d') measure and one response criterion (c). In even the simplest multidimensional case, with stimuli composed of two dimensions that can each take on one of two levels, the number of measures that need to be estimated increases dramatically. MSDA requires estimates of d' and c along each dimension, either specific to each level of the other dimension (marginal analyses) or conditionalized on the level of and response to the other dimension (conditional analyses). In addition, MSDA requires assessment of a set of equalities at the level of relative response frequencies. As a result, MSDA requires well over a dozen different statistical tests. The benefit of this requirement is the wealth of potentially convergent evidence in support of inferences regarding PI, PS, and DS. The cost is the number of results that need to be reported and summarized. We describe the most important tests in this section and focus on the most informative tests in the body of the article.

Statistical tests in MSDA are conducted at two levels: marginal and conditional. Marginal analyses compare each level of one dimension collapsed across both levels of the other dimension. The marginal tests include (a) a test for marginal response invariance and (b) tests of equivalence on marginal d' and marginal c values. The test of marginal response invariance evaluates whether the probability of correctly reporting one dimension is independent of the level of the other dimension; for example, is the probability of correctly reporting that the top is the same or different dependent on whether the bottom is the same or different? The marginal d' equivalence tests compare differences between d' or c values for each level of one dimension collapsed across both levels of the other dimension; an example is a comparison of the d' values when the bottom is the same versus when the bottom is different, collapsed across same and different tops. Equivalence of marginal d values can indicate that PS holds. Equivalence of marginal c

values can indicate that DS holds. If both PS and DS hold for the two dimensions, marginal response invariance should also hold (Kadlec & Townsend, 1992). Logically, this means that if marginal response invariance does not hold, then either PS or DS does not hold.

The conditional analyses are used to draw inferences about DS and PI. We do not fully consider the conditional analyses in this article, for two reasons. First, DS is already assessed in the marginal analyses, which are considered more reliable than are the conditional analyses, because they are based on more data (see Kadlec & Townsend, 1992). Second, when DS is violated, the conditional analyses are inconclusive with respect to PI (Ashby & Townsend, 1986).

PI can also be assessed with a test of sampling independence. Is the probability of reporting a joint event of the two stimulus components equal to the product of the marginal probabilities of reporting each stimulus component individually? For example, if sampling independence holds, the probability of reporting that the top is same and the bottom is same is equal to the probability of reporting that the top is same multiplied by the probability of reporting that the bottom is same. When DS holds, violations of sampling independence indicate violations of PI.

In our simulations and experiments, we conducted MSDA analyses by creating a 4×4 confusion matrix, with each test stimulus type (top same/top same, top same/bottom different, bottom different/top same, bottom different/top different) crossed with each possible response (top same/bottom same, top same/bottom different, bottom different/top same, bottom different/top different). In MSDA, the data matrices are analyzed with a computer program that implements MSDA logic (MSDA2; Kadlec, 1995, 1999; available at http://web.uvic.ca/psyc/); the output is marginal and conditional signal detection measures, along with a set of summary measures for marginal response invariance and sampling independence. We used these measures to guide inferences regarding possible violations of PI, PS, or DS. Within these analyses, we used z tests to test the equivalence of marginal response invariance and the sampling independence probability values. Equivalence of signal detection parameters was tested with nonparametric tests (Grier, 1971). See Kadlec and Townsend (1992) for additional information about MSDA and the truth tables used to make conclusions about PI, PS, and DS on the basis of these statistical tests.[3]

We illustrate MSDA by analyzing our Monte Carlo simulated data with respect to violations of PI, PS, and DS (see Figure 2, fourth column). For simplicity, we show only analyses for one value of variance ($\sigma = 1$); this simplification merely rescales the findings qualitatively. In the top panel, we plot the percentage of times that sampling independence was violated as a function of the magnitude of the simulated violation of PI (rho). As DS is not violated in this simulation, these systematic violations of sampling independence reveal the systematic violations of PI.

Critically, we are most interested in whether MSDA can discriminate violations of PS from violations of DS, as we showed earlier that both types of violation can lead to congruency effects. For simulated violations of both PS and DS, we plot marginal d' and marginal c values on the basis of the same–different status of the other part against the systematically manipulated parameter for that simulation (delta for PS and alpha for DS). As shown in the middle graphs, as PS is violated, differences in marginal d' values become larger, whereas marginal c values are indistinguishable. In contrast, as shown in the bottom graphs, as DS is violated, differences in marginal c values become larger, whereas marginal d' values are indistinguishable. Thus, MSDA reveals different kinds of violations that cannot be distinguished on the basis of congruency effects alone.

Overview of Experiments

We use the sequential responses task to relate the focused attention approach of the selective attention task, which is commonly used in the face literature, to the divided attention approach of the complete identification task, which allows GRT analyses. But we need first to demonstrate that the sequential responses task does not produce results that differ qualitatively from those produced by the two tasks it aims to relate theoretically and empirically.

Research using the selective attention task has consistently found that performance is worse on incongruent trials than it is on congruent trials (e.g., Farah et al., 1998; Gauthier et al., 2003; Richler, Tanaka, Brown, & Gauthier, 2007). Despite explicit instructions that participants should ignore the uncued part, participant performance is significantly affected by the same–different status of the uncued part; the typical interpretation of this result is that participants have difficulty comparing individual face parts, because they perceive the face as an integrated whole. In Experiment 1A, we compared the first response in the sequential responses task with the single response in the selective attention task to determine whether selectively attending to a single part is more difficult in the context of a task in which both parts are ultimately attended to in each trial.

Using a complete identification task, Wenger and Ingvalson (2002) analyzed their data with respect to GRT constructs and found consistent violations of DS, with limited evidence for violations of either PI or PS (see also Wenger & Ingvalson, 2003). Thus, effects that had been attributed to perceptual representations (e.g., Farah et al., 1998) were seen instead as being due to shifts in decisional criteria. In Experiment 1B, we compared the sequential responses task and the complete identification task with respect to differential patterns of violations of PI, PS, and DS. We specifically asked whether responding to each face part separately in the sequential responses task creates patterns of GRT violations different than those created by making a single unified response in the complete identification task.

In addition to comparing the three tasks with respect to violations of GRT constructs, the first experiment attempts to replicate the previous findings by Wenger and Ingvalson (2002, 2003) for violations of DS during face recognition, which are an indication of a decisional locus of holistic processing. But the second experiment goes beyond this previous research by systematically manipulating the alignment of the test face, a manipulation known to disrupt holistic processing (Boutet et al., 2002; Richler et al., 2007;

[3] MSDA is not the only approach for guiding inferences about GRT violations. Parametric model fitting has also been used (e.g., Ashby & Lee, 1991; Maddox, 2001; Maddox & Bogdanov, 2000; Thomas, 2001b). Unfortunately, there are not enough degrees of freedom in the data from the face composite paradigm for us to take this approach. Fortunately, parametric model fitting yields outcomes that are quite similar to those from MSDA (Copeland & Wenger, 2006).

Young et al., 1987). On the basis of previous work, we expected systematic misalignment of the test face would systematically decrease the magnitude of the congruency effect. The key question is whether and how misalignment would affect the magnitude of the various GRT violations. Specifically, if holistic processing has a decisional locus, manipulations known to systematically decrease the magnitude of holistic effects should systematically decrease the magnitude of the violations of DS.

We should note that complete identification tasks are often run with a small number of observers, with each participant being tested on a large number of trials over many sessions (e.g., Thomas, 2001a, 2001b). In the experiments reported here, we combined the data from many participants who had completed a smaller number of trials into a single data matrix (see also Wenger & Ingvalson, 2002, 2003). Because averaging across participants can lead to a mischaracterization of performance by individual participants (e.g., Ashby, Maddox, & Lee, 1994), we also conducted versions of Experiments 1B and 2, in which we tested a few participants over many sessions and analyzed data from each participant separately. Results from these individual participant experiments are not reported in full here, because they produced the same results as did their single-session counterparts.

Experiment 1

In Experiment 1, we aimed to validate the use of the sequential responses task as a bridge between the selective attention task and the complete identification task. In Experiment 1A, we compared the congruency effect in the selective attention task, in which subjects respond to only one face part, with the congruency effect elicited by the first response in the sequential responses task, in which subjects ultimately respond to both face parts. In Experiment 1B, we compared the sequential responses task with the complete identification task.

Method

Participants. Participants were 41 undergraduate students (12 male) at Vanderbilt University who earned course credit for participation. Ages ranged from 18 to 22 years ($M = 20.1$). Twenty-two participants completed Experiment 1A and performed both the selective attention task and the sequential responses task. Nineteen participants completed Experiment 1B and were randomly assigned to either the complete identification task ($n = 10$) or the sequential responses task ($n = 9$).

Stimuli. Stimuli were constructed from 200 faces (half male, half female) from the MPI face database (Troje & Bülthoff, 1996). Faces were converted to gray scale and were cut in half to produce 200 face tops and 200 face bottoms (each 256×128 pixels). Tops and bottoms were randomly combined on every trial. A white line 3 pixels thick separated the face halves to make a clear distinction between the top and bottom of the stimulus; if anything, the presence of this line should have facilitated selective attention (Gauthier et al., 2003; Gauthier & Tarr, 2002). Thus, each face was 256×259 pixels (the additional 3 pixels in height was due to the 3-pixel white line that separated the face halves). To eliminate cues derived from the shape of the head or chin, we presented faces inside an oval within a black rectangle and surrounded the rectangles with a white background. Study and test faces were separated by a 256×259 random pattern mask.

Procedure. In the selective attention task (see Figure 3), a fixation cross appeared in the center of the screen for 500 ms at the beginning of each trial. It was followed by a study face, which was shown for 400 ms. After a 2,000-ms mask stimulus, a test face appeared with a square bracket that cued either the top or the bottom half of the face. Participants were asked to press one key with their left hand if the cued part was the same as in the study face and another key with their right hand if the cued part was different; mapping of keys to responses was kept constant across participants. The next trial began as soon as a response had been made or after 2,500 ms if no response was made.

Two versions of the sequential responses task were used. The version used in Experiment 1A was the same as that used in the selective attention task, except that after a response had been made to the cued part, the uncued part was cued by a square bracket, and the participant was asked to make a second same–different response (see Figure 3). The second response could be made immediately after the first response or after 2,500 ms if no first response had been made.

In Experiment 1B, the sequential responses task included occasional catch trials, in which the test face changed between the two responses. Participants were not aware of the frequency of these catch trials, and we inserted an 80-ms blank screen between the test faces, so those changes were not immediately obvious (Rensink, O'Regan, & Clark, 1997). Furthermore, although only 12% of the trials in the experimental block were catch trials, 50% of the trials in the practice block were catch trials, which gave the impression that these trials would happen with a greater frequency. Second, the cue for the first response appeared 500 ms before the test face did, so participants knew which part they would have to respond to first before the test face was presented. If participants took longer than 1,000 ms to make their first response, they would see a red screen and hear a tone for 2,000 ms at the end of the trial. (Participants were informed in the instructions that this indicated that their first response had been too slow.) Although participants were warned if the reaction time for their first response was greater than 1,000 ms, responses up to 2,500 ms were accepted before a time-out was instituted by the experimental program. There was no penalty for taking longer than 1,000 ms on the second response.

In the complete identification task (see Figure 3), after a 500-ms fixation cross, a study face was presented for 400 ms, followed by a 2,000-ms mask. A test face was then presented, and participants were instructed to press one key if both its top and its bottom were the same as those in the study face, another key if the top was different and the bottom was same, a third key if the top was same and the bottom was different, and a fourth key if both the top and the bottom were different. Two of these responses (top same/bottom same, top different/bottom same) were made with the left hand, and the other two responses (top same/bottom different, top different/bottom different) were made with the right hand. The response-key mapping appeared on the screen during the test phase and was the same for all participants. Participants had a maximum of 5,000 ms to make their response. If no response was made in that time, the experiment continued to the next trial.

Participants in Experiment 1A completed two tasks, selective attention and sequential responses. The two tasks were blocked (400 trials per block), and the order was counterbalanced. Participants in Experiment 1B completed 720 trials of either the sequential responses task or the complete identification task. For all tasks, 16 practice trials preceded the experimental block. Time-outs were relatively rare in all tasks (less than 2.5%).

Results and Discussion

Congruency effects. Performance was measured by discriminability (d') for congruent and incongruent trials for the selective attention task. Both responses of the sequential responses task from Experiment 1A and of the sequential responses task and the complete identification task from Experiment 1B are plotted in the top row of Figure 4.

A 2 × 2 × 2 mixed-factors analysis of variance (ANOVA) was conducted on mean d' values for the selective attention task and the first response of the sequential responses task, with repeated-measures factors of task (sequential responses vs. selective attention) and congruency (congruent vs. incongruent) and a between-subjects factor of task order. The main effect of task did not reach significance, and order did not interact significantly with any factor. Importantly, there was a significant main effect of congruency, with greater discriminability for congruent versus incongruent trials, $F(1, 20) = 50.146$, $MSE = .232$, $p < .0001$, which did not interact with task.

We replicated the standard congruency effect in the selective attention task and observed a similar congruency effect in the first response of the sequential responses task. This finding gave us more confidence in using the sequential responses task to measure congruency effects in much the same way that they are measured in the bulk of the face recognition literature. But the sequential responses task also provided a second response that, when combined with the first response, allowed us to conduct more powerful GRT analyses.

We next compared the first and second responses of the sequential responses task in Experiment 1A. A 2 × 2 repeated-measures ANOVA on mean d' values, with factors of response (Response 1 vs. Response 2) and congruency (congruent vs. incongruent), revealed greater performance on congruent than on incongruent trials, $F(1, 21) = 33.912$, $MSE = .214$, $p < .0001$. The magnitude of this difference did not vary significantly between the two responses, $F(1, 21) = 2.796$, $MSE = .054$, $p = .109$. A significant congruency effect was observed for both responses.

A 2 × 2 mixed-factors ANOVA was conducted on sensitivity (d') from both tasks in Experiment 1B, with a repeated-measures factor of congruency (congruent vs. incongruent) and a between-subjects factor of task (sequential responses vs. complete identification). Performance was greater for congruent trials, $F(1, 17) = 45.005$, $MSE = .101$, $p < .0001$, compared with incongruent trials. There was a significant main effect of task, in which performance was better overall for the complete identification task, $F(1, 17) = 8.499$, $MSE = 2.129$, $p = .01$, but there was no significant Task × Congruency interaction. So, overall, performance in terms of congruency was comparable between the two tasks.

GRT analyses. We report the qualitative inferences regarding violations of PS, DS, and PI from analyses of an aggregate confusion matrix that combines data from all participants. We also ran MSDA on confusion matrices of individual participant data, extracted the marginal d' and c values, and conducted statistical tests on these values. Confusion matrices and complete MSDA output, including marginal and conditional analyses and summary measures from all the experiments, are available online (http://www.psy.vanderbilt.edu/faculty/palmeri/holistic2007/).

Marginal analyses from MSDA revealed consistent violations of PS in both versions of the sequential responses task and in the complete identification task. Statistical tests of marginal d' values (see Figure 4) were consistent with these qualitative results. A paired-samples t test revealed a significant difference in marginal d' in the sequential responses task from Experiment 1A (see Figure 4, center left panel), with better performance when the irrelevant part was same versus different, $t(21) = 5.090$, $p < .0001$. A 2 × 2 mixed-factors ANOVA on marginal d' from both tasks in Experiment 1B (see center right panel), with a within-subjects factor of status of the other part (same vs. different) and a between-subjects factor of task (complete identification vs. sequential responses), revealed a significant main effect of status of the other part, with better performance when the other part was same versus different, $F(1,17) = 23.393$, $MSE = .512$, $p < .0001$. Critically, although there was a main effect of task, such that performance was better overall for the complete identification task, $F(1,17) = 9.949$, $MSE = 2.364$, $p < .01$, there was no significant interaction between task and status of the other part.

Marginal analyses from MSDA showed consistent violations of DS in both versions of the sequential responses task and in the complete identification task. Statistical tests on marginal c values (see Figure 4) were consistent with these qualitative results. A paired-samples t test revealed a significant difference in marginal c values in the sequential responses task from Experiment 1A (see Figure 4, bottom left panel), such that participants were more likely to respond "same" when the irrelevant part was same as opposed to different, $t(21) = 6.908$, $p < .0001$. A 2 × 2 mixed-factors ANOVA on marginal c values from both tasks in Experiment 1B (see bottom right panel), with a within-subjects factor of the same–different status of the other part (same vs. different) and a between-subjects factor of task (complete identification vs. sequential responses), revealed a significant main effect of status of the other part, such that participants were more likely to respond "same" when the other part was same versus different, $F(1,17) = 61.476$, $MSE = 1.998$, $p < .0001$. Although there was a main effect of task, such that participants were more likely to respond "same" in the complete identification task, $F(1,17) = 11.538$, $MSE = 1.192$, $p < .01$, there was no significant interaction between task and status of the other part.

Conditional analyses from MSDA revealed no conclusive violations of PI in any task. The complete identification task and the sequential responses task produced the same pattern of results: violations of PS and DS but no violations of PI. Although overall levels of performance and response criteria differed a bit between the two tasks, holistic processing is assessed by differences in the signal detection parameters that are based on either congruency or the same–different status of the other part. The fact that the magnitude of these differences does not differ between the two tasks suggests that using an independent feature-report procedure (sequential responses) to obtain judgments on each dimension, as opposed to assigning a unique response to each combination of dimension levels (complete identification), does not affect the outcome of the inferences regarding holistic processing.

Like Wenger and Ingvalson (2002, 2003), we found consistent violations of DS but not of PI. However, unlike Wenger and Ingvalson, whose 2002 study showed only limited evidence for

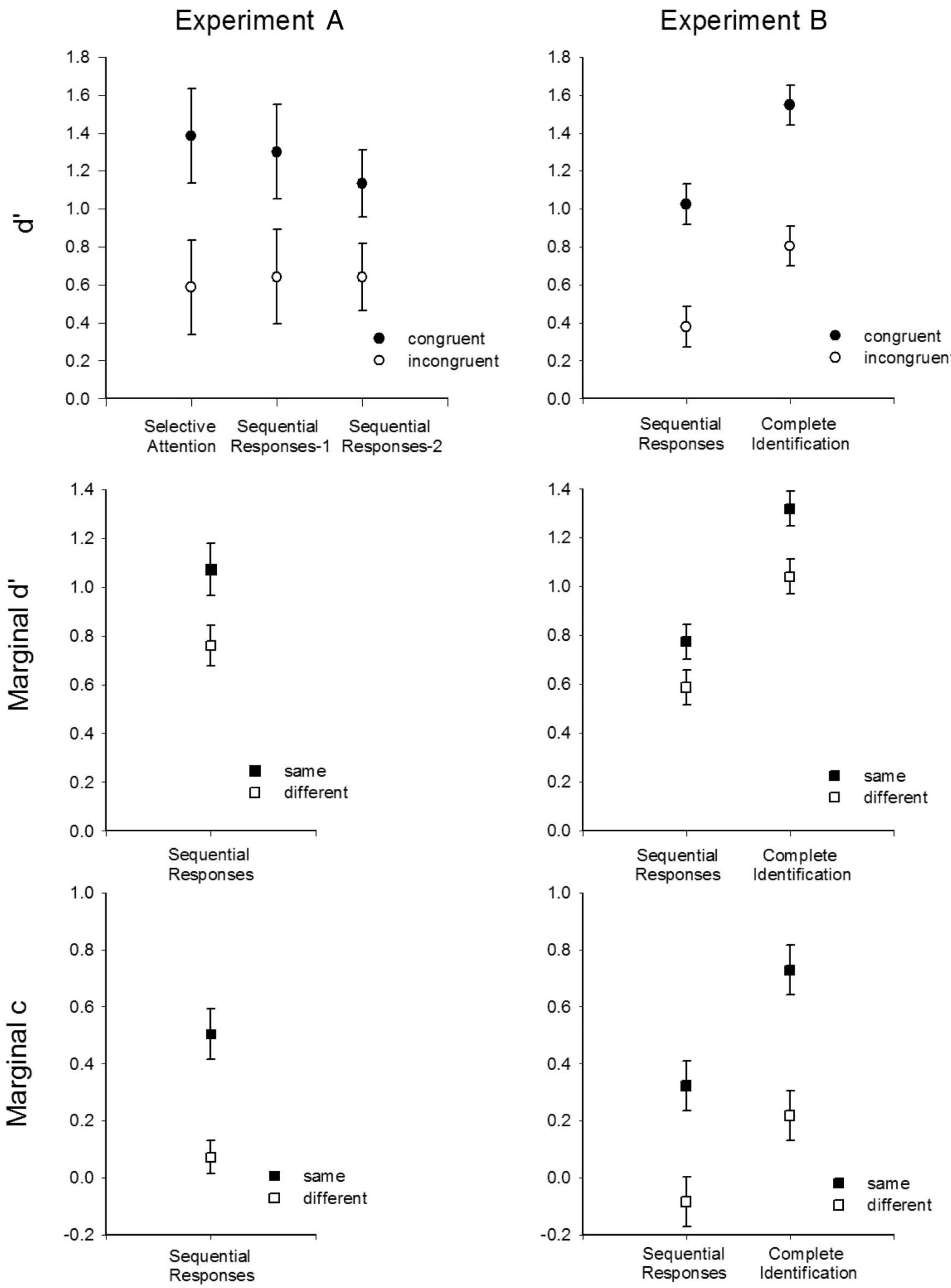

Figure 4. Performance in the selective attention and the sequential responses tasks in Experiment 1A (left column) and in the complete identification and the sequential responses tasks in Experiment 1B (right column). The top row shows discriminability (d') for congruent and incongruent trials. The middle row shows marginal d' values, which determine violations of PS. PS is violated when there is a significant difference in d' when the other part is same compared with when the other part is different. The bottom row shows marginal c values, which determine violations of DS. DS is violated when there is a significant difference in c when the other part is same compared with when the other part is different. Error bars show 95% confidence intervals of within-subjects effects.

violations of PS and whose 2003 study showed no violations of PS whatsoever, we observed consistent violations of PS in both tasks. This finding could perhaps be attributed to differences in the tasks or the stimuli: For example, Wenger and Ingvalson varied retention interval, but we did not; also, we varied entire face halves, whereas Wenger and Ingvalson had more subtle changes in eye or mouth position. For present purposes, however, a critical result is our documenting that variations in performance on this task reflect both perceptual and decisional factors. In the following section, we show that the measures that reveal violations of DS, but not of PS, vary with the magnitude of the observed congruency effect.

Experiment 2

In Experiment 1, we observed violations of both DS and PS but no violations of PI. Our earlier Monte Carlo simulations showed that violations of either DS or PS could give rise to a congruency effect. Although previous GRT results have shown evidence for violations of DS, it has not been shown whether violations of DS or PS are systematically affected by manipulations known to disrupt holistic processing. To determine whether variations in the magnitude of the congruency effect could be related to systematic violations of DS or PS, we manipulated the alignment of the test face halves and thus effectively manipulated the magnitude of holistic processing (Boutet et al., 2002; Richler et al., 2007). Would this manipulation lead to systematic variations in the measured violations of either DS or PS or of both?

Method

Participants. One group of participants ($n = 19$, 2 male, mean age 19.17 years) completed the sequential responses task for course credit. A second group of participants ($n = 22$, 8 male, mean age 25.09 years) completed the complete identification task in exchange for $12.[4]

Stimuli. Misaligned stimuli were adapted from stimuli in Experiment 1. For misaligned test faces, the top part was moved to the right and the bottom part was moved to the left, such that the edge of one part fell in the center of the other part. For very misaligned test faces, the top and bottom parts did not overlap, such that the edge of one part fell at the other edge of the other part (see Figure 5 for examples). In the sequential responses task, the locations of the cues were adjusted, so that they appeared immediately above or below the face top or bottom, respectively.

Procedure. The complete identification task was the same as the one we used in Experiment 1. For the sequential responses task, we used the version of this task from Experiment 1A. In both tasks, the study face was always aligned, but the test face was aligned, misaligned, or very misaligned. A 16-trial practice block preceded the 720-trial experimental block, which contained an equal number of aligned, misaligned, and very misaligned test faces.

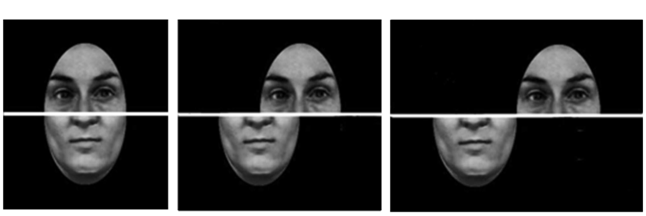

Figure 5. Examples of aligned, misaligned, and very misaligned faces used as stimuli in Experiment 2.

Results and Discussion

Performance, as measured by the congruency effect (d') and by marginal d' and marginal c values from MSDA as a function of alignment for both tasks, is plotted in Figure 6.

Congruency effect: Sequential responses task. A 3×2 repeated-measures ANOVA was conducted with factors of misalignment (aligned vs. misaligned vs. very misaligned) and congruency (congruent vs. incongruent). Performance was greater for congruent than for incongruent trials, $F(1, 18) = 24.394$, $MSE = .636$, $p < .0001$, and there was a significant Congruency $\times$ Misalignment interaction, $F(2, 36) = 16.412$, $MSE = .223$, $p < .0001$, such that the magnitude of the difference between congruent and incongruent trials decreased with misalignment.

Congruency effect: Complete identification task. A similar 3×2 repeated-measures ANOVA was conducted on these data. Performance was significantly greater for congruent than for incongruent trials, $F(1, 21) = 44.432$, $MSE = .170$, $p < .0001$, and there was a significant Congruency $\times$ Misalignment interaction, $F(2, 42) = 11.697$, $MSE = .051$, $p < .0001$, such that the magnitude of this difference between congruent and incongruent trials decreased with misalignment. There was also a significant main effect of misalignment, $F(2, 42) = 3.310$, $MSE = .025$, $p = .046$, with greater performance for aligned versus misaligned or very misaligned configurations. So, for both tasks, misalignment of the test face decreased the magnitude of the congruency effect, as expected.

GRT analyses: Sequential responses task. Analyses of the aggregate confusion matrix by MSDA revealed consistent violations of PS for aligned faces and, to a lesser degree, for misaligned faces (violated in one of two statistical tests) but not for very misaligned faces. A 3×2 repeated-measures ANOVA of marginal d' values, with factors of misalignment (aligned vs. misaligned vs. very misaligned) and status of the other part (same vs. different), revealed a significant main effect of status of the other part, $F(1, 18) = 11.803$, $MSE = .035$, $p < .01$, such that discriminability was greater when the other part was same versus different. We also observed a significant interaction of Misalignment $\times$ Status, $F(2, 36) = 3.580$, $MSE = .043$, $p < .05$, such that differences in discriminability based on status of the other part decreased with misalignment.

The marginal analyses in MSDA revealed that DS was violated across all misalignment conditions. A 3×2 repeated-measures

[4] In Experiment 2, we originally tested only participants in the sequential responses task. After completing this experiment, we decided that we should compare the congruency effects and the MSDA results for the sequential responses task with a version of the complete identification task, especially given that, in Experiment 1, we had observed violations of PS with the complete identification task, whereas Wenger and Ingvalson (2002, 2003) had not. As such, participants were not randomly assigned to task, but, for ease of exposition, we report the results as if the two tasks had been conducted within a single experiment. Because we report no between-subjects analyses of task, the analyses can be treated as if they were two separate experiments, which indeed they were.

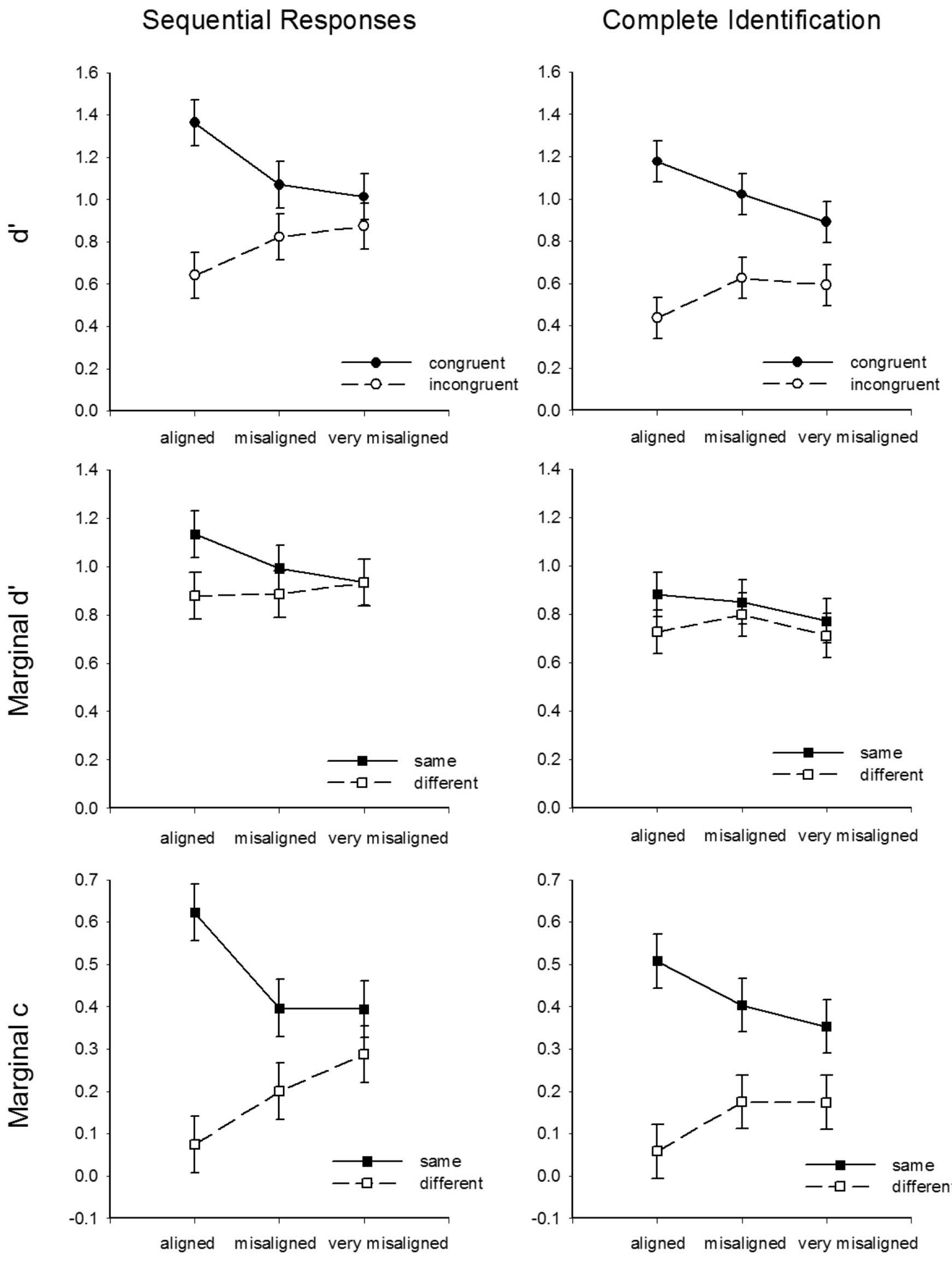

Figure 6. Performance in the sequential responses task (left column) and the complete identification task (right column) in Experiment 2 for aligned, misaligned, and very misaligned test stimuli. The top row shows discriminability (d') for congruent and incongruent trials. The middle row shows marginal d' values, which determine violations of PS. PS is violated when there is a significant difference in d' when the other part is same compared with when the other part is different. The bottom row shows marginal c values, which determine violations of DS. DS is violated when there is a significant difference in c when the other part is same compared with when the other part is different. Error bars show 95% confidence intervals of within-subjects effects.

ANOVA of marginal c values, with repeated measures of misalignment (aligned vs. misaligned vs. very misaligned) and status of the other part (same vs. different), revealed a significant main effect of status of the other part, $F(1, 18) = 49.019$, $MSE = .047$, $p < .0001$, such that participants were more likely to say "same" when the other part was same versus different. Most critically, there was a significant interaction of Misalignment × Status, $F(2, 36) = 25.088$, $MSE = .021$, $p < .0001$, such that the difference between marginal c values based on status of the other part decreased with misalignment. There were no conclusive violations of PI.

GRT analyses: Complete identification task. Analysis of the aggregate confusion matrix revealed that PS was violated only when the test face was aligned and in only one of two statistical tests. A 3 × 2 repeated-measures ANOVA of marginal d' values, with factors of misalignment (aligned vs. misaligned vs. very misaligned) and status of the other part (same vs. different), revealed significant main effects of misalignment, $F(2, 42) = 3.297$, $MSE = .025$, $p < .05$, with greater d' for aligned than for misaligned or very misaligned faces, and of status of the other part, $F(1, 21) = 8.309$, $MSE = .032$, $p < .01$, with greater d' when the other part was same. There was no significant interaction between misalignment and status of the other part.

Marginal analyses showed consistent violations of DS across all misalignment conditions. A 3 × 2 repeated-measures ANOVA of marginal c values, with repeated measures of misalignment (aligned vs. misaligned vs. very misaligned) and status of the other part (same vs. different), revealed a significant main effect of status of the other part, $F(1, 21) = 39.719$, $MSE = .067$, $p < .0001$, such that participants were more likely to respond "same" when the other part was same versus different, and a significant Other Part × Misalignment interaction, $F(2, 42) = 10.746$, $MSE = .022$, $p < .0001$, such that the difference in c when the other part was same versus different decreased with misalignment. There were no conclusive violations of PI.

GRT analyses: Summary. Although both tasks showed violations of PS and DS across alignment conditions, the marginal c values we used to determine violations of DS were significantly affected by misalignment, whereas the marginal d' values we used to determine violations of PS were less so. This finding suggests that changes in the magnitude of the congruency effect that were due to misalignment are linked to a decisional component as well as to, and perhaps instead of, a perceptual component of holistic processing. These results also illustrate that though qualitative conclusions about violations of GRT constructs are interesting and useful, they may be too coarse to pick up on quantitative changes in holistic processing across conditions of an experiment. These changes can, however, be detected by statistical estimation of the signal detection parameters that one uses to make qualitative conclusions in MSDA (see also Wenger & Ingvalson, 2003).

General Discussion

We examined holistic processing of faces in an interrelated set of experimental paradigms. In the selective attention task, participants are told to attend selectively to one face part, whereas in the complete identification task, attention is divided between both face parts. What is the relationship between the data obtained in these tasks? To find out, we related behavior in these tasks with behavior in a sequential responses task. Our ultimate goal was to relate violations of GRT constructs, which can be obtained with complete identification and sequential responses, to the observed congruency effect, which can be obtained with selective attention and sequential responses.

In Experiment 1, we replicated earlier work (e.g., Cheung, Richler, Palmeri, & Gauthier, 2007; Farah et al., 1998; Gauthier et al., 2003; Gauthier & Tarr, 1997; Richler et al., 2007) and found that, in the selective attention task, discriminability (d') on trials when both face halves are same or both halves are different (congruent trials) is higher than is discriminability on trials when one half is same and the other half is different (incongruent trials). More important, we found the same congruency effect for both responses in the sequential responses task. In Experiment 2, we also showed that the magnitude of the congruency effect in the sequential responses task decreased as the top and bottom face halves were systematically misaligned, a finding that replicated previous work (see Richler et al., 2007). Results such as these have been used to argue for a perceptual locus of holistic processing of faces (e.g., Hole, 1994; Young et al., 1987): Participants cannot selectively attend to a face half, because both halves are fused during perception. Misaligning the face prevents that fusion from taking place and allows participants to ignore the irrelevant half.

Research that uses the complete identification task approaches the question of holistic processing of faces from a somewhat different theoretical angle. Rather than identifying holistic processing through a failure of selective attention, the research analyzes the entire confusion matrix of same–different data generated by this task to reveal violations of GRT constructs (Ashby & Townsend, 1986). A violation of PS indicates that one face half may be perceived as more similar to or more different from the study face, depending on the same–different status of the other face half. A violation of DS indicates that the criterion for generating a same–different response for one face half depends on the same–different status of the other face half. A violation of PI indicates that variability in the perception of the same–different status of the two face halves is correlated.

In our experiments, for both the complete identification task and the sequential responses task, we observed violations of both PS and DS but very little evidence for violations of PI. The violations of DS replicate the results of Wenger and Ingvalson (2002, 2003). The violations of PS we obtained were not obtained consistently by Wenger and Ingvalson (2002, 2003).[5] In sum, although there may be a perceptual locus for congruency effects, as revealed by violations of PS, we found significant decisional effects, as revealed by violations of DS. Importantly, we showed that these

[5] In our experiments, we were fairly liberal with respect to participant inclusion criteria. However, we performed further analyses using an additional set of parametric and nonparametric estimates of marginal and conditional sensitivity and bias, in which we adopted a more conservative inclusion criteria: Specifically, participants with marginal d' values that were not reliably different from 0 were excluded from further analyses. This approach was similar to the criteria used by Wenger and Ingvalson (2002, 2003). These analyses revealed even larger and more consistent differences in marginal c values (indicating violations of DS) and smaller and less consistent differences in marginal d' values (indicating violations of PS). That is, these analyses showed stronger evidence for decisional effects and weaker evidence for perceptual effects than did the findings reported here.

violations occur in a task that requires sequentially focused attention; thus, they are not related to task demands that require dividing attention between both parts of the stimulus (in which case decisional holism might be less surprising).

Critically, this question of a decisional versus perceptual locus of holistic processing cannot be revealed by using the selective attention task most commonly used in the face recognition literature. Our Monte Carlo simulations demonstrate that there are multiple ways of obtaining congruency effects from the perspective of GRT. Without the data provided by same–different judgments about both face parts, we cannot know which of these possibilities—specifically, violations of PS or DS in any combinations—might be operative. The detailed analyses provided by MSDA are needed, so we can identify and distinguish between these possibilities.

In Experiment 2, we systematically misaligned the top and bottom halves of the face. Most critically, we examined quantitatively how misalignment affected the measures we used to assess violations of DS and PS and of marginal c and marginal d'. The systematic decrease in the magnitude of the congruency effect was accompanied by a systematic decrease in the differences between marginal c values, which is indicative of violations of DS. This finding is important, because prior research simply demonstrated violations of DS (Wenger & Ingvalson, 2002, 2003), whereas the present study shows that the magnitude of these violations goes hand in hand with the magnitude of the congruency effect. As such, the congruency measure of holistic processing based on selective attention may be strongly related to shifts in decisional criteria. However, as shown by our simulations, one needs more than the congruency effect to draw this inference.

One common assumption in the face literature is that holistic processing of a face during encoding creates a holistic representation of the study face in memory, in which the face parts are not explicitly represented but the entire face is represented as a whole or gestalt (Diamond & Carey, 1986; Lewis & Glenister, 2003; Murray, Young, & Rhodes, 2000; Tanaka & Sengco, 1997). However, the limited violations of PS and PI observed by Wenger and Ingvalson (2002, 2003) and our inconsistent violations of PS and limited violations of PI challenge a holistic encoding hypothesis. Instead, holistic effects may occur during the retrieval of the study face or during the comparative process. This finding is consistent with that of Richler et al. (2007), who showed that the congruency effect was larger for aligned versus misaligned test faces, regardless of whether the study face was aligned or misaligned; configural manipulations reduced the congruency effect when applied at test but not when applied at study.

Although we found some support for perceptual factors underlying holistic effects, the most consistent evidence was for a decisional locus of holistic processing. The specific conclusions regarding violations of DS depend on whether the data are consistent with the assumptions of the multivariate model, the statistical quality of the estimators of the parameters of that model, and the strength of the logic relating the theoretical constructs to the observable data (and vice versa). These factors, of course, are issues in relating any theoretical model to data. That said, the differences in marginal d' and marginal c values are transformations of the raw data in the confusion matrices. So, regardless of whether GRT is the appropriate model, any theory of face recognition needs to account for these specific patterns of response frequencies.

But what does it mean to suggest that holistic effects emerge from decisional and not perceptual factors? Although our data suggest that holistic effects have a decisional basis, our analyses cannot speak to the mechanisms producing these effects, any more than classic univariate signal detection theory can speak to the mechanisms of perceptual or cognitive processes. Our results suggest that holistic effects in face processing are decisional but cannot tell how they might be decisional. Furthermore, signal detection theory does not distinguish between any potential unconscious versus conscious influences or between task-related versus automatic influences on decision criteria (but see Snodgrass, 2002).

Whatever the theory of face processing, some aspects of decision criteria placement may be inherently top down, but other aspects could be influenced by the many years of experience individuals have had with faces. It is likely that, because of our extensive experience with faces, we have developed a deeply ingrained assumption that face parts change together: That is, when the top half of a face is different from a previously studied face, we assume it is likely from a different person, so the bottom half of the face should be different as well. This expectation that face parts change together may be so strong that it cannot be overridden during an experiment, even when participants are explicitly told to selectively attend to one part of the face while ignoring the other part. Even though participants are able to explicitly represent both parts of the face, their expectation about faces, gained through experience, creates a decisional bias that affects the percept. Decisional biases may be gained through experience, in much the same way that perceptual representations are gained through experience (Diamond & Carey, 1986; Gauthier & Tarr, 2002). Although we normally think of decisional components of processing as being highly malleable, they may become deeply ingrained and be relatively immune to task influences, especially in domains of expertise. Further empirical and theoretical work is needed to fully expand this speculation into a viable theoretical proposition.

References

Ashby, F. G., & Lee, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General, 120,* 150–172.

Ashby, F. G., & Maddox, W. T. (1994). A response time theory of separability and integrality in speeded classification. *Journal of Mathematical Psychology, 38,* 423–466.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional scaling and similarity choice model. *Psychological Sciences, 5,* 144–151.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93,* 154–179.

Boutet, I., Gentes-Hawn, A., & Chaudhuri, A. (2002). The influence of attention on holistic face encoding. *Cognition, 84,* 321–341.

Cheung, O. S., Richler, J. J., Palmeri, T. J., & Gauthier, I. (2007). *Revisiting the role of spatial frequencies in the holistic processing of faces.* Manuscript submitted for publication.

Copeland, A. M., & Wenger, M. J. (2006). An investigation of perceptual and decisional influences on the perception of hierarchical forms. *Perception, 35,* 511–529.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117.

Farah, M. J., & Tanaka, J. W. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology, 46,* 225–245.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. W. (1998). What is "special" about face perception? *Psychological Review, 105,* 482–498.

Gauthier, I., & Bukach, C. (2007). Should we reject the expertise hypothesis? *Cognition, 103,* 322–330.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6,* 428–432.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring face recognition mechanisms. *Vision Research, 37,* 1673–1682.

Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 431–446.

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

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin, 75,* 424–429.

Hole, G. J. (1994). Configurational factors in the perception of unfamiliar faces. *Perception, 23,* 65–74.

Hole, G. J., George, P. A., & Dunsmore, V. (1999). Evidence for holistic processing of faces viewed as photographic negatives. *Perception, 28,* 341–359.

Ingvalson, E. M., & Wenger, M. J. (2005). A strong test of the dual-mode hypothesis. *Perception & Psychophysics, 67,* 14–35.

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, NJ: Erlbaum.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science, 15,* 762–768.

Lewis, M. B., & Glenister, T. E. (2003). A sideways look at configural encoding: Two different effects of face rotation. *Perception, 32,* 7–14.

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111,* 835–863.

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.

Murray, J., Yong, E., & Rhodes, G. (2000). Revisiting the perception of upside-down faces. *Psychological Science, 11,* 492–496.

Nosofsky, R. M., & Palmeri, T. J. (1997). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception & Psychophysics, 59,* 1027–1048.

O'Toole, A., Wenger, M. J., & Townsend, J. T. (2001). Quantitative models of perceiving and remembering faces: Precedents and possibilities. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives of facial cognition: Contexts and challenges* (pp. 1–38). Mahwah, NJ: Erlbaum.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). The need for attention to perceive changes in scenes. *Psychological Science, 8,* 368–373.

Richler, J. J., Tanaka, J. W., Brown, D. D., & Gauthier, I. (2007). *Why does selective attention to part fail in face processing?* Manuscript submitted for publication.

Robbins, R., & McKone, E. (2003). Can holistic processing be learned for inverted faces? *Cognition, 88,* 79–107.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioral tasks. *Cognition, 103,* 34–79.

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396.

Snodgrass, M. (2002). Disambiguating conscious and unconscious influences: Do exclusion paradigms demonstrate unconscious perception? *American Journal of Psychology, 115,* 545–580.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition, 25,* 583–592.

Thomas, R. D. (2001a). Characterizing perceptual interactions in face identification using multidimensional signal detection theory. In M. J. Wenger & J. T. Townsend (Eds.), *Computational, geometric, and process perspectives on facial cognition: Contexts and challenges* (pp. 193–228). Mahwah, NJ: Erlbaum.

Thomas, R. D. (2001b). Perceptual interactions of facial dimensions in speeded classification and identification. *Perception & Psychophysics, 63,* 625–650.

Townsend, J. T., Hu, G. G., & Evans, R. J. (1984). Modeling feature perception in brief displays with evidence of positive interdependencies. *Perception & Psychophysics, 36,* 35–49.

Townsend, J. T., Hu, G. G., & Kadlec, H. (1988). Feature sensitivity, bias, and interdependencies as a function of energy and payoffs. *Perception & Psychophysics, 43,* 575–591.

Troje, N., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36,* 1761–1771.

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.

Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 755–779.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16,* 747–759.

Received March 7, 2007
Revision received September 11, 2007
Accepted October 18, 2007 ■