Linear Theory, Dimensional Theory, and the Face-Inversion Effect

Geoffrey R. Loftus, Martin A. Oberg, and Allyss M. Dillon
University of Washington

We contrast 2 theories within whose context problems are conceptualized and data interpreted. By traditional *linear theory*, a dependent variable is the sum of main-effect and interaction terms. By *dimensional theory*, independent variables yield values on internal dimensions that in turn determine performance. We frame our arguments within an investigation of the *face-inversion effect*—the greater processing disadvantage of inverting faces compared with non-faces. We report data from 3 simulations and 3 experiments wherein faces or non-faces are studied upright or inverted in a recognition procedure. The simulations demonstrate that (a) critical conclusions depend on which theory is used to interpret data and (b) dimensional theory is the more flexible and consistent in identifying underlying psychological structures, because dimensional theory subsumes linear theory as a special case. The experiments demonstrate that by dimensional theory, there is no face-inversion effect for unfamiliar faces but a clear face-inversion effect for celebrity faces.

Understanding the implications of any experimental outcome requires a fundamental *quantitative theory*, within whose context the numbers constituting the data may be transformed into conclusions about the underlying processes that generated the data. This article is about the general consequences for eventual conclusions of deciding to use one such theory or another and about the specific consequences for understanding a well-known phenomenon in the domain of face processing.

The article is divided into five sections. In the first section ("The Face-Inversion Effect"), we describe an extant psychological problem that serves as a vehicle for illustrating the points that we make regarding theories. In the second section ("Theories to Analyze Data"), we describe two quantitative theories: The first, traditional *linear theory*, is used almost universally within many disciplines, including psychology, whereas the second, *dimensional theory*, is considerably less known and less used. In the third section ("Simulations"), we describe three simulations, whose purpose is to illustrate some costs and benefits of interpreting data using linear and dimensional theory. In the fourth section ("Experiments"), we describe three experiments that, given the foundation we have established, allow us to make some tentative conclusions about the face-inversion effect in particular and about face processing in general. Finally, in the fifth section, our General Discussion, we

compare the two theories that we have been considering: We show formal mathematical relations between them, we comment on the advantages and disadvantages of using one versus the other as a tool for inferring the underlying processes that generated a data set, and we articulate the resulting implied conclusions about face processing.

These five sections are designed in pursuit of three interrelated goals. The first goal is to demonstrate (yet again) that traditional linear theory has severe limitations as a basis for conceptualizing problems and for analyzing and interpreting data. The second goal is to describe the usefulness of dimensional theory and to provide several kinds of tutorial information about how to use it as a basis for interpreting data. The third goal is to provide an insight and an associated quantitative theory for understanding the circumstances under which a face-inversion effect does or does not emerge.

The Face-Inversion Effect

The face-inversion effect (FIE) refers to a finding first reported by Yin (1969, 1970) and subsequently by many others (see, e.g., Bradshaw, Taylor, Patterson, & Nettleton, 1980; Diamond & Carey, 1986; Ellis & Shepherd, 1975; Leehey, Carey, Diamond, & Cahn, 1978; Phelps & Roberts, 1994; Phillips & Rawles, 1979; Valentine & Bruce, 1986) that the processing disadvantage of inverting a visual stimulus is worse for faces than for other kinds of visual stimuli (e.g., houses). The FIE is one of the primary bases for the claim that face processing is special, i.e., is qualitatively different from processing other visual stimuli: As noted in a classic review article, for example, ". . . the evidence from the effect of inversion . . . provides the most direct indication that face recognition may involve a unique process" (Valentine, 1988, p. 472). Although Valentine himself concluded otherwise, the debate is still far from settled (see, e.g., Kanwisher, McDermott, & Chun, 1997; Kanwisher, Tong, & Nakayama, 1998; Gauthier, Curran, Curby, & Collins, 2003).

In using the FIE to assess face-processing uniqueness or anything else, no one would argue that there is a *universal* FIE—that is, that a FIE emerges *whenever* faces are compared with some

other stimulus type. Rather, the question is almost invariably "under what circumstances does a FIE emerge or not emerge?" For instance, a FIE has been claimed to emerge for 10-year-old children but not 6- or 8-year-old children (Carey & Diamond, 1977), for faces learned normally but not for faces learned as parts (Farah, Tanaka, & Drain, 1995), in a recognition task but not in a matching task (Bruyer & Velge, 1981; see discussion by Valentine, 1988, p. 474), and for naïve observers but not for experts in the alternative stimulus class (Diamond & Carey, 1986). In short, the quest is usually for the effect of some other variable on the FIE, and seeking such interactions has been a staple of face-processing research. It is our hope that in the course of our discussion, we will both clarify the consequences of using one theory or another to interpret data (our major goal) and elucidate the issue of when the FIE does or does not emerge (our secondary goal).

Theories to Analyze Data

Below, we report three simulations and three analogous experiments concerning the FIE. The simulations and experiments all have the same general design, which we briefly describe here in order to provide a context for our discussion of theory. In the simulations and experiments, Faces or non-Faces (Houses or City-scapes) are shown in a study phase, followed by an old–new recognition test. At study, stimuli are shown upright or inverted for varying exposure durations. Thus, there are three independent variables: exposure duration at study, orientation at study, and stimulus type.

A comment about notation is useful here. We index duration by i, orientation by j, and stimulus type by k. In the simulations and experiments that we report, there can be an arbitrarily large number, I, of durations; hence, duration is designated d_i, where $i \in \{1, 2, \ldots, I\}$. There are always $J = 2$ orientation levels, Upright and Inverted; hence $j \in \{U, I\}$. There are always $K = 2$ stimulus types, Faces and non-Faces. Because non-Faces are mostly Houses, we find it notationally convenient to let $k \in \{F, H\}$. We use this indexing notation extensively throughout this article, it plays a central role in our discourse, and we urge the reader to memorize it.

We now discuss the two theories within whose context we will analyze and interpret the data from these experiments: linear theory and dimensional theory. For ease of discourse, we establish here the subtle but useful distinction between a structure and a theory: A *structure* is a representation of the information and information processing that characterizes some system, whereas a *theory* is a set of assumptions concerning the structure's nature that is used to guide the analysis of that structure. Below, we will discuss linear theories designed to describe linear structures and dimensional theories designed to describe dimensional structures.

Linear Theory

The vast majority of experiments in many disciplines, including psychology, are designed, analyzed, and interpreted within the context of general linear theory. Linear theory, described at least implicitly in any statistics text and explicitly in any mathematically based statistics text (e.g., Hays, 1973), holds that the dependent variable in an experiment is the sum of terms corresponding to main effects of and interactions among independent variables. In the present experiments with three independent variables, probability, p_{ijk} in the condition defined by level i of exposure duration, level j of orientation, and level k of stimulus type, is

$$p_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \lambda_{ij} + \nu_{ik} + \kappa_{jk} + \xi_{ijk} + \text{error terms},$$

(1)

where μ is the grand mean; α_i, β_j, and γ_k are main effects of exposure duration, orientation, and stimulus type, respectively; λ_{ij}, ν_{ik}, and κ_{jk} are two-way interaction effects; ξ_{ijk} is the three-way interaction effect; and "error terms" are all the error terms applicable given the particular design of the experiment. A FIE within the context of this linear theory consists of a linear structure that includes a set of nonzero, appropriately signed orientation × stimulus type interaction terms, κ_{jk}; conversely, lack of a FIE consists of a linear structure with all $\kappa_{jk} = 0$.

Although linear theory has been a standard tool in understanding myriad data sets, it has some serious disadvantages (see, e.g., Loftus, 2002, pp. 343–344). Briefly, they are as follows. First, linear theory is entirely *linear*, hence its name. Although linearity may bear an approximation to some actual psychological relations, many other such relations are decidedly nonlinear, which means that investigating them within the context of linear theory produces profoundly misleading results. A classic example of this problem involves interpretations of interactions (see, e.g., Bogartz, 1976; Loftus, 1978, 1985; Loftus & Bamber, 1990): As interpreted within the context of linear theory, nonordinal interactions observed with one dependent variable (e.g., recognition performance) can disappear or reverse with another dependent variable (e.g., d') or a theoretical construct (e.g., "memory strength") that is monotonically but nonlinearly related to the dependent variable.

The second disadvantage of linear theory is more subtle but also more insidious: Because linear theory is both seductively plausible and almost universally welcomed, it blinds investigators to alternative theories that might better elucidate underlying psychological processes. We will argue that dimensional theory includes many such theories as special cases.

Dimensional Theory

The general idea of dimensional theory is that independent variables combine at various stages into internal psychological *dimensions* that underlie performance. By determining (a) how many such dimensions are necessary to account for performance in a given situation along with (b) the nature of the mathematical functions that describe how the independent variables combine to produce the dimensional values, the underlying nature of the relevant structures can be unveiled.

Dimensional theory is related to conjoint measurement (e.g., Krantz, Luce, Suppes, & Tversky, 1971; Krantz & Tversky, 1971; Tversky & Russo, 1969), functional measurement (e.g., Anderson, 1974, 1979), multidimensional scaling (e.g., Kruskal, 1964; Shepard, 1962), the concept of integral and separable dimensions (Garner, 1974), and the concept of "mental modules" (Pinker, 1997). The incarnation of it with which we are concerned here was independently described by Bamber (1979) and Dunn and Kirsner (1988), although it was not called "dimensional theory" in either article. Dimensional theory has proven useful in illuminating various psychological phenomena, including visual displacement dis-

crimination (Palmer, 1986a, 1986b); the relation between iconic memory and visible persistence (Loftus & Irwin, 1998); the relation between confidence and accuracy in face recognition (Busey, Tunnicliff, Loftus, & Loftus, 2000); the relation between degree of original learning and forgetting rate (Loftus, 1985; Loftus & Bamber, 1990); the relations among stimulus duration, stimulus contrast, confidence, and accuracy in visual recognition (Harley, Dillon, & Loftus, 2004); and the logical, theoretical, and empirical underpinnings of the dissociation technique (Dunn & Kirsner, 1988). Dunn and James (2003) have recently offered a technique founded on dimensional theory called *signed difference analysis* and have illustrated its use in addressing three problems within cognition: First, does the "remember-know" distinction in recognition memory reflect qualitatively different cognitive states or different regions on some unidimensional scale? Second, do three measures of spatial attention—visual search, texture segregation, and location precuing—reflect the same underlying process? Third, is a "dual-route model" of dyslexia proposed by Coltheart (e.g., Coltheart, 1985) viable?

Dimensional theory is central within vision science. Two examples of its use there are as follows. The first is that of color metamers: that an indefinitely large number of independent variables, in the form of different monochromatic hues, reduce, in the form of cone quantum-catch values, to three retinal-output dimensions whose values determine color perception. The second is Bloch's law: that the two independent variables of stimulus duration and stimulus intensity combine multiplicatively into a single dimension of "total intensity" whose values determine brightness perception.

D-*Dimensional Theory*

By dimensional theory, a particular combination of M independent variables in some experiment yields D values—one on each of D internal psychological dimensions. Thus,

$$
\begin{aligned}
V_2 &= f_1(IV_1, IV_2, \ldots, IV_M)\\
V_2 &= f_2(IV_1, IV_2, \ldots, IV_M)\\
&\vdots\\
V_D &= f_D(IV_1, IV_2, \ldots, IV_M),
\end{aligned}
\tag{2}
$$

where the IV's are the independent variables, the V_d's are the values on the D dimensions, and the f_d's are unconstrained functions. The D values are then mapped to N dependent variables,

$$
\begin{aligned}
DV_1 &= g_1(V_1, V_2, \ldots, V_D)\\
DV_2 &= g_2(V_1, V_2, \ldots, V_D)\\
&\vdots\\
DV_N &= g_N(V_1, V_2, \ldots, V_D),
\end{aligned}
\tag{3}
$$

where the DV_n's are the dependent variables. When $D = 1$, that is, when there is only a single dimension, $g_1 = g$ is monotonic. When $D > 1$, the situation is somewhat more complicated, but the g_n's are still constrained in a manner described by Dunn and James (2003). Equations 2 and 3 thereby define a D-dimensional structure. If the inferred number of internal dimensions, D, is less than the number of independent variables, M, one concludes that at least two of the independent variables have lost their unique represen-

tations somewhere in the structure, when they merge into fewer dimensions.

We are by no means the first to suggest this kind of structure. The distinction between dimensional values and the relations between these values and the dependent variable has been embodied in the distinction between a *structural model* and a *measurement model* (Busemeyer & Jones, 1983) or between an *integration function* and a *response function* (Anderson, 1981). Similarly, and most recently, Dunn and James (2003) used the term *structural mapping* in reference to a set of functions relating the $V_1, \ldots, V_D$ to performance expressed in terms of an abstract, unobservable, psychological metric and the term *measurement mapping* in reference to a set of monotonic functions relating the psychological metric to observable performance. Dunn and James's characterizations have the advantage of making a very clear distinction between the central psychological assumptions of the theory (expressed by the structural mapping) and the necessary, although less interesting, relation between the performance expressed in terms of the psychological metric and performance expressed in terms of the actual dependent variable.

An illustration of a dimensional structure is found in color metamers, mentioned above. Suppose that a mixture of M monochromatic hues of differing intensities is presented to the visual system. Each hue acts as an independent variable; thus the different experimental conditions correspond to the different combinations of hue intensity. It has long been known that the sensory result of any combination of such hue intensities can be described completely by the three numbers corresponding to the three cone quantum catches engendered by that combination. Therefore, any dependent variable used to measure color perception in this kind of experiment depends only on $D = 3$ dimensions corresponding to the output of the three cone classes. This finding, embodied in the classic color-matching experiment, was pivotal in color science: It formed the basis for the trichromacy theory of color vision and laid the groundwork for the eventual discovery of cone photoreceptors.

Unidimensional Theory

The simplest dimensional theory is a *unidimensional* theory in which $D = 1$; that is, only a single dimension is required to account for a data set. An illustration of unidimensional theory is Bloch's law, mentioned above, which describes the results of an experiment in which different stimulus durations and stimulus intensities are combined. Bloch's law states that when duration is less than some threshold (around 100 ms), performance—measured, for example, by detection—depends only on the product of duration and intensity; in other words, the two independent variables, duration and intensity, combine (multiplicatively) to produce a value on the $D = 1$ dimension of "summed intensity." This means that neither the physical value of duration nor the physical value of intensity is represented within the sensory system; rather, only their product is represented. This finding was useful in understanding how the sensory system processes intensity; viz., for a certain period, about 100 ms, the system simply integrates arriving photons over time and maintains a representation of the photon sum.

A unidimensional theory is somewhat akin to a null hypothesis within a linear theory: It is a default version of the theory that in general, is tested first. However, a unidimensional theory is unlike

most null hypotheses in an important respect. Usually a null hypothesis is deliberately constructed so as to be uninteresting, and the investigator's goal is to reject it. In contrast, a unidimensional theory such as Bloch's law is generally viewed as simple and elegant: If a unidimensional theory is correct, it implies strong and interesting constraints on how some system works. Therefore, an investigator generally wants to confirm a unidimensional theory. If one must reject unidimensional theory, then one must add theoretical dimensions until the data can be accounted for.

Depending on the questions to be addressed and the experimental design, many kinds of unidimensional theories may be constructed. Here we are concerned with the FIE. Accordingly, a useful unidimensional theory to test would be one in which there is no FIE. A unidimensional theory designed to accomplish this goal is as follows. First, the two perceptual variables, duration, d, and orientation, O, combine to produce a value on a single dimension that we will somewhat arbitrarily term $S = $ "Strength"; thus,

$$S = f(d, O), \tag{4}$$

where the function f, although in principle unconstrained, would be expected to yield higher S values with longer durations, and for upright compared to inverted stimuli. A fundamental element of such a theory has been suggested by Valentine (1988), who noted that ". . . the available evidence suggests that inversion and brief presentation affect perception in similar ways" (p. 483). The critical element that Valentine does *not* suggest is that these "similar ways"—formally embodied here in Equation 4, which states that duration and orientation both operate in the common currency of Strength—are identical for Faces and non-Faces.

The next component in our unidimensional theory is the mapping from Strength to recognition performance, which is assumed to be monotonic. However, it makes little sense to assume that Face performance and non-Face performance are related to Strength in identical ways. There are many reasons to believe that the Strength-performance function might differ for the two stimulus types; for instance, the set of Faces may be more homogeneous than the set of non-Faces (or vice versa), or Faces may be more intrinsically recognizable than non-Faces (or vice versa), and so on. Thus we allow separate monotonic functions relating recognition performance to Strength, namely,

$$p = \begin{cases} m_F(S) & \text{for Faces} \\ m_H(S) & \text{for non-Faces} \end{cases}, \tag{5}$$

where p is recognition performance and m_F and m_H are monotonic functions.

We emphasize that within the context of this unidimensional theory, the mechanism affected by orientation is embodied only in the function, f, in Equation 4 that maps a duration × orientation combination to Strength. At that point, stimulus type has not yet come into play; rather the stimulus-type effect is embodied in the functions, m_F and m_H, of Equation 5, which operate only on Strength. Therefore, by this unidimensional theory, there is no FIE; that is, there is no direct connection between orientation on the one hand and stimulus type on the other. Below, we describe a formal and unambiguous prediction of this unidimensional theory.

Simulations

The simulations and experiments that we report are designed to simulate or to investigate the FIE with respect to memorization of

faces in anticipation of an old-new recognition test. As we have noted, we treat both simulations and experiments primarily as a vehicle for discussing theory. However, they are also designed to illuminate a specific aspect of face processing and of the effects of face inversion. In particular, Valentine (1988) has asserted that "the orientation of the inspection series [in a recognition procedure] does not appear to be critical [in producing a FIE]," and went on to conjecture, "Therefore it is possible that the disproportionate effect of face inversion only emerges when the task involves recognizing a face as one stored in memory" (p. 474). If true, this would suggest that face inversion does not differentially affect encoding of faces for subsequent recognition but rather affects processing only of already stored faces. Thus our experiments are designed to address this issue—to determine whether there is a FIE in a very powerful and stringently controlled experiment in which orientation is varied only in the study phase of a recognition procedure.

Our simulated data were designed to be qualitatively similar to the data that emerged from our experiments. In each simulation, the data issued from a particular structure, and the data were analyzed within the contexts of both linear theory and dimensional theory. In Simulation 1, we generate data from a unidimensional structure, in which the single dimension is "Strength." In Simulations 2 and 3, we generate data from a linear structure. The linear structure generating the Simulation-2 data includes a FIE, whereas the linear structure generating the Simulation-3 data does not include a FIE.

To foreshadow our findings, the simulations serve as demonstrations that data interpretation using linear theory produces a somewhat chaotic pattern of results, whereas interpretation using dimensional theory produces a consistent pattern of results. We argue that this superiority of dimensional theory over linear theory is general rather than specific to our demonstrations, and we describe underlying reasons why this is so.

Simulation 1: Unidimensional Structure

The method used in all simulations is as follows. Two stimulus sets, Faces and Houses, are presumed to have been assembled. Target stimuli are shown, one by one, in the study phase of a recognition experiment. Half the target stimuli are shown upright, while the remainder are shown inverted. A test phase follows in which the target stimuli are randomly intermingled with distracters, also consisting of Faces and Houses, drawn from the same population as the targets. All test pictures are shown upright, and old-new recognition performance is measured. There are 14 levels of study exposure duration, ranging from 30 to 420 ms in 30-ms increments. So here, as in all our simulations and experiments, there are three independent variables: stimulus duration (here, 14 levels) × orientation (Upright/Inverted) × stimulus type (here Faces/Houses).

In Simulation 1, we generated data from a unidimensional structure defined by Equations 4 and 5. We did so by implementing specific versions of the functions f, m_F, and m_H, selected to produce data that are qualitatively similar to the data that emerged in the experiments. In particular, to generate Strength, S_{ij} for duration d_i and orientation j,

$$S_{ij} = \begin{cases} M_U \times d_i & \text{for Upright stimuli} \\ M_1 \times d_i & \text{for Inverted stimuli} \end{cases}. \tag{6}$$

To generate recognition performance, p_{ijk},

$$p_{ijk} = \begin{cases} [G(S_{ij}/B_F) - 0.5] \times Y_F & \text{for Faces} \\ [G(S_{ij}/B_H) - 0.5] \times Y_H & \text{for Houses} \end{cases}, \quad (7)$$

where $G(Z)$ is the cumulative Gaussian probability of a normal deviate, Z. The free parameters in Equations 6 and 7 are interpreted as follows. M_U and M_I ($M_j > 0$) are proportionality constants mapping exposure duration to Strength for Upright and Inverted stimuli. The four parameters, B_F, B_H, Y_F, and Y_H, specify the monotonic functions mapping Strength to probability: The B_k's ($B_k > 0$) are scaling parameters for the Gaussian transformations, and the Y_k's ($0 < Y_k \leq 2$) are asymptotic values (actually asymptotic values multiplied by 2, as the bracketed portions of Equation 7 produce values between 0.0 and 0.5). These parameter values are designed to produce p values in the 0–1 range, which may be interpreted as probabilities. The six parameter values were set to $M_U = 1.50$, $M_I = 0.75$, $B_F = 400$, $B_H = 250$, $Y_F = 2.0$, and $Y_H = 1.3$.

We emphasize, as we did earlier with respect to the general case, there is no FIE within the context of this unidimensional structure.

Orientation combines with duration to produce a Strength value. Once Strength has been generated, orientation no longer has a unique representation. Face and House performance then depend only on Strength.

For simplicity, we assume that there is no statistical error, and that the experiment is perfectly counterbalanced. These assumptions allow us to directly compare linear theory and dimensional theory in their ideal forms when population parameters are known rather than estimated. We do not, of course, deny the existence of statistical error in the real world; however, in the context of the present discussions, such error constitutes an add-on complication whose consideration serves no useful purpose.

Results

The structure defined by Equations 6 and 7 yields the data shown in Figure 1. Figures 1A and 1B show the main results for Faces and Houses: memory performance as functions of stimulus duration (ignore for the moment the arrows and associated text). Within each panel, there are two curves: In this, as in subsequent
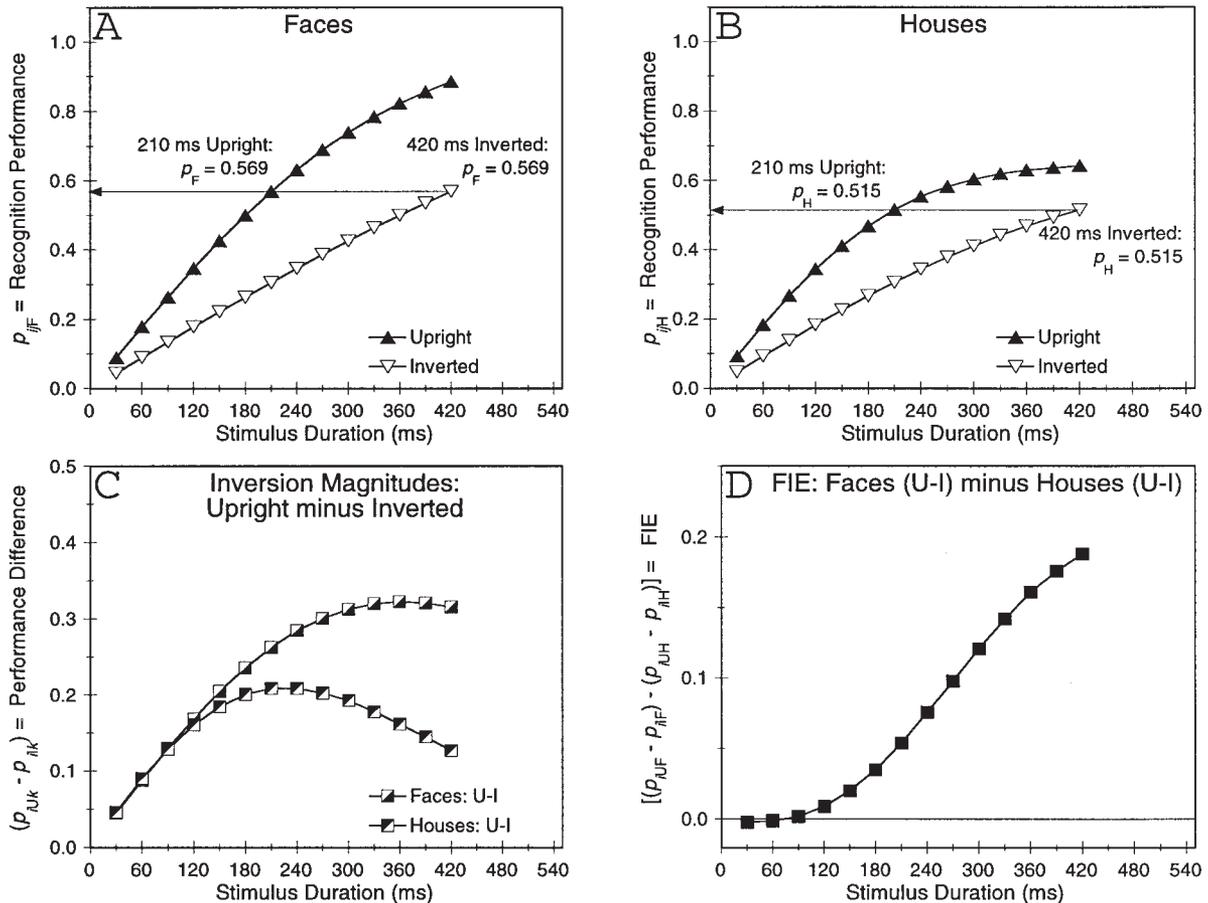

Figure 1. Simulation 1: Hypothetical data generated by a unidimensional structure. All data points are as functions of study exposure duration. A and B: Recognition performance for Faces and Houses, respectively; the two curves in each panel are for Upright and Inverted stimuli. C: Inversion magnitudes (Upright minus Inverted performance) for Faces and Houses. D: Face-inversion effect (FIE) magnitude (Face minus House inversion magnitudes). Subscripts in ordinate labels refer to duration (i), Orientation ($j \in \{U, I\}$), and Stimulus Type ($k \in \{F, H\}$).

figures, solid upright triangles represent Upright conditions, and open inverted triangles represent Inverted conditions. Figure 1C shows what we refer to as the *inversion magnitudes*—Upright performance minus Inverted performance—for both Faces and Houses, also as functions of duration. Finally, Figure 1D shows the FIE directly, in the form of the difference between the Face and House inversion magnitudes.

Although we of course know how these data were generated, we will for the moment act as if, just like normal investigators, we do not know, would like to find out, and will try to do so using an appropriate analysis and interpretation technique.

Analysis and interpretation within the context of linear theory. We first analyze and interpret the data within the context of traditional, i.e., linear theory. The equation for linear theory in this experiment is Equation 1 above (minus the error terms, as we are assuming no error), where the α_i's, β_i's, and γ_k's represent the effects of duration, orientation, and stimulus type.

A FIE would be implied by an appropriately signed orientation × stimulus type interaction. This interaction—the difference between Face and House inversion magnitudes—is shown as a function of exposure duration in Figure 1D. Its value, averaged across durations, is 0.077. Overall, therefore, a FIE would be inferred from the results of this experiment.

However, Figure 1D also indicates that the story is more complicated: There is an orientation × stimulus type × exposure duration interaction; that is, FIE magnitude varies with exposure duration. In particular, the FIE is virtually nil for stimuli shown for an exposure duration of less than about 100 ms and then increases with increasing exposure duration. One might therefore conclude that stimuli must be shown for some threshold duration in order that the FIE has a chance to develop.

Analysis and interpretation within the context of dimensional theory. In Simulation 1, we are, as noted, in the felicitous position of *knowing* the unidimensional structure that produced the data because we created it. Normally of course, the epistemological situation is reversed: We collect the data and then analyze them in quest of inferring the structure that generated them. How would we have confirmed that a unidimensional structure underlies the Simulation-1 data if we didn't know so to begin with?

The answer is that even a weak unidimensional theory of the sort expressed in Equations 4 and 5 makes a strong and unambiguous prediction that can be tested via a technique described by Bamber (1979) and Dunn and Kirsner (1988). Following Bamber's terminology, we refer to the technique as *state-trace analysis*, a tutorial on which is provided by Harley et al. (2004; see Harley et al.'s Appendix).

To understand the prediction and the analysis, consider Figure 1A. Two conditions—the 210-ms Upright condition and the 420-ms Inverted condition—happen to yield identical performance for Faces ($p_F = 0.569$). This is indicated by the horizontal arrow connecting the two data points. By Equation 5, this means that these two conditions must have produced the same Strength value, specifically $S = m_F^{-1}(0.569)$, where the superscript "-1" signifies "inverse." Equation 5 then also implies that because these two conditions have the same Strength value, they must produce equal performance values for Houses, specifically $p_H = m_H(S) = m_H[m_F^{-1}(0.569)]$. This prediction is confirmed, as indicated by the horizontal arrow in Figure 1B connecting the corresponding data points: House performance for both conditions is $p_H = 0.515$.

To maximize this example's comprehensibility, we deliberately chose a set of durations and theory parameters that would produce convenient instances of shorter-duration Upright conditions (e.g., 210-ms Upright) and longer-duration Inverted conditions (420-ms Inverted) yielding equal performance. However, the success of state-trace analysis does not require that one be lucky enough (in real life) to find such pairs. Instead, unidimensional theory is generally tested by constructing a *state-trace plot* which like the familiar receiver-operating characteristic (ROC) of signal-detection theory, is a scatterplot, over experimental conditions, of one dependent variable (here, House performance) against another dependent variable (here, Face performance). The prediction is that across all conditions—all 28 duration × orientation conditions in this instance—the scatterplot points must form a monotonic function. The reason for this monotonicity prediction is as follows. Consider two duration × orientation conditions, $C_{11} = (d_1, O_1)$ and $C_{22} = (d_2, O_2)$. Condition C_{11} gives rise to Strength S_{11} and to performance values p_{11F} and p_{11H}, whereas Condition C_{22} gives rise to Strength S_{22} and to performance values p_{22F} and p_{22H}. Suppose that for Faces, Condition C_{11} produces poorer performance than Condition C_{22}, that is, $p_{11F} < p_{22F}$. Because the m_k's are monotonic, they can be inverted to recover Strength, which has the same ordering: $S_{11} = m_F^{-1}(p_{11F}) < m_F^{-1}(p_{22F}) = S_{22}$. This, in turn, means that the ordering of house performance must also be the same because $p_{11H} = m_H(S_{11}) < m_H(S_{22}) = p_{22H}$. In other words, if this unidimensional theory is correct, then any two scatterplot points ordered in one way for faces must be ordered in the same way for houses. This defines a monotonic relation between the two.[1]

A more complete and formal treatment of this issue is provided by Bamber (1979). Briefly, one assumes that the measured data points are samples from an underlying continuous function whose form can be estimated from the data. Because the continuous function is not measured completely but only sampled—the sample corresponding here to the chosen set of stimulus durations—we must estimate the underlying function by "connecting the dots" in the scatterplot. Unidimensional theory's prediction remains that the sampled scatterplot points be monotonic. We will see instances of such monotonic state-trace plots in Simulations 1 and 3 and in Experiment 1. If the scatterplot points are nonmonotonic, we reject unidimensional theory and are required to posit additional dimensions. We will see instances of this process in Simulation 2 and in Experiments 2 and 3.

The state-trace plot for the Simulation-1 data is shown in Figure 2. Again, the solid and open curve symbols represent Upright and Inverted conditions. On a state-trace plot, two contrast × duration conditions that happen to be equal for both Face and House performance appear as two scatterplot points that fall atop one another. The two such conditions that we discussed earlier, and that are indicated in Figures 1A and 1B, indeed appear as two overlapping points enclosed by the small square in Figure 2. Overall, the scatterplot points form a monotonic function, just as

---

[1] We have one caveat: The foregoing assumes that the functions m_k are monotonically increasing. If instead they were monotonically decreasing, the ordering would change in the inverse transition from p_{ijF} to S and then would change back in the transition from S to p_{ijH}; thus, the end result would be the same.
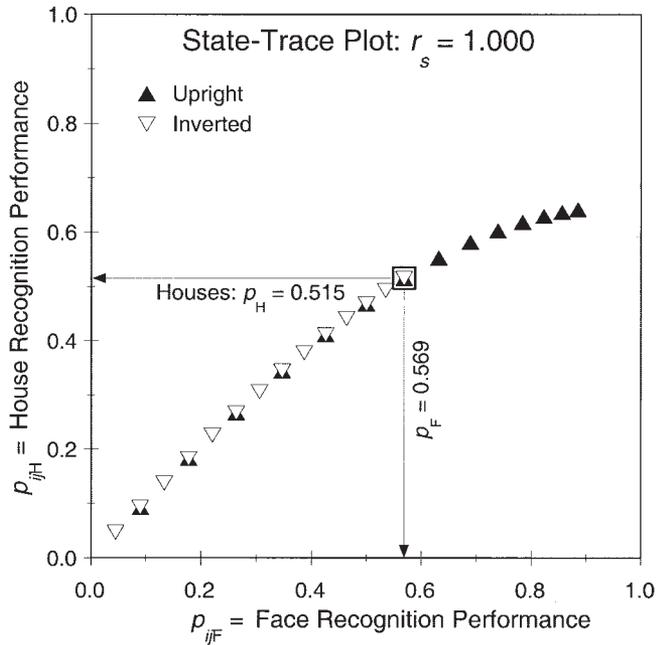
Figure 2. Simulation 1: State-trace plot for the Figure-1 hypothetical data. House performance is plotted as a function of Face performance. Each data point represents a particular duration for an Upright stimulus (solid upright triangles) or an Inverted stimulus (open inverted triangles). The two overlapping boxed points represent the 210-ms Upright and 420-ms Inverted conditions indicated in Figures 1A and 1B by the arrows and associated text: These two conditions produce equal performance for both Faces ($p_F = 0.569$) and for Houses ($p_H = 0.515$). The r_s of 1.000, which refers to the rank-order correlation between all House conditions and all Face conditions, signifies that the state-trace plot is perfectly monotonic, thereby correctly confirming a unidimensional theory.

they must, given the underlying unidimensional structure that generated them.

The simplest way of representing degree of monotonicity is with a Spearman rank-order correlation, termed r_s. If $r_s = 1.0$, then the function is monotonic; if $r_s < 1.0$, then the function is nonmonotonic. The Simulation-1 state-trace plot monotonicity is indicated at the top of Figure 2 by "$r_s = 1.000$."

Discussion

We began by creating a simple, unidimensional theoretical structure within whose context there was no FIE: Duration and orientation combined to produce a single value of Strength, and once the duration-orientation combination took place, neither duration nor orientation was represented separately. Recognition performance was then determined only by Strength, with different functions mapping Strength to recognition performance for the two stimulus types, Faces and Houses. Thus, the performance difference between Faces and Houses is not connected to orientation; rather it occurs only because of differences in the monotonic functions, m_F and m_H, mapping Strength to performance.

Analysis using linear theory. Despite the FIE-less underlying structure, analyzing and interpreting the data within the context of traditional linear theory resulted in the inference of a FIE, based on

the orientation × stimulus type interaction. Moreover, this FIE was inferred to be somewhat complicated in that its magnitude depends on the study exposure duration of the stimuli: There is inferred to be no FIE for durations less than about 100 ms, whereas for durations greater than 100 ms, the magnitude of the FIE is inferred to increase with exposure duration. If such data were observed and thusly analyzed in real life, they would doubtless spark a spirited General Discussion that would include confirmation of the FIE's existence in this paradigm, along with speculations about the intriguing manner in which the FIE appears to develop following the first 100 ms of stimulus processing. Unbeknownst to the investigator, alas, these conclusions and speculations would be entirely without foundation. (We justify this gloomy assessment below.)

Analysis using dimensional theory. In contrast, analyzing the data within the context of dimensional theory—that is, constructing a state-trace plot and making conclusions accordingly—immediately reveals the underlying unidimensional structure that produced the data. One would conclude, given this analysis and inferred structure, that there is no FIE; that is, that in this experiment, Faces and Houses are not differentially affected by orientation. This conclusion is, of course, contrary to that issuing from standard analysis within the context of linear theory, wherein a FIE was inferred from the orientation × stimulus type interaction. Equally important is that the elegance and simplicity of the theoretical structure actually responsible for the data is entirely lost with the classical linear-theory analysis: As we have seen, within the context of linear theory, a rather complex and untidy theoretical picture emerges.

The generality issue. Above we used the phrase, "without foundation" referring to conclusions about the Figure-1 data based on linear theory. Is the situation really this grim?

One might argue that because an investigator in real life does not know the nature of the structure under investigation, classical linear theory provides a safe and comfortable haven within whose context data should be analyzed and interpreted. That is, in order to foster communication among investigators and uniformity across conclusions, it makes sense to agree on a basic theory (linear theory) and then stick to it. On the one hand, we might all agree that linear theory is not quite appropriate to describe many cognitive structures, but on the other hand, we might also agree that perhaps it provides reasonable approximations to such structures. Given that a particular effect like the FIE can be inferred to exist only within the context of some theory, we may as well infer the state of its existence using the theory that most of us grew up with.

A crucial presupposition of such an argument is that there is some degree of *generalizability*: That is, small variations on the basic structure that underlies the data—variations that would come about from different materials, observers, retention intervals, settings, and so on—must, if this argument is to be tenable, produce results that are at least qualitatively similar to one another. To test this assumption, we carried out the following exercise. Note first that Equations 6 and 7, which generated the Simulation-1 data include six free parameters. On each of a series of iterations, we randomly selected new values of these parameters, drawing each new parameter from a normal distribution whose mean was the original assigned value (indicated in the paragraph just below Equation 7) and whose standard deviation was 20% of that mean

value. We constrained each parameter set so as to produce probabilities only in the allowable 0–1 range. We then generated data from this new parameter set and examined the FIE. We ran 100,000 iterations of this procedure. Each of these 100,000 iterations can be viewed as a replication of the original experiment that produced the Figure-1 data.

There are several noteworthy results of this enterprise, which are shown in the "Simulation 1" column of Table 1 in the "Linear-theory generalizations for the FIE" section and in Figure 3. To begin with, we calculated the mean FIE—the difference of Face and House inversion magnitudes averaged over durations—for each iteration. Recall that this value was 0.077 for the Figure-1 data. Over all 100,000 iterations, this mean FIE averaged 0.040 with a standard deviation of 0.052. Approximately 20% of the mean FIE values were negative, while the remaining 80% were positive.

Figure 3 shows six instances, generated by this process, of FIE magnitude as a function of duration. They are analogous to the Figure-1D curve, and for comparison, one of them *is* the original Figure-1D curve, with the original solid square curve symbols. All six of these patterns—the original and the remaining five—are representative of some of the different general patterns that emerged across the 100,000 iterations. The five new patterns were not particularly unusual; all turned up within the first 30 iterations of our 100,000-iteration simulation. Our main point here is that these patterns are (loosely speaking) all over the map. Sometimes the FIE is uniformly negative over all durations, whereas other
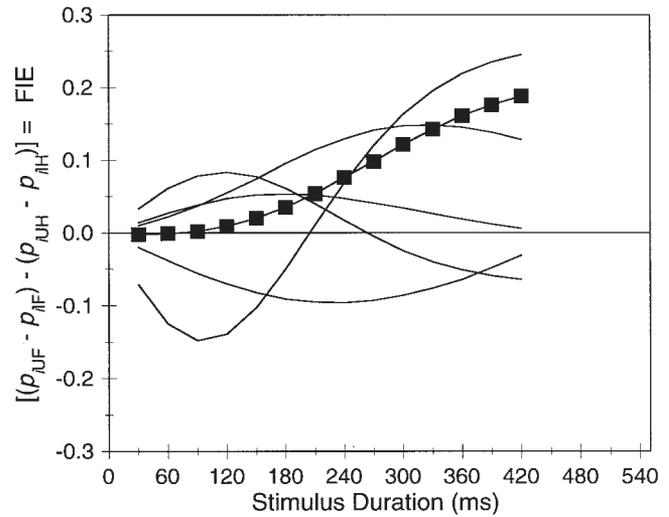

Figure 3. Simulation 1: Face-inversion effect (FIE) magnitudes as functions of study exposure duration generated by unidimensional structures. The curve with the solid square symbols is the same as the one in Figure 1D. The remaining curves result from small variations in the theory parameters that generated the Figure-1 data.

times it is uniformly positive. Sometimes it monotonically increases or decreases over exposure duration, whereas other times it is nonmonotonic. All in all, there is not one qualitative aspect of the FIE or of the interaction between FIE and duration that remains constant across relatively small variations in parameter values.

In contrast to this mess, the monotonic state-trace plot implied by the unidimensional structure that generated the data to begin with is guaranteed to show up in every iteration. This fact is underscored in the "Simulation 1" column of Table 1 in the "State trace: Distributions of r_s" section where 100% of the 100,000 r_s scores are equal to 1.0.

In short, analysis using linear theory to analyze any particular iteration, that is, any particular experiment in this simulation, would produce one sample from a profoundly chaotic underlying reality, whereas analysis using dimensional theory would consistently reveal the underlying unidimensional structure.

Simulation 2: Positive-FIE Linear Structure

In Simulation 1, we generated data with a unidimensional structure and found that the data were more accurately and consistently described by dimensional theory than by linear theory. Obviously, it is not especially surprising that the theory corresponding to the structure which generated the data is better at describing them than a theory corresponding to the structure which did not generate the data. In Simulations 2 and 3, we generated data from a linear structure.

In Simulation 2, we wanted to define a simple linear structure in which there were the expected main effects of orientation (better performance for Upright than for Inverted stimuli) and of duration (better performance for longer durations) along with a positive FIE (greater inversion magnitude for Faces than for Houses), but without any other interactions and without a stimulus-type effect.

Table 1
Various Measures of Generality for Simulations 1–3

Measure	Simulation 1[a]	Simulation 2[b]	Simulation 3[c]
Linear-theory generalizations for the FIE			
M	0.040	0.085	−0.023
SD	0.052	0.046	0.037
Maximum	0.335	0.320	0.172
Minimum	−0.287	−0.113	−0.231
% < 0	19.9	3.1	74.6
% > 0	80.1	96.9	25.4
State trace: Distribution of r_s (%)			
r_s < 0.95	0.0	33.4	0.0
0.95 ≤ r_s < 0.96	0.0	18.7	0.0
0.96 ≤ r_s < 0.97	0.0	0.0	0.0
0.97 ≤ r_s < 0.98	0.0	44.1	0.0
0.98 ≤ r_s < 0.99	0.0	0.0	0.0
0.99 ≤ r_s < 1.00	0.0	3.7	0.0
% r_s			
<1.00	0.0	100.0	0.0
1.00	100.0	0.0	100.0
% FIE implied			
Positive		100.0	
Negative		0.0	

Note. All data are based on 100,000-iteration simulations. FIE = face-inversion effect.
[a] This simulation used a unidimensional structure (no FIE). [b] This simulation used a linear structure with a positive FIE. [c] This simulation used a linear structure with no FIE.

Achieving this goal proved to be a challenge because the presumed dependent variable in our experiment—recognition probability—is, like many dependent variables, bounded—in this case between 1.0 and 0.0—whereas a linear structure, such as the one defined by Equation 1, is unbounded. In subsequent discussion sections, we have more to say about this issue. For the moment, we solved the problem by creating the linear structure in terms of a variable that we again call Strength and we again map Strength to probability using a cumulative Gaussian transform.

In particular, for the ith duration level, jth orientation level, and kth stimulus-type level we define

$$S_{ijk} = \mu + d_i + \beta_j + \kappa_{jk}, \tag{8}$$

where $j \in \{U, I\}$ and $k \in \{F, H\}$. In keeping with common linear-theory notation, we define β_j, the one-degree-of-freedom orientation main effect, to be β_U, and $\beta_I = -\beta_U$. Similarly, we define κ_{jk}, the one-degree-of-freedom orientation $\times$ stimulus type interaction terms, to be as shown in Table 2. A positive FIE implies that $\kappa_{UF} > 0$, whereas no FIE implies that $\kappa_{UF} = 0$ and therefore that all κ_{jk}'s be zero.

To map Strength to performance, we again use a Gaussian transform,

$$p_{ijk} = \begin{cases} [G(S_{ijk}/B)] \times Y_F & \text{for Faces} \\ [G(S_{ijk}/B)] \times Y_H & \text{for Houses} \end{cases}, \tag{9}$$

where $G(Z)$ is the cumulative Gaussian probability of a normal deviate, Z. The six free parameter values were set to $\mu = -250$, $\beta_U = 50$, $\kappa_{UF} = 19$, $B = 150$, $Y_F = 0.7$, and $Y_H = 0.9$. As in Simulation 1, we assume infinite statistical power and perfect counterbalancing.

Results

The structure defined by Equations 8 and 9 yields the data shown in Figure 4, which is organized like Figure 1.

Analysis and interpretation within the context of linear theory. As with Simulation 1, we first analyze and interpret the data within the context of linear theory. The conclusions are much the same as they were in Simulation 1: There are main effects of duration and orientation along with a positive FIE whose magnitude averaged over duration is 0.078. Superficially at least, these effects are to be expected given that they were built into the structure that produced them, via Equation 8. What is superficially less expected because they were not explicitly built into the structure are a duration $\times$ orientation interaction (as shown in Figure 4C, the difference between Upright and Inverted curves changes over exposure duration for both Faces and Houses) and a duration $\times$ orientation $\times$

Table 2
Organization of the Orientation $\times$ Stimulus Type Interaction Terms (κ_{jk})

Orientation	Faces	Houses
Upright	κ_{UF}	$\kappa_{UH} = -\kappa_{UF}$
Inverted	$\kappa_{IF} = -\kappa_{UF}$	$\kappa_{IH} = \kappa_{UF}$

Note. There is only one free parameter, κ_{UF}.

stimulus type interaction, as shown by the inverted U-shaped FIE curve in Figure 4D.

Analysis and interpretation within the context of dimensional theory. Neither a casual look at the Figure-4 data nor a casual look at Equations 8 and 9 that generated them provides much insight about whether the data do or do not conform to a unidimensional theory. Intuitively, one would expect that a unidimensional theory is sufficiently constrained that it would rarely if ever characterize data that didn't actually issue from an underlying unidimensional structure. Such intuition is confirmed by the state-trace plot, which is shown in Figure 5. Clearly it is nonmonotonic ($r_s = 0.977$), which indicates that the Simulation-2 data do not issue from a unidimensional structure. The particular pattern in the state-trace plot is instructive: Note that the Inverted data points are displaced to the left of the Upright data points. This means that if two duration-orientation conditions of the sort indicated in Figure 5—a long-duration Inverted condition and a short-duration Upright condition—are found for which House performance is identical, then for the same two conditions, Face performance is poorer for the Inverted condition compared with the Upright condition. We have more to say about this kind of data pattern in conjunction with its appearance in Experiments 2 and 3; for now, we note that this is the kind of state-trace plot pattern that within the context of dimensional theory implies a positive FIE—a greater deleterious effect of inversion on Faces than on Houses.

Generality. We have reported two main findings for Simulation 2. First, within the context of linear theory, there was an orientation $\times$ stimulus type interaction, of the sort that implied a positive FIE. Within the context of dimensional theory, there was a systematically nonmonotonic state-trace plot, also of the sort that implied a positive FIE. As with Simulation 1, we investigated the generality of these findings using the procedure described in conjunction with Simulation 1. We generated data sets using variations of the parameter set used in Equations 8 and 9. Each parameter in a parameter set was drawn from a normal distribution again with a standard deviation equal to 20% of the mean parameter value. In addition to constraining the parameter sets to produce only valid probabilities, we also constrained them to allow only positive FIE's; that is, we excluded any instances of $k_{UF} \leq 0$. We again ran 100,000 iterations.

The results are shown in the "Simulation 2" column of Table 1 and in Figure 6. Table 1 indicates that within the context of linear theory, 96.9% of the iterations produced a positive FIE, whereas, despite the constraint that $\kappa_{UF} > 0$, the remaining 3.1% produced a negative FIE. Figure 6, like Figure 3, indicates that even when the underlying structure is explicitly set up to be linear and simple, the observed FIE as well as the relation between FIE and exposure duration is chaotic over small variations of the underlying structure.

Within the context of dimensional theory, 100% of the state-trace plots were nonmonotonic, that is, had a less-than-1.0 r_s, thereby disconfirming a unidimensional theory. However, there are two interrelated issues to be contended with. First, disconfirmation of a unidimensional theory implies only that there is *some* kind of FIE, without specifying whether it is positive or negative. Second, rejection of unidimensional theory is based on a strict criterion that one rejects unidimensional theory whenever r_s is less than 1.0. One might argue that this criterion is *too* strict, and in real life, one might fail to reject the null hypothesis of a FIE when r_s
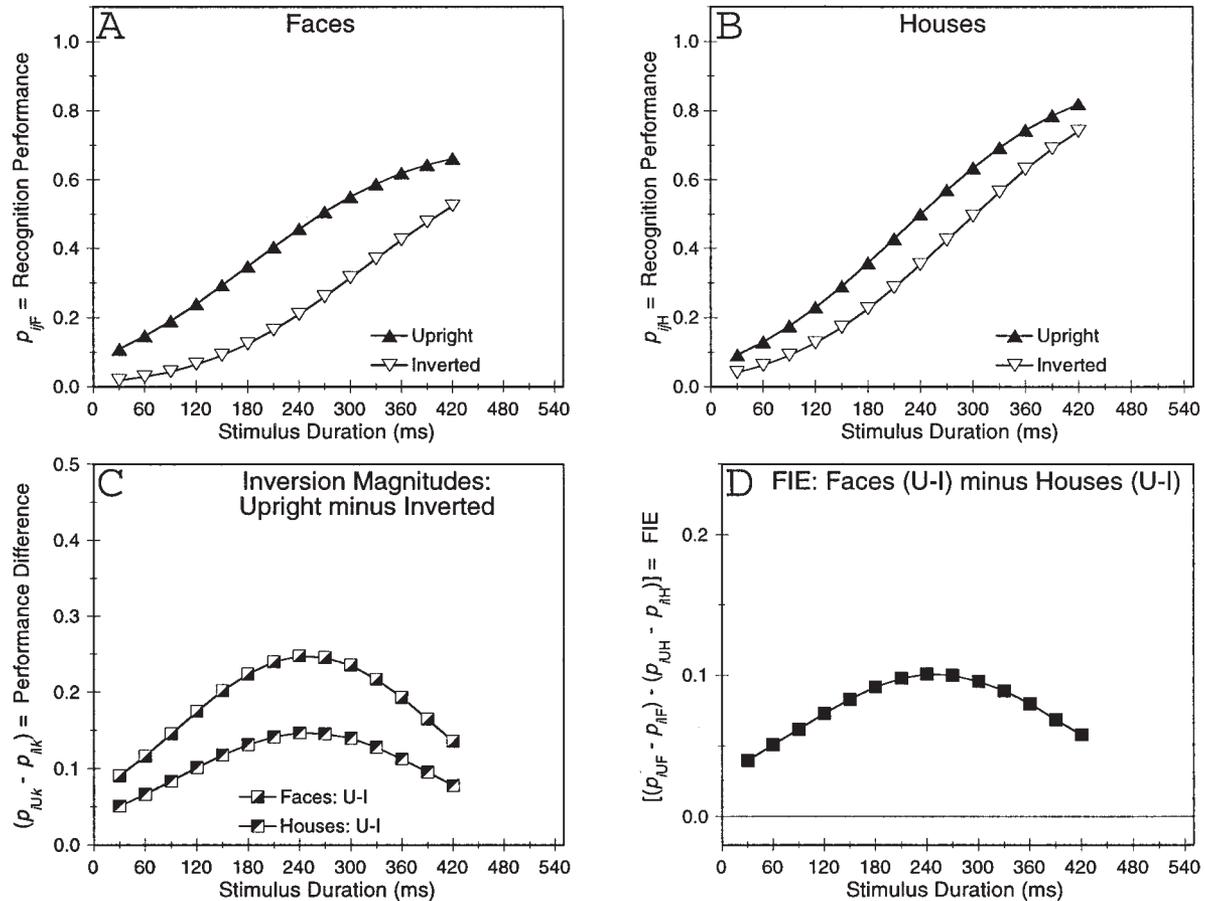
Figure 4. Simulation 2: Hypothetical data generated by a linear structure that includes a positive face-inversion effect (FIE). A and B: Recognition performance for Faces and Houses, respectively; the two curves in each panel are for Upright and Inverted stimuli. C: Inversion magnitudes (Upright minus Inverted performance) for Faces and Houses. D: FIE magnitude (Face minus House inversion magnitudes). Subscripts in ordinate labels refer to duration (i), orientation ($j = \{U, I\}$), and stimulus type ($k = \{F, H\}$).

is close to 1.0, while not requiring r_s to *be* 1.0. We address these issues in three ways.

First, and most straightforwardly, this argument does not apply to our simulations, which assume no statistical error: For each iteration, the r_s value (and every other statistic) can be accepted as 100% accurate. This means that there is no need to equivocate about the value of r_s—either it is 1.0 or it is not.

Second, even in real life, where there is statistical error, the Spearman r_s, because it uses ranked data, is a discrete or "grainy" measure; thus, it can yield a value of 1.0 even when the data are somewhat noisy (unlike, for example, a Pearson r). The graininess in r_s is exacerbated to the degree that there are few data points. Even with the 28 data points in our present scatterplots, the graininess is evident in the "Simulation 2" column of Table 1, where there are no r_s values between 0.96 and 0.97 or between 0.98 and 0.99. Thus it is reasonable to use a strict criterion of $r_s <$ 0 to disconfirm a unidimensional theory.

Our third set of comments directly addresses both the issues that we have raised. We note that although r_s is the most straightforward measure expressing the degree to which a unidimensional

structure may be inferred from a state-trace plot, it is not the *only* measure. Simply eyeballing the data can provide a start at detecting instances when even very high values of r_s—even, in fact, an r_s value of 1.0—although necessary are not sufficient to reject a single-dimension theory. This is illustrated in Figure 7, which shows scatterplots corresponding to various r_s values. In every case, even when $r_s = 1.0$, there is notable divergence between the Upright and Inverted curves.

More important, it is often possible to find empirical functions that characterize each of the implied functions within the scatterplot—in the case of the Figure-5 state-trace plot, for instance, the functions relating Houses to Faces for Upright stimuli (solid-triangle data points) and for Inverted stimuli (open-triangle data points). Such functions can then be used to infer more precisely the nature of the underlying structure that produced the data. Again harkening to a familiar example, the usual observed cumulative Gaussian form of an ROC curve provides a tool that allows an investigator to infer the ratio of signal to noise distribution variances. We will see an instance of this technique in a later section when we describe specific (linear) functions fitting the real state-
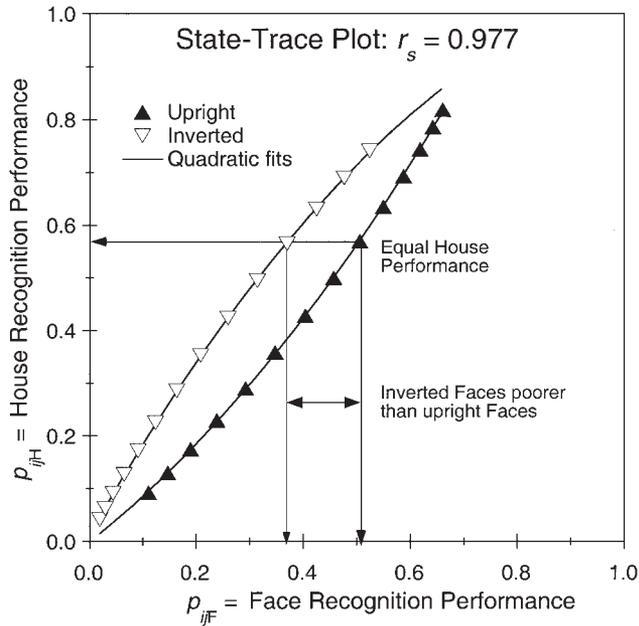
State-Trace Plot: $r_s = 0.977$

Figure 5. Simulation 2: State-trace plot for the Figure-4 hypothetical data. Solid lines through the data points are best quadratic fits, as explained in the text.

trace plots obtained in Experiments 1–3, along with theoretical constraints that are implied by these observed linear functions.

Such functions can be used to determine whether a rejected single-dimension theory implies a positive or a negative FIE. In the present instance, we made such a determination as follows. First, we observed that the individual curves in all our simulated state-trace plots could be fit very well by quadratic functions, as shown by the solid lines in Figure 5. Thus, we were able to compute such quadratic curves for each iteration and to determine the mean amount by which the Inverted curve was displaced to the left of the Upright curve. As we have already described, such a pattern implies a positive FIE. This mean displacement amount was positive, thereby implying a positive FIE, for 100% of the 100,000 iterations.

Discussion

The Simulation-2 data were generated by a linear structure in which there was a positive FIE. This positive FIE was correctly inferred from approximately 97% of variants on the data by using linear theory and from 100% of variants on the data by using dimensional theory. Even though these data sets were generated by a linear structure, dimensional theory is perfect in producing the correct inference of a positive FIE, whereas linear theory is imperfect. As was the case in Simulation 1, the relation between the FIE and exposure duration was chaotic over minor changes in the data structure.

It is apparent that even when the data are generated from a linear structure to begin with, data interpretation within the context of dimensional theory produces more coherent and consistent results than interpretation within the context of linear theory. A major reason for this is that the actual data being analyzed depart from

the structure that underlies them via a monotonic but nonlinear (Gaussian) measurement mapping. As we discussed earlier and as we elaborate below, there are, however, no reasonable alternatives to this.

Simulation 3: Zero-FIE Linear Structure

In Simulation 3, as in Simulation 2, we generated data from a linear structure. Unlike Simulation 2, there was no FIE in the Simulation-3 structure. This was accomplished by using Equations 8 and 9, which generated the Simulation-2 data, with the same parameter values as in Simulation 2, except that for Simulation 3, all $\kappa_{jk} = 0$.

Results

The Simulation-3 data are shown in Figure 8, which is organized like Figures 1 and 4.

Analysis and interpretation within the context of linear theory. As with Simulations 1 and 2, we first analyze and interpret the data within the context of linear theory. Even though there was no FIE in the structure that generated the data, the data themselves imply a negative FIE whose mean magnitude is -0.039 and which as shown in Figure 8D, varies in a U-shaped fashion with exposure duration.

Analysis and interpretation within the context of dimensional theory. The state-trace plot corresponding to the Simulation-3 data is shown in Figure 9. It is completely monotonic, which implies no FIE. We consider this observation to illustrate a critically important point, about which we say more below.

Generality. We ran the same kind of 100,000-interation simulation described for Simulations 1 and 2. In addition to the constraints already described, κ_{UF} was set to 0, thereby implying
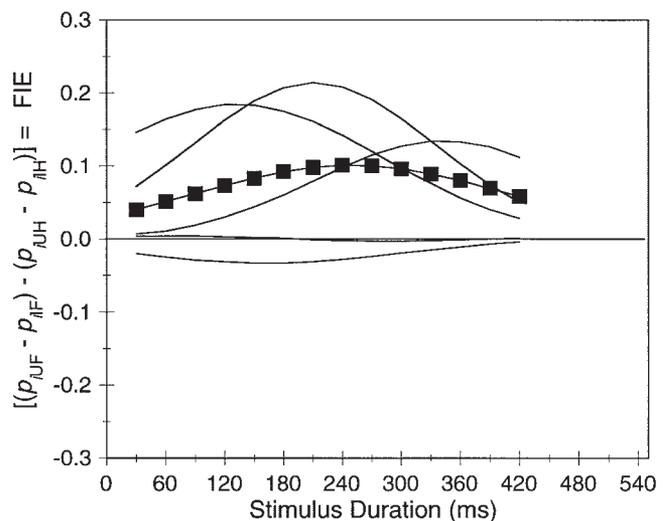

Figure 6. Simulation 2: Face-inversion effect (FIE) magnitudes as functions of study exposure duration for data generated by linear structures within whose contexts a FIE exists. The curve with the solid square symbols is the same as the one in Figure 4D. The remaining curves resulted from small variations in the theory parameters that generated the Figure-4 data.
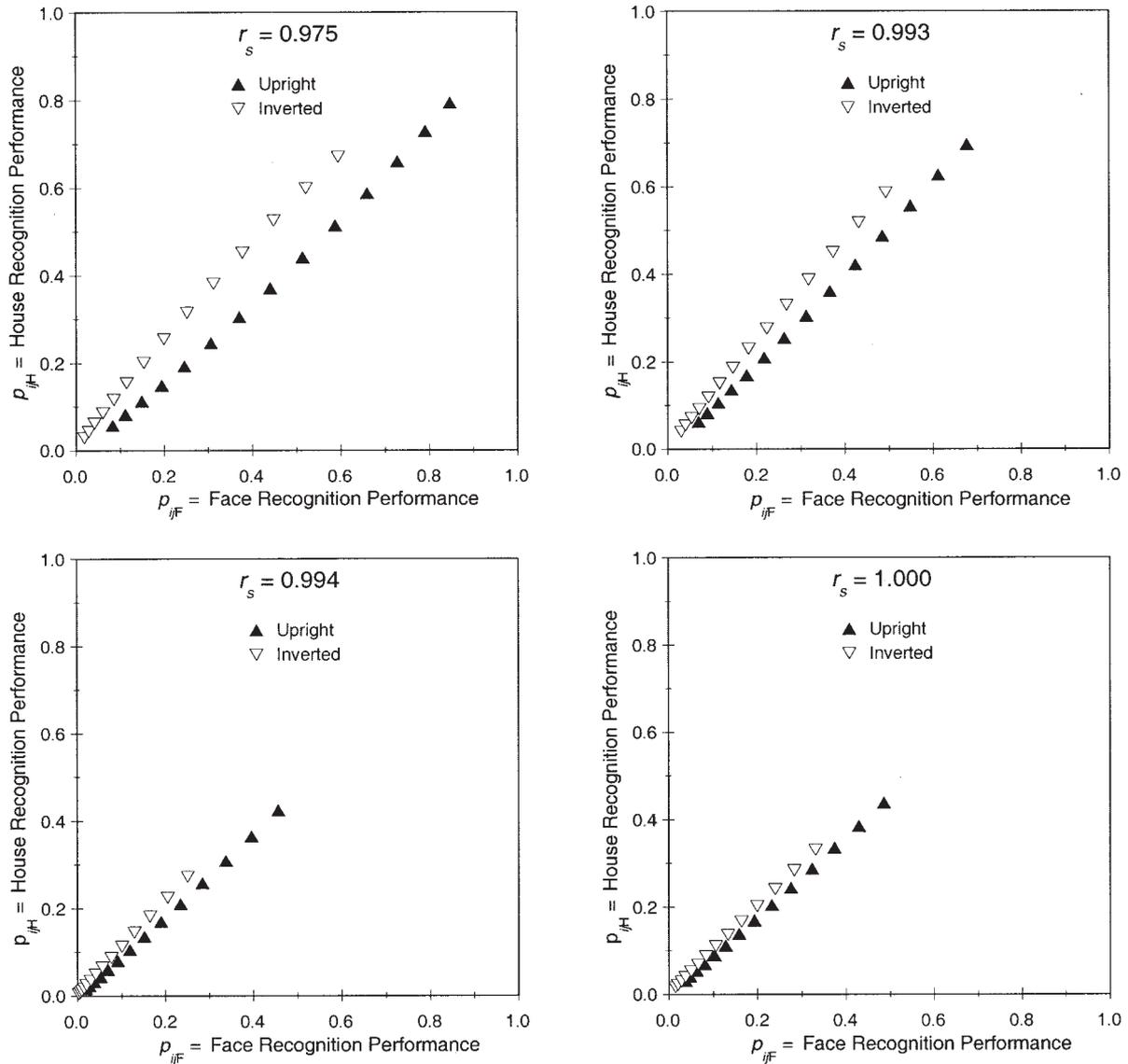
Figure 7. Examples of state-trace plots corresponding to various rank-order correlation values. Each curve was generated using minor variations of the parameters used by the linear theory that generated the Figure-4 data.

no FIE, for all iterations. The "Simulation 3" column of Table 1 shows the results. Across the 100,000 iterations, the mean magnitude of the FIE was -0.023, with a standard deviation of 0.037. About 25% of the FIE's were positive, and the rest were negative. Within the context of dimensional theory, the value of r_s was 1.0, thereby implying no FIE for every iteration. Figure 10, which is organized like Figures 3 and 6, shows that the FIE × duration interaction is again chaotic over small variations in the data structure.

Discussion

Although the data were generated by a linear structure that did not include a FIE, a small negative FIE would be inferred by analysis assuming a linear theory. Analysis assuming dimensional theory, however, yielded the correct inference of no FIE, both with the data shown in Figure 8 and with all 100,000 variants of the Figure-8 data.

This latter finding turns out to be entirely understandable. The FIE-less linear-theory equation generating the Simulation-3 data is

$$S_{ij} = \mu + d_i + \beta_j. \tag{10}$$

That is, it is the same as Equation 8, which generated the Simulation-2 data, except that there are no interaction terms, κ_{jk}. Therefore Equation 10, wherein S_{ij} depends only on duration and orientation, is an instance of Equation 4, which describes a unidimensional theory. Mapping Strength to performance is accomplished with Equation 9, which is an instance of Equation 5. Therefore, Equations 10 and 9, which generated the Simulation-3
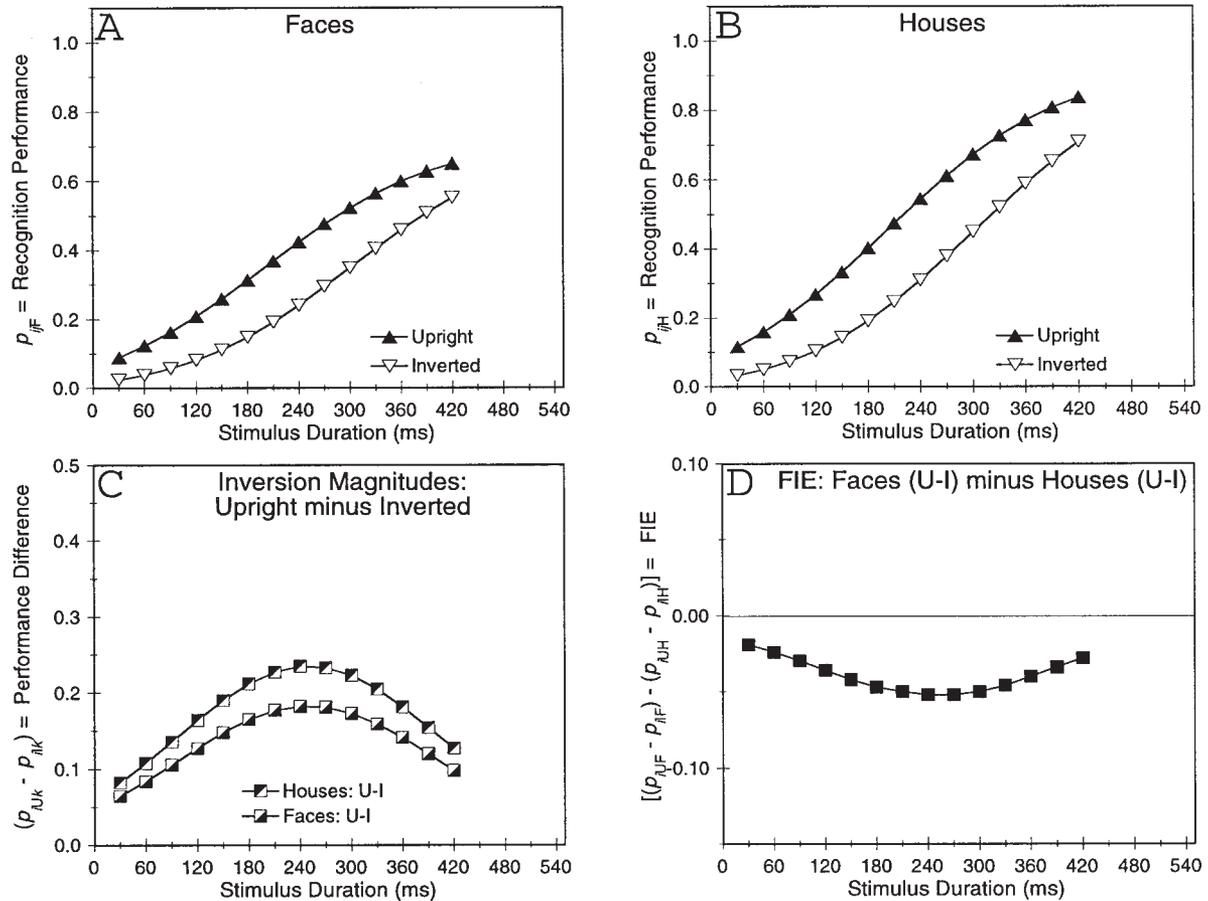
Figure 8. Simulation 3: Hypothetical data generated by a linear structure that includes no face-inversion effect (FIE). A and B: Recognition performance for Faces and Houses, respectively; the two curves in each panel are for Upright and Inverted stimuli. C: Inversion magnitudes (Upright minus Inverted performance) for Faces and Houses. D: FIE magnitude (Face minus House inversion magnitudes). Subscripts in ordinate labels refer to duration (i), orientation ($j \in \{U, I\}$), and stimulus type ($k \in \{F, H\}$).

data, are instances of Equations 4 and 5, which define a unidimensional theory. In short, the Simulation-3 data were generated by a disguised unidimensional structure, which is reflected in the universally monotonic state-trace plots.

Despite being an extraordinarily simple linear structure, the pattern of the FIE over exposure duration was again chaotic. Again, this results from the nonlinear relation between the underlying Strength structure and the measured probabilities. More generally, as we will eventually detail in our General Discussion, a linear structure involving no stimulus-type interactions implies a monotonic state-trace plot.

Discussion: Simulations

Our three simulations were designed to generate data from either a unidimensional structure or a linear structure and then to determine how effective are both linear theory and dimensional theory as a context for inferring the existence of a FIE.

Summary

In Simulation 1, data were generated from a unidimensional structure that included no FIE. Analysis within the context of

linear theory never produced the inference of no FIE; rather, it led to the inference of a positive FIE about 80% of the time and to the inference of a negative FIE the remaining 20% of the time. Analysis within the context of linear theory also produced inconsistent inferences about the relation between FIE magnitude and exposure duration. Analysis within the context of dimensional theory produced, as it logically had to, the inference of no FIE 100% of the time.

In Simulation 2, data were generated from a linear theory that included a positive FIE. Analysis within the context of linear theory produced the inference of a positive FIE approximately 97% of the time but also produced inconsistent inferences about the relation between FIE magnitude and exposure duration. Analysis within the context of dimensional theory produced the inference of a positive FIE 100% of the time.

In Simulation 3, data were generated from a linear theory that included no FIE. Analysis within the context of linear theory never produced the inference of no FIE; rather, it led to the inference of a positive FIE about 25% of the time and to the inference of a negative FIE the remaining 75% of the time. As in Simulations 1 and 2, analysis within the context of linear theory also produced
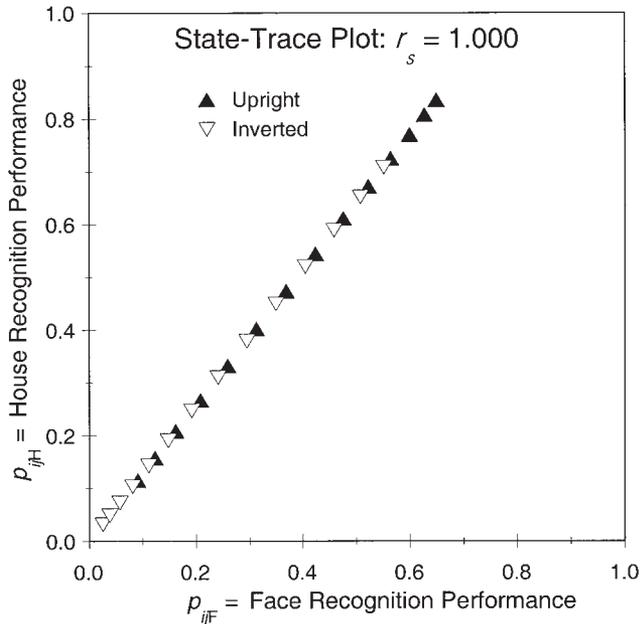
Figure 9.   Simulation 3: State-trace plot for the Figure 8 data.

inconsistent inferences about the relation between FIE magnitude and exposure duration. Analysis within the context of dimensional theory produced the correct inference of no FIE 100% of the time. This occurred because the linear structure that generated the data was an instance of a unidimensional structure.

The Unnatural Nature of a Linear Structure

Earlier, we mentioned that a linear structure cannot accurately characterize actual perceptual or cognitive data when some form of probability or other bounded variable is used as the response measure. We view this as an "elephant in the living room" problem: It is a big, serious, in-your-face problem that is rarely acknowledged. It was to solve this problem that in Simulations 2 and 3, we devised the linear structure in terms of "Strength," and then transformed Strength to probability via a nonlinear (Gaussian) mapping. The consequences of such nonlinearities have been discussed in the literature (see Loftus, 1978) and can lead to all manner of illusion and inconsistency, particularly when interpretation of interactions are at issue. The inconsistencies, apparent over iterations in all simulations when analyzed within the context of a linear theory, are simply examples of this general problem.

One might complain that in Simulations 2 and 3 we short-changed linear theory by not making the actual data structure—the probabilities—linear. The problem is that there is no realistic alternative to this procedure. If the linear structure is constructed in terms of probabilities, then the equations, unless constrained, will routinely produce illegal less-than-zero or greater-than-one values. Constraining probabilities to be between 0.0 and 1.0 creates functions that are, again, nonlinear. One could very carefully constrain the linear-theory parameters and values of the independent variables in such a way that no probability exceeds legal bounds, but this procedure implies very limited theory construction—it severely constricts the range of realistic situations to which linear

theory could potentially apply. Finally, one might argue that a hypothesized underlying linear structure could be recovered by applying an appropriate nonlinear transformation to the raw probabilities that will, at the very least, unbound them. This is done, for example, when probability is transformed to d'. It could have been done in the present experiments by applying the inverses of Equation 9 to the Simulation-2 and Simulation-3 data. But this is not a viable general procedure because there is an infinite number of possible nonlinear transformations, and one can never know a priori which one should be applied. Equation 9 illustrates this point: Its inverse would be far from obvious if one didn't know exactly what it was to begin with.

Conclusions Based on Our Simulations

Given the results of our simulations, we make several conclusions. First, even if a linear structure underlies the data, a linear theory does poorly in identifying its nature. In Simulation 2, we constructed the simplest possible linear structure consistent with the undebated effects of exposure duration and orientation, plus a FIE. Analysis within the context of linear theory was inconsistent with respect to both recovering the FIE itself and with respect to determining the relation between the FIE and other variables, particularly exposure duration. It is true that we could construct other linear structures besides the one that we chose; for example, we could add a stimulus-type main effect and/or other interactions to the structure. However, any other such linear structure would be more complex—it would include the one that we examined as a special case—and the data would therefore be even less consistent across variation in parameter values. Analogous arguments apply to the results of Simulation 3. Here analysis within the context of a linear theory was similarly inconsistent and was similarly poor in detecting the lack of a FIE and the lack of a FIE × exposure
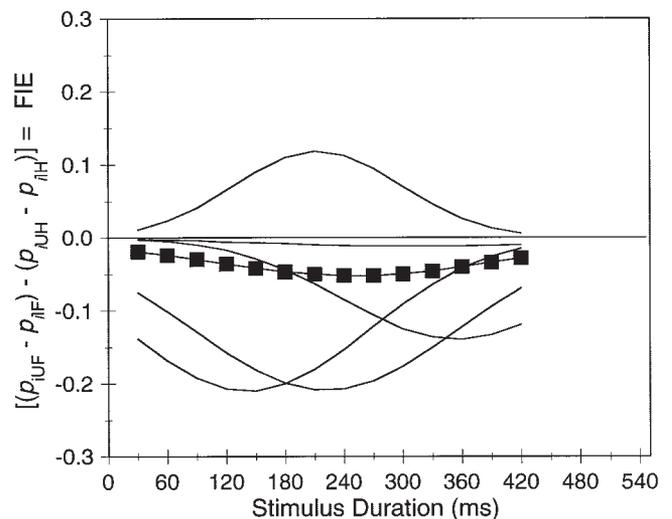

Figure 10.   Simulation 3: Face-inversion effect (FIE) magnitudes as functions of study exposure duration for data generated by linear structures within whose context there is no FIE. The curve with the solid square symbols is the same as the one in Figure 8D. The remaining curves resulted from small variations in the theory parameters that generated the Figure-8 data.

duration interaction that characterized the underlying linear structure.

In contrast, analysis within the context of dimensional theory was considerably more successful in determining whether a particular effect—in this case a FIE—exists. Given that a unidimensional structure describes the data, as in Simulations 1 and 3, a monotonic state-trace plot emerges with a 100% probability, thereby producing an unambiguous conclusion. Even when a linear structure generated the data, as in Simulation 2, analysis within the context of dimensional theory was still 100% effective—better, in other words, than linear theory—in revealing the positive FIE that was built into the linear structure.

Experiments

With our simulations, we have accomplished much of what we set out to do in this article: On the basis of our demonstrations, we have argued that dimensional theory should be considered as a preferable alternative to standard linear theory as a means of designing experiments, along with thinking about, analyzing, and interpreting data. We now change tack: We return to the real world of real data and describe three real experiments whose purpose is to address a specific question about the genesis of the FIE. The designs of the experiments were similar to those of the simulations.

There are two points of connection between the first part of this article (i.e., the simulations and what we have discussed so far about linear theory versus dimensional theory) and the second half (i.e., the real data and their interpretation). The first point of connection is the Experiment-1 data, which, as it turned out, resemble the Simulation-1 data: A conclusion about whether there is a FIE depends on whether the data are interpreted within the context of linear theory or within the context of dimensional theory. Because interpretation of the Experiment-1 data is pivotal in interpretation of the Experiments 1–3 data taken as a whole, we must choose one theory or the other. For reasons that we have already discussed, we argue that interpretation within the context of dimensional theory is more meaningful. The second point of connection is the use of dimensional theory itself: We show how to develop a set of successively stronger dimensional theories to account for our data.

Our primary goal in Experiments 1–3 was to investigate whether the FIE is confined to retrieval of already stored faces or whether it emerges during encoding of novel faces as well. In quest of this goal, we varied orientation in the study phase but not in the test phase of an old-new recognition procedure. In Experiment 1, we ensured that only encoding of faces occurred at study by using computer-generated faces that were unfamiliar to our observers. In Experiment 2, we induced face retrieval at the time of study by using faces of known celebrities. Experiment 3 was a replication and generalization of Experiment 2 using complex, heterogeneous Cityscapes instead of homogeneous Houses as our alternative stimulus set.

Experiment 1: Unfamiliar Faces and Houses

In Experiment 1, we compared computer-generated Faces with Houses. The general design was much the same as that described in conjunction with the simulations.

Method

Observers. A total of 366 University of Washington undergraduates participated in 48 groups of 6 to 8 observers per group, in exchange for course credit.

Stimuli. Two stimulus sets were created: Faces and Houses. The FACES Identikit program (IQ Biometrix, Fremont, CA) was used to generate a set of 144 lifelike faces: 73 male faces and 71 female faces. Houses were 144 photos of similar houses in a middle-class Seattle neighborhood. All stimuli were rendered as grayscale images, 450 pixels high $\times$ 400 pixels wide. Each picture's luminance ranged from 0.2 to 14.8 cd/m^2 across pixels. Examples of the stimuli are provided in the top two rows of Figure 11.

Apparatus. Stimuli were displayed on a flat-white wall via an LCD projector interfaced to a Macintosh G4. Stimulus displays were accomplished within MATLAB, using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Observers sat in eight seats arranged in two rows of four. Stimuli subtended visual angles of 44° or 26° vertically and 57° or 34° horizontally for observers in the front or back row. When no stimulus was displayed (a blank screen), luminance was 3.7 cd/m^2.

Design. Each group participated in an experimental session that incorporated two subsessions. For 24 of the 48 groups, Faces were stimuli in the first subsession, while Houses were stimuli in the second subsession. The Faces-Houses subsession order was reversed for the other 24 groups. Except for the different stimuli, the two subsessions were identical.

The basic unit of each subsession, called a *tray* (honoring its physical slide-tray ancestry), consisted of 24 virtual "slides" or pictures. There were six trays in each subsession. For each tray, 12 target pictures were shown in a study phase. A test phase immediately followed in which all 24 pictures in the tray—the 12 just-viewed targets, randomly intermingled with 12 never-seen distracters—were presented in an old-new recognition procedure.

Each target stimulus fell into one of 12 study conditions, defined by combining Upright/Inverted with six stimulus durations: 17, 33, 50, 83, 150, and 250 ms. During the study phase of each tray, each of the 12 study conditions was assigned to one of the 12 target stimuli in a manner that although constrained by counterbalancing procedures (see below), appeared random to the observers: Thus the observers never knew the duration or orientation of an upcoming target picture. They did, however, know its stimulus type because, as indicated, subsessions were blocked by stimulus type.

Procedure. An experimental session began with the display of four practice pictures (Faces or Houses depending on which stimulus type was to be used during the first subsession) shown in a randomly selected 4 of the 12 study conditions, along with detailed study instructions and general test instructions. After the first study phase, detailed test instructions were provided. Following the six trays in the first subsession, observers had a brief break before proceeding through the second subsession.

Each study trial began with a 1,000-ms fixation cross in the center of the screen, accompanied by a 1000-Hz warning tone. Next came a 300-ms blank screen, then the target picture in its assigned orientation for its assigned duration, and then the blank screen again. To maintain moderate performance, we reduced target-picture contrast by scaling pixel luminance to range from 2.4 to 5.6 cd/m^2 for each picture. Following each study picture, observers rated their confidence that they would subsequently be able to recognize the picture on a scale from 1 (*definitely no*) to 4 (*definitely yes*). The purpose of this rating procedure was to encourage the observers to attend to the target pictures, and the study-phase rating data are not further considered in this article.

A single test picture was shown during each test trial. Observers rated their confidence that they had seen it during the previous study phase, using a scale value of 1 (*definitely no*), 2 (*probably no*), 3 (*probably yes*), or 4 (*definitely yes*). All test pictures were shown upright at full contrast and remained visible until everyone had responded.
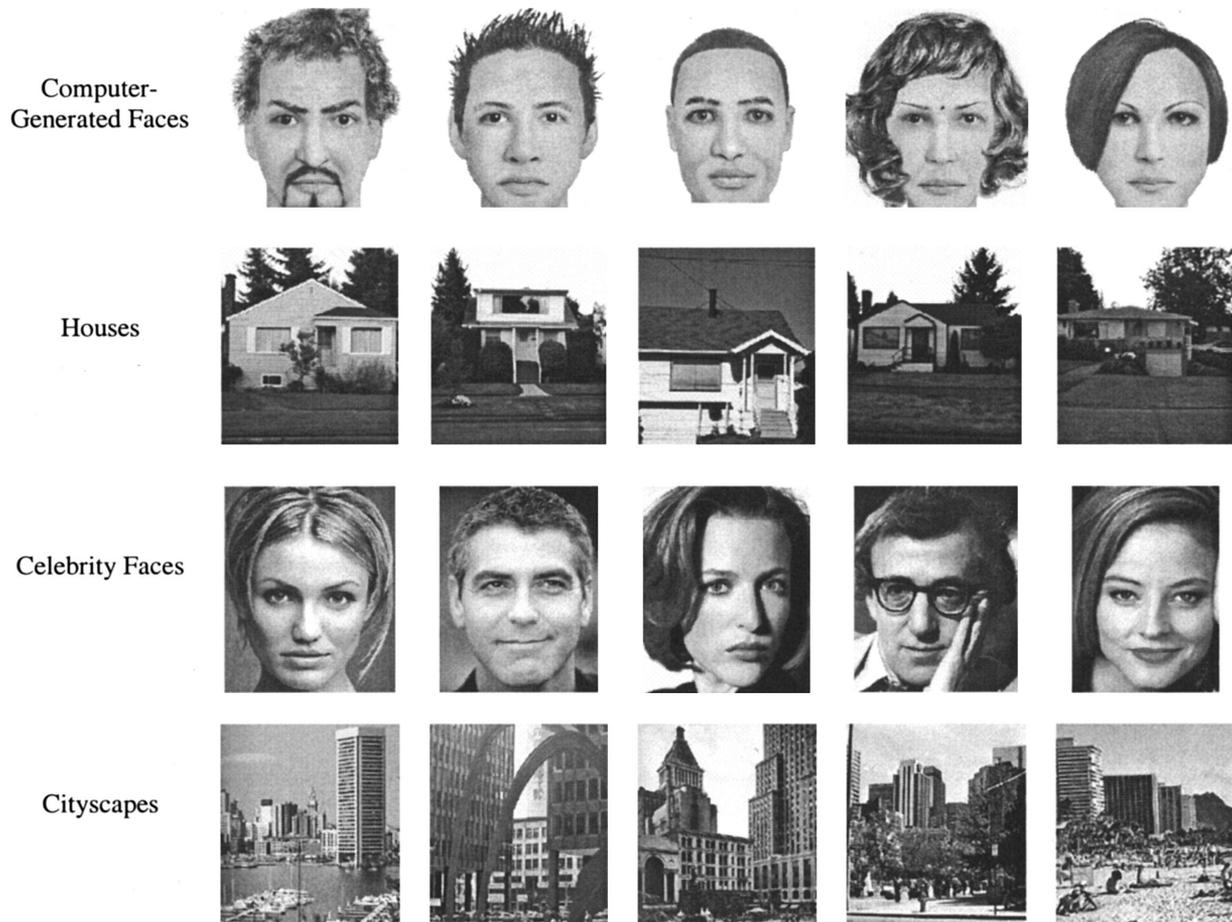
Figure 11. Examples of stimuli used in Experiment 1 (Rows 1 and 2), Experiment 2 (Rows 2 and 3), and Experiment 3 (Rows 3 and 4). From "Why Is It Difficult to See in the Fog? How Contrast Affects Visual Perception and Visual Memory," by E. M. Harley, A. Dillon, and G. R. Loftus, 2004, *Psychonomic Bulletin & Review, 11,* p. 213. Copyright 2004 by the Psychonomic Society. Adapted with permission.

Counterbalancing. Each stimulus picture was seen as a target by 24 groups and as a distracter by the remaining 24 groups. Each picture was seen twice in each of the 12 study conditions over the 24 groups that saw it as a target. This counterbalancing was done independently for each of the 12 trays.

Results

The dependent variable was mean test confidence rating, which ranged from 1–4. For each stimulus type, Faces and Houses, there were 13 separate mean confidence ratings: 12 for targets in the 12 study conditions, which are akin to hit rates, plus the mean confidence rating for distracters, which is akin to a false-alarm rate. To produce comparable performance across the two stimulus types, we carried out the following procedure. First, we transformed confidence rating (*CR*) to proportion (*pr*) by the transformation, $pr = (CR - 1)/3$; thus for each stimulus type, we had 12 hit rates (H_{ijk}) plus 1 false-alarm rate (FA_k), $k \in \{F, H\}$. Then, for each stimulus type, we generated a single performance measure for each of the 12 study conditions. This measure, which we term *p*, combines the hit rates and the false-alarm rate, using the equation

$$p_{ijk} = \frac{H_{ijk} - FA_k}{1 - FA_k}. \tag{11}$$

Note that although the high-threshold assumptions embodied in Equation 11 are dubious, this transformation causes no serious interpretational difficulties because (a) we are not interested in differences between the two stimulus types per se, (b) we are not interested in differences between the two false-alarm rates, and (c) because there is only a single false-alarm rate for each stimulus type, the transformation embodied in Equation 11 preserves the ordering among the hit rates within each stimulus type. Moreover, p_{ijk}, which ranges from 0.0 corresponding to no memory to 1.0 corresponding to perfect memory, is a convenient measure to use as a basis for evaluating theoretical predictions that we describe following our presentation of Experiment 2.

The main results—performance as functions of exposure duration—are shown in Figures 12A and 12B. In Figure 12, as in all our real-data figures, error bars represent standard errors. The theoretical curves through the data points are described following our presentation of Experiment 2. As expected, performance in-
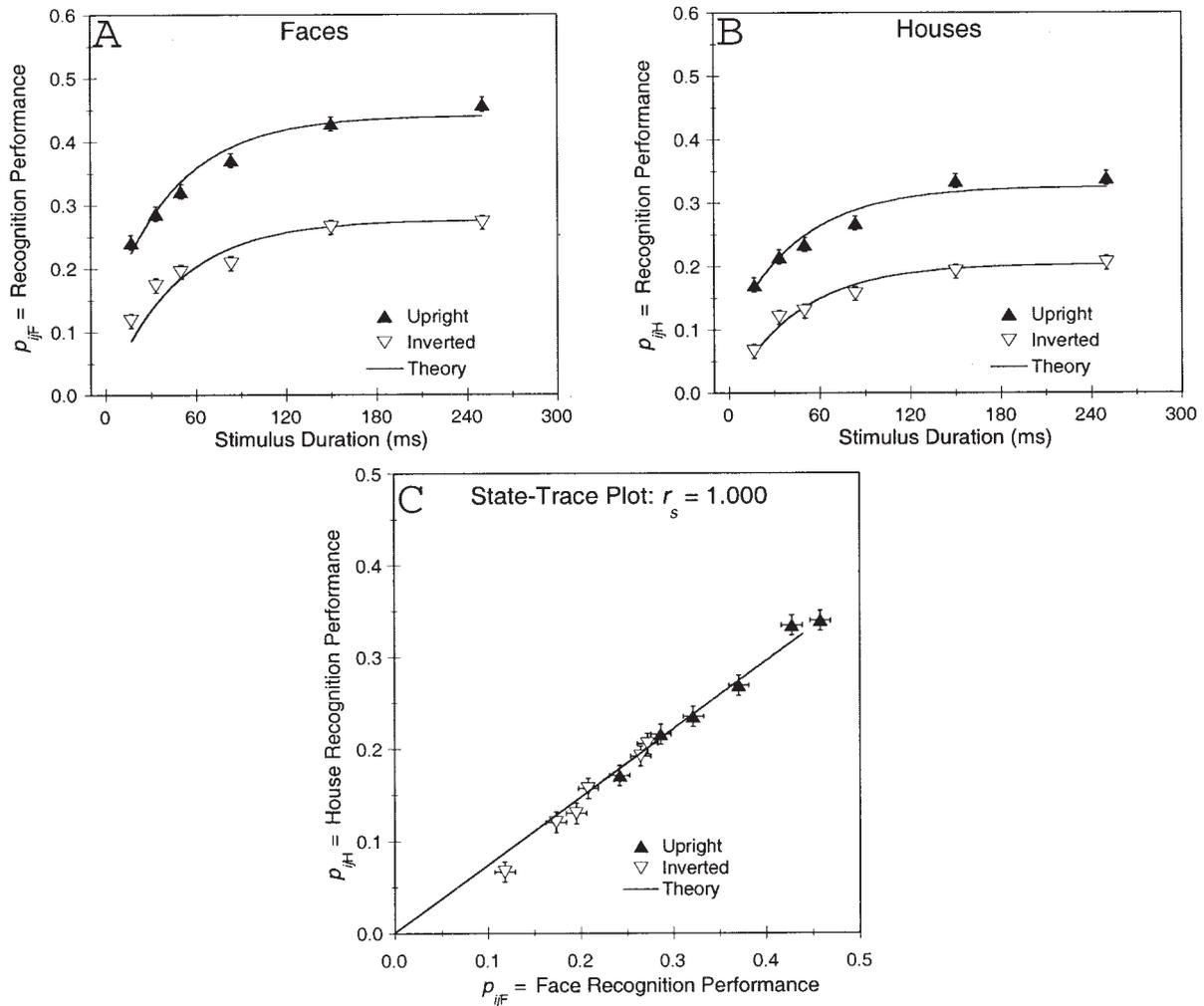
Figure 12. Experiment-1 data. A and B: Basic data, performance as functions of duration for Faces and Houses, respectively. C: The state-trace plot. Solid lines represent predictions from the theory described in the text. Each data point is based on 2,196 observations. Error bars are standard errors.

creases with exposure duration and is greater for upright than for inverted pictures.

Analysis and interpretation within the context of linear theory. As is apparent by comparing the difference between the two curves in Figure 12A with the difference between those of Figure 12B, there is a stimulus type × orientation interaction: Inversion affects Faces more than Houses. The interaction's mean magnitude is 0.042 ± 0.022 (95% confidence interval). Accordingly, within the context of linear theory, one would conclude that there is a small but unequivocal FIE in Experiment 1.

Analysis and interpretation within the context of dimensional theory. Figure 12C shows the state-trace plot from Experiment 1. It is clearly monotonic ($r_s = 1.0$), and there is nothing about the nature of the plot that would lead one to conclude that it reflects anything but a unidimensional structure. Accordingly, within the context of dimensional theory, one would conclude that there is no FIE in Experiment 1. The data are perfectly consistent with the unidimensional theory expressed in Equations 4 and 5: Orientation

and duration combine at perception to form a single value on a single dimension ("Strength"), and Face and House recognition performance are then determined by different monotonic functions of Strength.

Discussion

In short, the Experiment-1 data resemble the Simulation-1 data. A reasonable conclusion is that this happens because the Experiment-1 data reflect the kind of structure that generated the Simulation-1 data: an underlying unidimensional structure but including, within the context of a linear theory, a positive orientation × stimulus type interaction. It is with this kind of structure that the conjunction of results seen in Experiment 1—a positive FIE (within the context of linear theory) along with a monotonic state-trace plot (within the context of dimensional theory)—emerges.

Given that the Experiment-1 data reflect such a structure, the conclusion about the existence or nonexistence of a FIE depends, as it did in Simulation 1, entirely on which theory within whose context one chooses to interpret the data: If one operates within the context of linear theory, one infers a FIE in Experiment 1, whereas if one operates within the context of dimensional theory, one infers that there is no FIE. For reasons that we have discussed earlier in conjunction with our simulations, and to which we return in our General Discussion, we find it most parsimonious and sensible to operate within the context of dimensional theory, to conclude that the Experiment-1 data are well described by a unidimensional theory, and to therefore infer there is no FIE in Experiment 1. This provides support for Valentine's (1988) assertion that there is no FIE when orientation is manipulated only in the study phase of a recognition experiment. It suggests that whatever processing is responsible for the FIE applies to retrieval of known faces but not to encoding of unfamiliar faces.

Experiment 2: Celebrity Faces and Houses

In Experiment 1, we concluded that there is no FIE when stimulus sets are unfamiliar faces and homogeneous Houses and when orientation is manipulated in the study phase of a recognition procedure. We hypothesized that this comes about because the FIE results from processing that occurs when faces are retrieved from memory rather than processing that occurs when facial features are stored in memory. In Experiment 2, we used the same paradigm as in Experiment 1 but with a change in stimuli: Instead of unfamiliar computer-generated faces, we used celebrity faces. With the Experiment-1 computer-generated faces, the observer's task at study was to encode visual information from a target face in hopes of being able to eventually match such information when the same face is eventually shown at test. In contrast, with celebrity faces, the observer's task becomes one of recognizing the celebrity and encoding the celebrity's appearance as a target (e.g., "Oh, that's Jennifer Lopez; I'd better remember that she showed up during this study phase"). At test, it then suffices to re-recognize the celebrity and simply remember her appearance during the test phase (e.g., "Hmmm, it's Jennifer Lopez; oh yes, I saw her during the study phase"). If the FIE occurs when already known faces are retrieved from memory, then a FIE should be observed in Experiment 2.

Method

Observers. A total of 295 University of Washington undergraduates participated in 48 groups of 5 to 8 observers per group, in exchange for course credit.

Stimuli. A new stimulus set, Celebrity Faces was created. These were 144 photos of recognizable celebrities: 90 men and 54 women. All were rendered as grayscale images. Faces were 450 pixels high and ranged in width from 294 to 466 pixels. Examples are provided in the third row of Figure 11. Houses were the same as in Experiment 1.

Apparatus. The apparatus was the same as that used in Experiment 1.

Design, procedure, and counterbalancing. Procedures and counterbalancing measures were the same as in Experiment 1. The design was identical to that of Experiment 1 except that exposure durations were longer by a factor of 3 for the Inverted than for Upright pictures. For Upright pictures, exposure durations were 17, 33, 50, 83, 150, and 250 ms, whereas for Inverted pictures, they were 50, 100, 150, 250, 450, and 750 ms. This procedure was implemented as a compromise between two conflicting goals: The first goal was to allow more performance overlap

between the Upright and Inverted scatterplot points on the state-trace plot so as to allow a more stringent test of monotonicity—and hence of unidimensional theory—and the second goal was to allow enough exposure-duration overlap between the Upright and Inverted orientation levels to be able to reasonably estimate linear-theory parameters.

Results

The results are shown in Figure 13, which is organized like Figure 12. The orientation × stimulus type interaction cannot be computed for the complete data set because of the different exposure durations for the two orientations. However, three exposure durations—50, 150, and 250 ms—are common to both orientations. Considering these three durations, the orientation effect is greater for Faces than for Houses: The interaction's mean magnitude is 0.224 ± 0.031. Inspection of Figures 13A and 13B indicates that allowing curve interpolation, the interaction is present over exposure duration at least up to 250 ms. Similarly, curve extrapolation strongly suggests that it would exist for longer durations as well. In short, there is every reason to infer a strong interaction, which, by linear theory, would imply a FIE in Experiment 2.

Figure 13C shows the state-trace plot, which is very different from its Experiment-1 counterpart. Here it is nonmonotonic: The Inverted stimulus points are displaced leftward—that is, toward poorer Face performance—compared with the Upright stimuli. The interpretation of this finding is, as in Simulation 2, that when two orientation × duration conditions lead to the same performance for Houses, inverted performance is worse for Faces than for Houses. We emphasize that this is what it means, when rejecting unidimensional theory, to say that "inversion affects Faces more than Houses."

Discussion

The Experiment-2 results are unambiguous: The Inverted points are displaced to the left of the Upright points, which indicates a positive FIE. This finding, in conjunction with a lack of FIE in Experiment 1, supports the proposition that when retrieval of known faces does not occur, as with the Experiment-1 computer-generated faces, no FIE will emerge; conversely when retrieval of known faces does occur, as with the Experiment-2 celebrity faces, a FIE will emerge.

A Two-Dimensional Theory

We now describe a multidimensional—specifically a two-dimensional—theory that is designed to account for our data. The theory has two purposes. First, it provides a tutorial demonstration of how specific, quantitative theories may be constructed within the context of general dimensional theory. Second, it provides a bona fide account of the FIE in our real experiments, thereby constituting a foundation for further testing and further research.

We first describe our theory at a general level. We then add specific quantitative assumptions that allow it to make numerical predictions.

A General Theory

Our general theory, which is schematized in Figure 14, draws from numerous nonquantitative theories of face processing. In
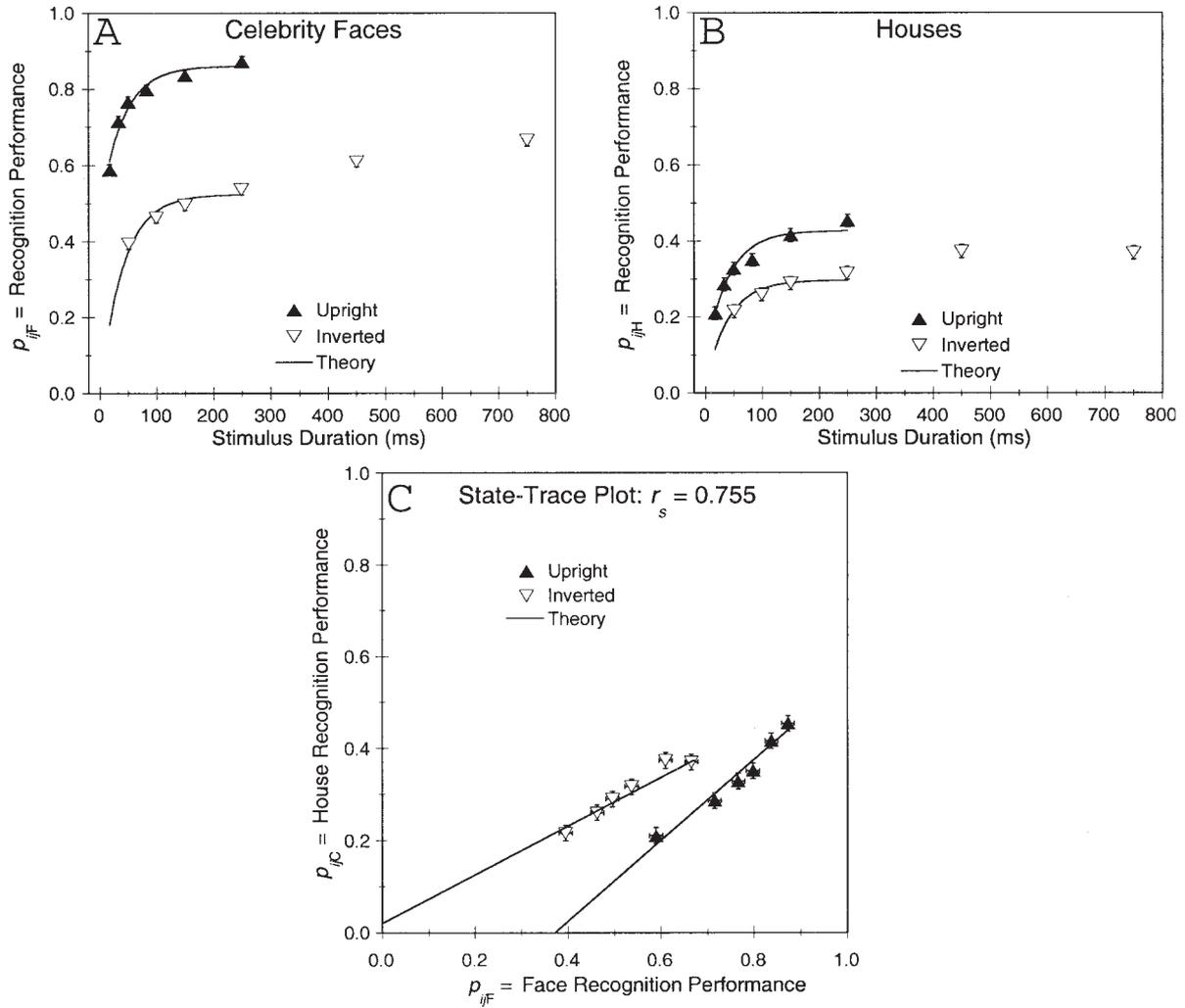
Figure 13. Experiment-2 data. A and B: Basic data, performance as functions of duration for Celebrity Faces and Houses, respectively. C: The state-trace plot. Each data point is based on 1,770 observations. Error bars are standard errors.

particular it incorporates the idea that face processing is often claimed to be done in a way that is different from processing of other visual stimuli. Although the proposed nature of this difference varies from one theory to another, there is a pervasive general notion that face processing is based on two kinds of information. Depending on the specific theory, face processing is assumed to be, more than other kinds of visual stimuli, based on "second-order features," based on "configural information" (e.g., Awh et al.,
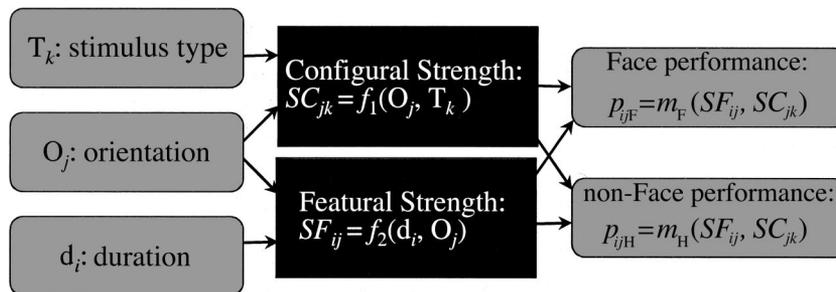

Figure 14. A two-dimensional theory to account for the Experiment-2 results. See the text for further explanation.

2004; Carey & Diamond, 1977; Collinshaw & Hole, 2000; Reed, Stone, Bozova, & Tanaka, 2003; Rhodes, 1988), based on "second-order relational information" (e.g., Diamond & Carey, 1986), based on "holistic" or "gestalt" information (Farah et al., 1995; Farah, Wilson, Drain, & Tanaka, 1998; Sargent, 1984), or "carried out in parallel" (e.g., Bradshaw & Wallace, 1971). This processing mode is contrasted to *featural* or *serial* processing that is assumed to be relatively more important in processing non-Face stimuli.

Given these considerations, we retain the notion of "Strength" but assume two Strength dimensions. The first is featural Strength (SF_{ij}), which depends only on duration (d_i) and orientation (O_j). The second is configural Strength (SC_{jk}), which depends only on orientation (O_j) and stimulus type (T_k). Face and non-Face performance are determined by monotonic functions m_F and m_H of these two kinds of Strength.

This theory is not meant to favor any one of the many proposed theories of face processing and of the FIE that are in the literature but rather represents an attempt to incorporate the general idea of two processing modes—one relatively more important for face processing—that are common to most of the theories. We recognize that the particular terms *configural* and *featural* are used in some theories but not others, but we note that even in the specific theory that we present in the next section, we use the terms only as somewhat arbitrary stand-ins for what we intend to be more general concepts.

A Specific Theory

The next theory-construction step was to specify the functions f_1, f_2, m_F, and m_H (see Figure 14) such that they would produce quantitative predictions. Our choice of these specifications was based on two considerations. First, we designed our theory to be one of a quite successful class of information-processing theories, variants of which have been described by numerous investigators (e.g., Busey & Loftus, 1994; Loftus, Busey, & Senders, 1993; Massaro, 1970; Rumelhart, 1970; Shibuya & Bundesen, 1988), that assume random sampling over time of information from a visual stimulus. Such theories imply some internal measure, for example, "Information" or "Strength", and/or observed performance to increase with stimulus duration by the function $(1 - e^{-kd})$, where d is duration and k is a constant.

The second consideration driving theory design was based on the data obtained in Experiments 1 and 2. We have already shown why the monotonic state-trace plot in Figure 12C and the nonmonotonic state-trace plot in Figure 13C confirm a unidimensional theory and imply a multidimensional theory, respectively. However there is additional information in these state-trace plots that allows stronger inferences: As is apparent in Figures 12C and 13C and as is indicated in Table 3, both the Upright and the Inverted functions are approximately *linear*. Accordingly, we decided to restrict our choice of quantitative theory to those that imply linear state-trace functions. This decision strongly constrains our potential theories.

Equations for featural Strength and configural Strength. To generate quantitative predictions, we specify the various functions of the Figure-14 theory. For featural Strength SF_{ij} in duration condition i and orientation condition j ($j \in \{U, I\}$),

Table 3

Measures of Linearity for State-Trace Plot Functions for Experiments 1–3

Orientation	Slope	Intercept	r^2
Experiment 1			
Upright	0.800	−0.020	0.986
Inverted	0.873	−0.034	0.987
Experiment 2			
Upright	0.851	−0.309	0.956
Inverted	0.615	0.020	0.956
Experiment 3			
Upright	1.687	−0.882	0.982
Inverted	0.795	−0.034	0.934

$$SF_{ij} = \begin{cases} (1 - e^{-d_i/b}) \times Y_U & \text{for Upright stimuli} \\ (1 - e^{-d_i/b}) \times Y_I & \text{for Inverted stimuli} \end{cases}, \quad (12)$$

where d_i is duration. The parameter b, constrained to be positive, is an exponential growth rate common to all stimuli; it may be viewed as representing a low-level characteristic of the system that is ignorant of stimulus meaning. The parameters Y_U and Y_I, constrained to fall between 0 and 1, reflect asymptotic featural Strength that can differ for Upright compared to Inverted stimuli. Note that SF_{ij} must fall between 0 and 1 and can therefore be treated as a probability.

Configural Strength, SC_{jk} for orientation j and stimulus-type k ($k \in \{F, H\}$), is defined to be

$$SC_{jk} = \begin{cases} C_{Uk} & \text{for Upright stimuli} \\ C_{IF} = 0 & \text{for Inverted faces} \\ C_{IH} & \text{for Inverted non-Faces} \end{cases}, \quad (13)$$

where the C_{jk} are free parameters, constrained to fall between 0 and 1, and therefore interpretable as probabilities. The assumption incorporated in Equation 13 that $C_{IF} = 0$ means that no information is acquired from inverted Faces. This assumption is motivated in part by Valentine's (1988) observation that, ". . . configural information is seen as a means of encoding upright faces, but configural information cannot be extracted from an inverted face" (p. 480). Equation 13 also incorporates the idea that configural Strength can be different for Upright Faces compared with Upright non-Faces; that is, C_{UF} does not necessarily equal C_{UH}. In short, featural Strength is the same for Faces and non-Faces, whereas configural Strength can differ in systematic ways for the two stimulus types.

Note that Equation 13 also implies that configural Strength is acquired essentially instantaneously; that is, it does not depend on stimulus exposure duration. In the limit of course, this assumption must be false; *some* time must be required to acquire any kind of information from a stimulus. Realistically, this assumption becomes more precisely that in our experiments, all the configural Strength that will ever be acquired can be acquired from an unmasked stimulus that is as brief as 17 ms.

Performance based on independent featural-Strength and configural-Strength contributions. We assume, as suggested by Collinshaw and Hole (2000), that recognition can be carried out

independently on the basis of either featural or configural Strength, and accordingly we combine them via a probability-summation rule. We also include, because our data require it, a different performance asymptote for Faces and non-Faces. Thus,

$$p_{ijk} = \begin{cases} [SF_{ij} + (1 - SF_{ij})SC_{jF}]Y_F & \text{for Faces} \\ [SF_{ij} + (1 - SF_{ij})SC_{jH}]Y_H & \text{for non-Faces} \end{cases}, \quad (14)$$

where the asymptotes, Y_K ($K \in \{F, H\}$), are free parameters between 0 and 1.

Temporal boundary conditions. We note that this theory is only applicable to information acquisition within a single eye fixation: At the beginning of the second and each subsequent eye fixation on a stimulus, new featural Strength would become available, and the acquisition process embodied in Equation 12 would begin anew. It would be possible to model this process in the present experiments wherein eye fixations have not been explicitly monitored. However, such modeling would be quite cumbersome, and accordingly, we do not attempt to do so. Instead, we confine our data fit to conditions involving durations 250 ms or less. This process excludes only the two longest-duration Inverted conditions of Experiment 2.

Outcomes that imply a unidimensional theory. If C_{UF} and C_{UH} are equal and in addition $C_{IH} = C_{IF} = 0$, the theory collapses into a unidimensional theory. The reasoning that leads to this assertion is as follows. From Equation 12, it is evident that featural Strength, SF_{ij}, is always a function only of duration (i) and orientation (j). Likewise, from Equation 13, when $C_{UF} = C_{UH}$ and $C_{IH} = 0$, SC_{jk} becomes a function only of orientation (k): For Upright stimuli, we can define $C_{UF} = C_{UH} = C_U$, whereas for Inverted stimuli, $C_{IH} = C_{IF} = C_I = 0$. Thus we are left with only two states of C: C_U and C_I. Given these constraints, an overall measure of "Strength," which depends only on duration, i, and orientation, j, can be derived from the bracketed portions of Equation 14 to be

$$S_{ij} = SF_{ij} + (1 - SF_{ij})C_j, \quad (15)$$

where $j \in \{U, I\}$. Performance is then simply

$$p_{ijk} = \begin{cases} S_{ij}Y_F & \text{for Faces} \\ S_{ij}Y_H & \text{for non-Faces} \end{cases}. \quad (16)$$

Note that Equations 15 and 16 are instances of Equations 4 and 5, respectively, which define a unidimensional theory.

State-trace plot linearity. Earlier, we noted that we constrained our choice of theories to those that predicted linear state-trace plots. Appendix A shows that by the theory defined by Equations 12–14, the state-trace functions relating House to Face performance are linear with predicted slope and intercept values for orientation j of

$$\text{Slope}_j = \frac{Y_H(1 - C_{jH})}{Y_F(1 - C_{jF})} \quad (17)$$

and

$$\text{Intercept}_j = \frac{Y_H(C_{jH} - C_{jF})}{(1 - C_{jF})}. \quad (18)$$

Note that a zero intercept value for some orientation, j, implies that $(C_{jH} - C_{jF}) = 0$ or that $C_{jH} = C_{jF}$. In particular, this must happen for both orientation functions when the theory collapses to a unidimensional theory: Here, as described above, $C_{UH} = C_{UF}$, and $C_{IH} = C_{IF} = 0$.

Theory Fits

We first apply the theory to the Experiment-2 data and show that it fits reasonably well. We then apply it to the Experiment-1 data whereupon, as we demonstrate, it collapses into a unidimensional theory, as implied by the monotonic Experiment-1 state-trace plot.

Theory fit to Experiment 2. We fit the theory to the Experiment-2 data using a least mean-square criterion. The best-fitting values of the eight free parameters, along with fit measures are shown in the third row of Table 4. The theoretical predictions are shown as the smooth lines in Figure 13. As can be seen, the fit is reasonably good as shown by the relatively small root-mean-square error (0.022) and the high data-theory correlation ($r^2 = 0.994$). Accordingly, it is appropriate to interpret the estimated parameter values in meaningful ways.

There are several noteworthy aspects of the fits. First, Upright configural Strength, C_{Uk}, is greater for Faces ($C_{UF} = 0.478$) than for Houses ($C_{UH} = 0.166$). Second, Inverted configural Strength is

Table 4
Theory Fits for Experiments 1–3

Experiment	b (ms)[b]	Y_U	Y_I	C_{UF}	C_{UH}	C_{IH}	Y_F	Y_H	Root-mean-square error (df)	Pearson r^2
				Theory parameter					Fit measure[a]	
1[c]	45.3	0.988	0.621	0.294	0.270	0.000	0.444	0.330	0.018 (16)	0.978
1[d]	45.3	0.838	0.551	0.253	0.253	0.000	0.502	0.371	0.018 (17)	0.978
2	40.4	0.734	0.524	0.478	0.166	0.036	1.000	0.550	0.022 (12)	0.994
3	46.9	0.910	0.613	0.619	0.116	0.000	0.931	0.693	0.023 (12)	0.992

Note. b = featural Strength growth rate; Y_U = featural Strength Upright asymptote; Y_I = featural Strength Inverted asymptote; C_{UF} = configural Strength (Upright faces); C_{UH} = configural Strength (Upright non-Faces); C_{IH} = configural Strength (Inverted non-Faces); Y_F = performance asymptote (Faces); Y_H = performance asymptote (non-Faces).
[a] Root-mean-square error is total squared theory-data deviations divided by number of data points minus number of free parameters. Pearson r^2 is between the predicted and observed data points over the 24 duration $\times$ orientation $\times$ stimulus type conditions. The fits in Experiments 2 and 3 exclude the longest-duration conditions for both Faces and non-Faces; accordingly, they are based on 20 rather than 24 conditions. [b] A lower value implies faster feature acquisition rate. [c] C_{UF} and C_{UH} vary independently. [d] C_{UF} and C_{UH} are constrained to be equal.

close to zero for Houses ($C_{IH} = 0.036$). We also note that featural Strength, SF_{ij}, asymptotes at a higher level for Upright stimuli ($Y_U = 0.734$) than for Inverted stimuli ($Y_I = 0.524$), thereby implying (unsurprisingly) that more featural Strength can be acquired from Upright than from Inverted stimuli.

Theory fit to Experiment 1. We fit the same theory to the Experiment-1 data. We did so in two ways. First, we allowed all eight parameters to vary freely, and second, we constrained the Upright configural-Strength values to be equal for Faces and Houses. The results are shown in the first two rows of Table 4. When the C_{Uk}'s are allowed to vary freely, they are estimated to be almost identical ($C_{UF} = 0.294$ and $C_{UH} = 0.270$). When the C_{Uk}'s are constrained to be equal, the resulting estimated parameter values change somewhat, but the fit measures are unchanged. The Inverted-House configural Strength value, C_{IH}, was estimated to be zero in both cases. The theory fits shown in Figure 12 are from the fit in which the C_{Uk}'s are constrained to be equal. It seems safe to conclude that as we inferred from the state-trace plots, a unidimensional version of our theory can account for the Experiment-1 data (where $C_{UF} = C_{UH}$ and $C_{IH} = C_{IF} = 0$) but not for the Experiment-2 data (where $C_{UF} > C_{UH}$).

Discussion: Theory Fits

Our theory fit both the Experiment-1 and the Experiment-2 data quite well. The estimated parameter values are instructive in several ways. First, configural Strength for Inverted Faces was estimated to be zero or close to zero for both experiments. This observation extends to Houses, Valentine's (1988) assertion that configural Strength can reasonably be construed to be zero for Inverted Faces. Confirming our conclusions based on the state-trace plots, the parameter estimates indicated that two dimensions, featural Strength and configural Strength, were necessary and sufficient to account for the Experiment-2 data (C_{UF} was estimated to be greater than C_{UH}), whereas in Experiment 1, the near equality of C_{UF} and C_{UH} (in conjunction with $C_{IF} = C_{IH} = 0$) indicated that a single-dimension theory suffices to account for the data.

Experiment 3: Celebrities and Cityscapes

Experiment 2 indicated that when face retrieval is required in the study phase of a recognition experiment, a FIE emerges. Experiment 3 was designed to extend and replicate Experiment 2 under somewhat different conditions: The non-Face stimuli consisted of pictures of heterogeneous Cityscapes rather than the homogeneous Houses used in Experiments 1 and 2.

We had, in particular, three goals in Experiment 3. First, as indicated, we wanted to replicate Experiment 2 and show that it generalizes to a stimulus class beyond Houses. Second, our quantitative theory was developed *post hoc*, based in part on the Experiment-1 and Experiment-2 data patterns. Accordingly, it is desirable to confirm that it can also fit data from a new experiment, that is, that with its eight free parameters, it was not simply fitting noise from Experiments 1 and 2. Third, both Experiments 1 and 2 indicated that within the context of our theory, configural Strength was zero for Houses as well as for Faces. Is configural Strength equal to zero for inverted stimuli from *any* mono-oriented stimulus class? We cannot, of course, answer this question completely without carrying out the impossible task of testing all such stim-

ulus classes. However, generalization to a quite different stimulus class, Cityscapes, would strongly suggest a phenomenon that is fairly general, not specific to homogeneous stimuli such as Faces and Houses.

Method

Observers. A total of 160 University of Washington undergraduates participated in 24 groups of 6 to 11 observers per group, in exchange for course credit.

Stimuli. A new stimulus set, Cityscapes was created. These were 144 photos of skylines, individual buildings, street scenes, and bridge–water scenes from numerous cities (excluding Seattle), all rendered as grayscale drawings 450 pixels high and 400 pixels wide. Examples are provided in the bottom row of Figure 11.

Apparatus. The apparatus was the same as that used in Experiments 1 and 2.

Design, procedure, and counterbalancing. Procedures and counterbalancing measures were the same as in Experiment 2 except that only 24 groups were run: In all groups, Faces were shown in the first subsession, followed by Cityscapes in the second subsession. This was an efficiency measure implemented because Experiments 1 and 2 indicated no stimulus-type order effects, and the additional power gained by doubling the number of observers didn't appear to be needed.

Results

The results are shown in Figure 15, which is organized like Figures 12 and 13. For the three durations that are common to both orientations, the orientation effect is greater for Faces than for Cityscapes: The interaction's mean magnitude is 0.161 ± 0.028.

Figure 15C shows the state-trace plot, which like its Experiment-2 counterpart, is nonmonotonic: The Inverted stimulus points are shifted leftward—that is, toward poorer Face performance—compared with the Upright stimuli. Again, we reject unidimensional theory and conclude that there is a FIE, that is, that inversion affects Faces more than Cityscapes.

We fit our quantitative theory to the Experiment-3 data. The best-fitting parameter values are shown in the fourth row of Table 4, and the predictions are shown as the smooth curves through the Figure-15 data points. The fits and the parameter values are similar to those of Experiment 2.

Discussion

Experiment 3 nicely replicates Experiment 2, thereby providing another demonstration of a FIE when face retrieval is required during the study phase of a recognition experiment. The estimated parameter values are qualitatively quite similar to those of Experiment 2. Of most note is that configural Strength for Inverted Cityscapes is estimated to be zero, thereby adding credence to the proposition that configural Strength, whatever it is defined to be, is not acquired from inverted stimuli in general.

General Discussion

In this article, we have developed two quite distinct themes. The first focuses on a comparison between linear theory and dimensional theory, and the second focuses on a specific question about the circumstances under which the FIE does or does not emerge.
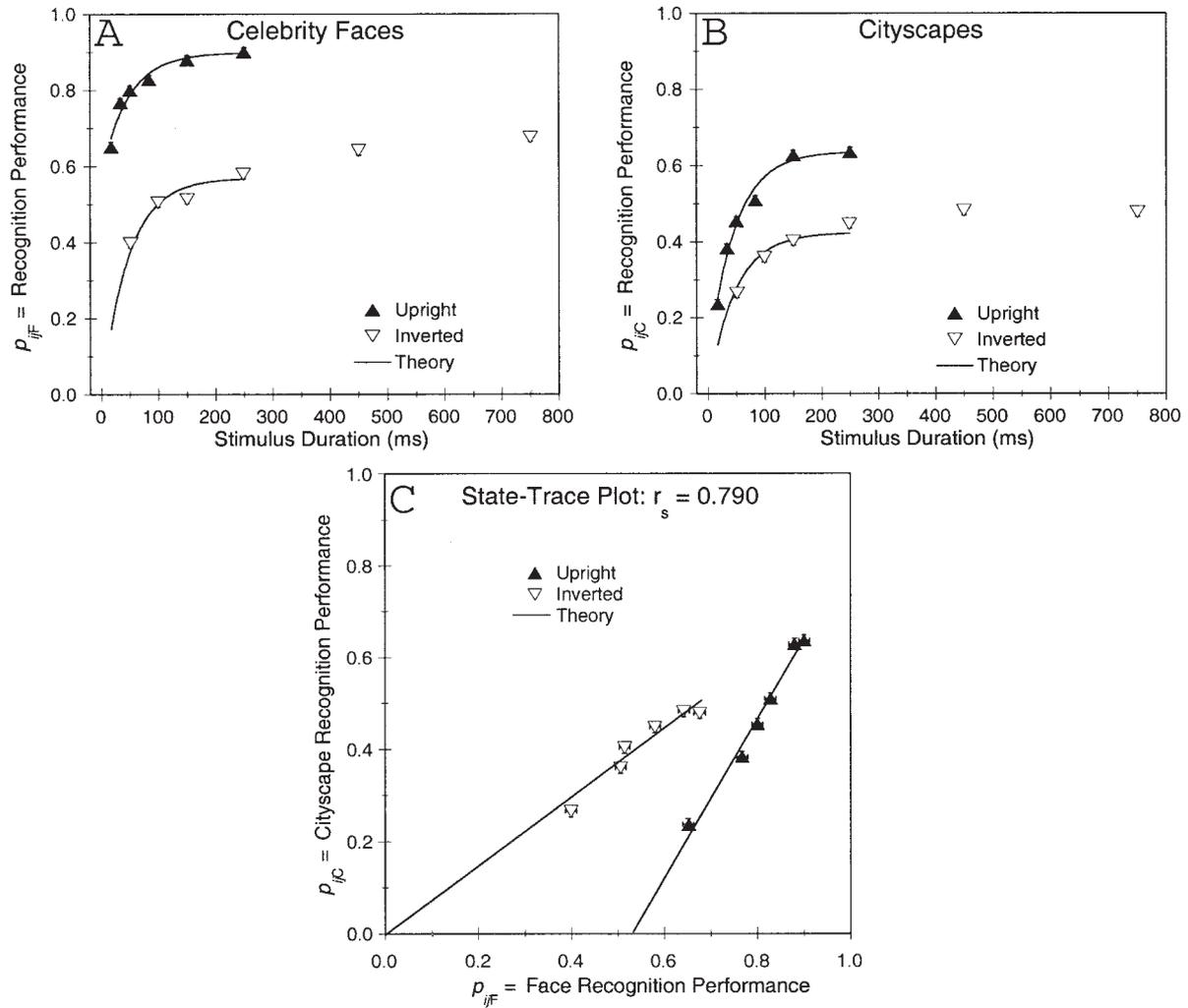
Figure 15. Experiment-3 data. A and B: Basic data, performance as functions of duration for Celebrity Faces and Cityscapes, respectively. C: The state-trace plot. Each data point is based on 960 observations. Error bars are standard errors.

These two themes are united by a critical observation: One can make a decision about whether some effect, for example, a FIE, has emerged only within the context of some fundamental theory. Simulation 1 and Experiment 1 are pivotal in demonstrating the consequences of this observation: In both, the data patterns were such that a FIE was inferred to have emerged if one interpreted the data within the context of linear theory, but was inferred to have not emerged if one interpreted the data within the context of dimensional theory. If this kind of data pattern had been observed only in Experiment 1, it might have been dismissed as some kind of weird statistical fluke. That the same data pattern was observed in the noise-free Simulation-1 data demonstrates that the conundrum raised by it is real.

In what follows, we first summarize the relations between linear theory and dimensional theory. We conclude, as we tentatively did following discussion of our simulations, that dimensional theory is more flexible and more realistic than linear theory as a context for thinking about psychological problems, for designing experiments,

and for analyzing and interpreting data. After offering these arguments, we turn to the FIE and summarize what we have learned about it.

Linear Theory and Dimensional Theory

In this section, we compare the benefits of using linear theory versus dimensional theory as a foundation for data interpretation.

Linear Theory Is a Special Case of Dimensional Theory

Earlier we characterized dimensional theory as comprising two components: a mapping from the independent variables to dimensional values (Equation 2) and a mapping from the dimensional values to the dependent variables (Equation 3). Linear theory can be construed as having the same structure. In the most general case, there are $2^M - 1$ dimensions corresponding to the main and interaction effects of the M independent variables. In the present

case of $M = 3$ independent variables, there are $2^3 - 1 = 7$ dimensions, and we can recast Equation 1, the linear-theory representation in terms of the dimensional-theory representation of Equations 2 and 3. Assume some condition C_{ijk} corresponding to level i of IV_1, level j of IV_2, and level k of IV_3. Then, we can denote the value of the dth dimension for C_{ijk} as V_{dijk}, and the instantiation of Equation 2 becomes

$$
\begin{aligned}
V_{1ijk} &= f_1(IV_1, IV_2, IV_3) = f_1(IV_1) = \alpha_i \\
V_{2ijk} &= f_2(IV_1, IV_2, IV_3) = f_2(IV_2) = \beta_j \\
V_{3ijk} &= f_3(IV_1, IV_2, IV_3) = f_3(IV_3) = \gamma_k \\
V_{4ijk} &= f_4(IV_1, IV_2, IV_3) = f_4(IV_1, IV_2) = \lambda_{ij} \\
V_{5ijk} &= f_5(IV_1, IV_2, IV_3) = f_5(IV_1, IV_3) = \gamma_{ik} \\
V_{6ijk} &= f_6(IV_1, IV_2, IV_3) = f_6(IV_1, IV_2) = \kappa_{jk} \\
V_{7ijk} &= f_7(IV_1, IV_2, IV_3) = f_7(IV_1, IV_2, IV_3) = \xi_{ijk}
\end{aligned}
\quad . \quad (19)
$$

The seven values then simply sum to produce a single dependent variable for C_{ijk}; so the instantiation of Equation 3 is

$$
p_{ijk} = \sum_{d=1}^{D} V_{dijk}. \quad (20)
$$

This treatment makes it clear that Equation 19 is a special case of Equation 2, and Equation 20 is a special case of Equation 3. Therefore, linear theory is a special case of dimensional theory.[2] It is, indeed, a special case that does not have a great deal of psychological validity, in that linearity is rarely a property of psychological processes. As we have discussed, for example, probability correct is not gracefully predicted by linear processes in which sums of effects, unlike probabilities, are not constrained to lie between 0.0 and 1.0. It is for this reason that the simulated probabilities in our simulations could not themselves issue from a linear structure, which, in turn, was responsible for the chaotic behavior of the results.

In our view, this stark demonstration of the relation between linear theory and dimensional theory underscores the point that the more general and flexible dimensional theory, rather than the more specific and inflexible linear theory, should be viewed as a starting point for construction of theory within psychology, or, for that matter, within most sciences. This point should come as no surprise. The literature is bespeckled with warnings about limitations of linear theory (e.g., Loftus, 2002), particularly when interpretations of nonordinal interactions are at issue (e.g., Bogartz, 1976; Loftus, 1978). The linear-theory based inconsistencies that we have reported in conjunction with Simulations 1–3 are simply explicit manifestations of these limitations.

Linear theory and dimensional theory as models of the FIE. A major point of this article was to determine circumstances under which a FIE does and does not emerge. To do so requires a specific model of when there is no FIE, which can then be either confirmed or disconfirmed by the data. Both linear theory and dimensional theory offer a means way of modeling lack of a FIE: Within the context of linear theory, lack of FIE is specified by the absence of an orientation $\times$ stimulus type interaction, whereas within the context of dimensional theory, lack of FIE is embodied in a unidimensional theory. How useful are these two models? Some insights can be gleaned by considering the formal relations between linear theory and unidimensional theory.

The specific unidimensional theory that we have considered throughout this article is one in which two independent variables, duration and orientation, combine into a single measure (Strength) that then determines performance for both Faces and Houses. In Appendix B, we elaborate on the relations between linear theory, dimensional theory, and the particular unidimensional theory that we have considered. In Appendix B, we demonstrate that when, within the context of linear theory, there are any nonzero interactions involving stimulus type, linear theory is a multidimensional theory. The converse, as implied by the results of Appendix B and as illustrated in Simulation 3, is that when there are no interactions involving stimulus type, linear theory is a unidimensional theory. This is a useful result: Confirmation of unidimensional theory, in the form of a monotonic state-trace plot (see Figures 2, 12, and especially 9), implies the existence of a linear structure involving no stimulus-type interactions—which, of course, implies no FIE, even within the context of linear theory.

Suppose, in other words, that one observes a classical stimulus type $\times$ orientation interaction in the data. As we have shown, most specifically in Simulation 3, such an observation does not necessarily imply that there is an interaction in the underlying linear structure: Any nonlinear mapping from the underlying linear structure to the observed data almost certainly introduces some kind of interaction even if there was none there to begin with. However, an observed monotonic state-trace plot implies that there are no underlying interactions involving stimulus type.

Reassessing data. The foregoing is such a critical point that we would like to elaborate on it in the form of revisiting some of the data presented earlier. First, consider Simulation 3. Suppose that Simulation 3 were a real experiment, that is, that the Figure-8 data were observed in real life. Again, supposing that we had essentially infinite statistical power, we would have inferred a small negative FIE, which is seen most directly in Figure 8D. Switching now to our actual status of the data's creators, we know that the structure underlying the data, expressed in Equation 10, included no FIE and indeed no interactions at all involving stimulus type: It was only the nonlinear (Gaussian) measurement mapping that produced the FIE observed in the data.

However—and this is what we wish to emphasize—even if we did not know about the underlying FIE-less linear structure to begin with, we could have inferred it from the monotonic state-trace plot of Figure 9. Indeed, the same point holds true with the Simulation-1 data (see Figure 1). Even though the Simulation-1 structure was not originally expressed as a linear structure, we know from the monotonic state-trace plot (see Figure 2) that it could have been expressed as a FIE-less linear structure—in fact a linear structure with no stimulus-type interaction terms at all— along with a nonlinear mapping from the structure to performance.

Finally, we can make the same inference with our real Experiment-1 data (see Figure 12). Even though, as in Simulation 1, there is an observed FIE in the form of an orientation $\times$ stimulus type interaction, we can infer from the monotonic state-trace plot (see Figure 12C) that there exists an underlying FIE-less linear structure, along with a nonlinear mapping from this structure to the data. This is a very powerful conclusion and affords an alternative perspective on the assertion that despite the FIE observed in the data, there is no FIE in the underlying structure.

---

[2] We thank John Dunn for pointing out this relation between linear theory and dimensional theory, along with its implications.

Difficulties in Interpreting Data From Simpler Designs

The simulations and experiments in this article all involve designs in which stimulus duration is manipulated along with the two other independent variables—orientation and stimulus type—that are central in assessing the FIE. However, an investigator embarking on a standard investigation of the FIE would normally consider inclusion of duration, or any other independent variable, to be superfluous. Of critical interest, from the investigator's perspective, is the interaction between orientation and stimulus type. Thus the standard FIE investigation would entail a simple 2×2 design in which only orientation and stimulus type are manipulated. Such 2×2 designs are a staple of FIE investigations and of psychological research in general. One might ask how the logic that we have described—particularly the logic underlying state-trace plots and associated tests of dimensional theories—would apply to such simpler designs.

Scaling problems (serious ones): Comparisons of difference scores are not robust. The answer is that a 2×2 design is insufficient to test theories of any complexity. The reason is this: Most scales used in psychology are generally valid only up to a monotonic transformation; that is, two potential dependent variables, for example, probability and d'—or a potential dependent variable and an internal construct, for example, probability and "Strength"—can, in general, be assumed to be only monotonically related. Therefore, differences on such scales cannot be meaningfully compared; two differences that are, say, equal in one scale are usually unequal in some other equally valid but nonlinearly related scale, and vice versa. This leads directly to the difficulties in interpretations of interactions that have been discussed in the past (e.g., Bogartz, 1976; Loftus, 1978): An interaction observed with some particular dependent variable, for example, probability correct, can disappear or even reverse when expressed either in terms of some other dependent variable or in terms of some hypothetical internal construct.

The virtues of equality. So, comparisons of differences are unstable across monotonic transformations. In contrast, *equality* is preserved across all monotonic transformations: Two conditions that are equal on one scale, for example, probability, are likewise equal on any other monotonically related scale, be it another dependent variable (e.g., d') or some internal construct of interest (e.g., "Strength"). Therefore, robust conclusions follow when performance is equated for one pair of conditions—Upright and Inverted Houses, say—and then compared for another pair of conditions—Upright and Inverted Faces. Corresponding equality of Upright and Inverted Faces would unequivocally imply one conclusion, whereas corresponding inequality of Upright and Inverted Faces would unequivocally imply another conclusion.

To equate levels on a variable such as orientation, however, requires trading off with levels of some other variable, for example, duration or contrast, such that the intrinsically better level of orientation (Upright) is combined with an intrinsically worse level of duration (short), whereas the intrinsically worse level of orientation (Inverted) is combined with an intrinsically better level of duration (longer). By using multiple duration levels, state-trace curves are mapped out such that equality of Upright and Inverted Houses can be observed over a wide range of performance levels, and relative performance for Faces can be compared across this range (see, e.g., Figures 2 and 5). Robust conclusions then ensue:

In the present examples, no FIE is inferred if Upright and Inverted Face performance is equal when Upright and Inverted House performance is equal (see Figure 2), whereas a standard, positive FIE is inferred when equal Upright and Inverted House performance is associated with poorer Inverted compared with Upright Face performance (see Figure 5). We note that other investigators have manipulated other variables to attain equality. For instance, Biederman and Tsao (1979) equated Chinese and American observers on a control task in a Stroop paradigm by eliminating fast American observers and slow Chinese observers.

Impoverished information with 2×2 designs. Another way of conceptualizing the difficulty in interpreting standard 2×2 designs is by reference to Figure 1D. Assuming a 2×2 design, the exposure duration would have to be selected to have been *something*, and the observed magnitude of the FIE would then depend on which particular duration value had been selected. As is evident in Figure 2D, the observed FIE would have been observed to be nil had the selected duration been less than about 100 ms and would have been greater with greater selected durations. As if this were not bad enough, even *these* conclusions are tenuous because as is apparent in Figure 3, even the relation between duration and FIE magnitude is not very robust.

Nested Dimensional Theories

We now turn to a brief review of the usefulness of thinking in terms of dimensional theories. The complex designs of Simulations 1–3 and Experiments 1–3 have guided the construction of three nested dimensional theories—that is, theories devised such that each successive theory is a special case of the preceding one—within whose context we have sought to understand the results of our experiments. In what follows, we summarize how these theories provide an increasingly specific understanding of the nature of the FIE. In the present experiments, these theories were used to investigate whether the FIE is characteristic of storage of face information and/or retrieval of face information; however, the theories, or variants of them, could be used to assess any question about face processing, the FIE, and the circumstances under which the FIE does or does not emerge.

Theory I: Generic Dimensional Theory

Theory I, the most general, is generic dimensional theory (Equations 2 and 3), wherein we simply assume that the independent variables in our experiments result in some number of internal dimensions, the values of which then determine performance. Like generic linear theory, Theory I is not in and of itself particularly useful in understanding the FIE or any other process; it simply provides a theoretical shell within which more specific theories may be formulated.

Theory II: Configural and Featural Strength

We deem that actual insight into the nature of the FIE begins with Theory II: the two-dimensional theory presented in Figure 14. Here, we identify constructs that have been used in the past to explain face processing in general and the FIE in particular. Specifically, we postulated two kinds of Strength—drawn from many past accounts of face processing—that are assumed to accrue

during inspection of a visual stimulus. In this incarnation of the theory, we assume that featural Strength accrues over time in a manner that is indifferent to the nature of the stimulus but nevertheless depends on stimulus duration and stimulus orientation. Configural Strength is, in keeping with the notion of face-processing "specialness," allowed to be different for Faces and non-Faces as well as for different orientations.

Theory III: A Quantitative Version of Theory II

In Theory III, we fleshed out these assumptions with specific functions that allowed quantitative fits to the data from our three real experiments. Our decisions about exactly how to structure this strong version of the theory were somewhat but not entirely arbitrary: As noted, they were constrained by the nature of past quantitative visual-information-acquisition theories and independently by the observed approximately linear forms of the Experiments 1–3 state-trace plot functions. The theory fits to the data sets were quite good (see Figures 12, 13, and 15), and the estimated parameter values (see Table 4) therefore constitute a useful tool for interpreting the subtle and complex language spoken by the data.

Face Recognition and the FIE

Experiments 1–3 were designed to assess Valentine's (1988) assertion that a FIE does not emerge when orientation is manipulated in the study phase of a visual-recognition experiment. What can we conclude about this issue?

In our experiments, we manipulated orientation only during the study phase of a visual-recognition procedure: A FIE did not emerge when unfamiliar computer-generated Faces were compared with Houses (Experiment 1), but a FIE did emerge when photographs of familiar (celebrity) Faces were compared with Houses (Experiment 2) or Cityscapes (Experiment 3). This finding is consistent with the proposition that the FIE emerges when familiar faces are retrieved from memory, as almost certainly happens during the study phase of Experiments 2 and 3, but does not emerge when unfamiliar faces are encoded for subsequent recognition, as must happen during the study phase of Experiment 1. Such a proposition is reasonable given the findings that processing of familiar faces differs in many respects from processing of unfamiliar faces (see, e.g., Bruce, Henderson, Newman, & Burton, 2001; Hancock, Bruce, & Burton, 2000).

However, this conclusion must be tempered by (at least) two caveats. First, this pattern of results may not be unique to faces but, rather, may characterize any stimulus set that can occur in a familiar or an unfamiliar configuration. An example is buildings, which, like faces, come in unfamiliar versions (i.e., most buildings) and familiar versions (e.g., the Empire State Building, the Taj Mahal, etc.). Although, to our knowledge, the experiment has not been done, it seems perfectly plausible that a "building-inversion effect" would emerge when familiar but not unfamiliar buildings are compared with a uniform set of unfamiliar stimuli, such as houses.

The second caveat is that as is evident from even a quick glance at Figure 11, the computer-generated faces used in Experiment 1, though guaranteed to be unfamiliar, differ in many respects from the celebrity faces used in Experiments 2 and 3. Possibly our observers were not treating these computer-generated faces as actual faces, and that is why they didn't differ from houses with respect to the inversion effect. Numerous research projects suggest themselves to test this possibility.

On the Prevalence of the FIE

In the introduction to this article, we provided numerous references to reports of a FIE measured by the traditional stimulus type × orientation interaction. Generally speaking, these findings appear to have formed a robust and reasonably harmonious pattern that has assisted us in understanding face processing. Are we suggesting that because of the interpretational difficulties implied by Simulations 1–3, these findings simply be declared invalid until they are reinvestigated within the context of dimensional theory? And where, by the way, are all those *negative* FIE's whose existence is suggested by the chaotic demonstrations illustrated in Figures 3, 6, and 10?

These questions do not have obvious answers. We do, however, make a general observation, which is that the greater the magnitude of the FIE as characterized by the traditional interaction, the less likely that it will be associated with a lack of a FIE as demonstrated by confirmation of a unidimensional theory: Note that the Experiment-1 data involved quite a small-magnitude interaction, whereas the data from Experiments 2 and 3 involved much larger-magnitude interactions. As for the lack of negative FIE's in the literature, there are (at least) two plausible explanations. The first is that we have made faulty assumptions—perhaps, unlike our simulation parameters, the structural parameters that govern real-life FIE data are constrained in such a way as to not allow negative FIE's. The second explanation rests on the "file-drawer problem" (Rosenthal, 1979): Given no a priori reason to expect a negative FIE, it is unlikely that anyone who obtained such a result would publish it.

Conclusions

In this article, we have tried to make two main points. The first, which in our view is the less important, has to do with the FIE: We have provided evidence suggesting that a FIE emerges when known faces are retrieved from memory but not when unknown faces are encoded in memory. Although consistent with prior conjectures, our data are nevertheless subject to a number of alternative interpretations and should thus be viewed as a starting point for future research projects rather than as grounds for unambiguous conclusions about the nature of face processing.

The second point, which we consider the more important, is quite radical: We have asserted and have tried to demonstrate that standard, off-the-shelf linear theory is highly limited in its usefulness as a foundation for conceptualizing psychological problems and analyzing psychological data. We have described an alternative, dimensional theory, which as we have shown, subsumes linear theory: Dimensional theory is more general both in terms of the functions mapping independent variables to presumed internal constructs and in terms of the functions mapping the presumed internal constructs to the observed dependent variables.

As we have indicated, tests of specific dimensional theories generally require relatively complex designs; the standard 2 × 2, for example, is generally inadequate. We are convinced, however, that such complex experiments, carried out in conjunction with

carefully specified dimensional theories, will bear fruit in the form of deeper and more robust understandings of all manner of psychological phenomena.

References

Anderson, N. H. (1974). Information integration theory: A brief survey. In D. Krantz, R. C. Atkinson, D. Luce, & P. Suppes (Eds.), *Contemporary developments in mathematical psychology* (Vol. 2., pp. 236–305). New York: Freeman.

Anderson, N. H. (1979). Algebraic rules in psychological measurement. *American Scientist, 67,* 555–563.

Anderson, N. H. (1981). *Foundations of information integration theory.* New York: Academic Press.

Awh, E., Serences, J., Laurey, P., Dhaliwal, H., van der Jagt, T., & Dassonville, P. (2004). Evidence against a central bottleneck during the attentional blink: Multiple channels for configural and featural processing. *Cognitive Psychology, 48,* 95–126.

Bamber, D. (1979). State trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19,* 137–181.

Biederman, I., & Tsao, Y. C. (1979). On processing Chinese ideographs and English words: Some implications from Stroop task results. *Cognitive Psychology, 11,* 125–132.

Bogartz, R. S. (1976). On the meaning of statistical interactions. *Journal of Experimental Child Psychology, 22,* 178–183.

Bradshaw, J. L., Taylor, M. J., Patterson, K., & Nettleton, N. C. (1980). Upright and inverted faces, and housefronts in two visual fields. *Journal of Clinical Neuropsychology, 2,* 245–257.

Bradshaw, J. L., & Wallace, G. (1971). Models for the processing and identification of faces. *Perception & Psychophysics, 9,* 443–448.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology: Applied, 7,* 207–218.

Bruyer, R., & Velge, V. (1981). Lesions cérébrale unilateral et trouble direct examination la perception des visage: Specificité du deficit? *Acta Neurological Belgica, 81,* 321–332.

Busemeyer, J. R., & Jones, L. E. (1983). Analysis of multiplicative combination rules when the causal variables are measured with error. *Psychological Bulletin, 88,* 237–244.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review, 101,* 446–469.

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review, 7,* 26–48.

Carey, S., & Diamond, R. (1977, January 21). From piecemeal to configural representation of faces. *Science, 195,* 312–314.

Collinshaw, S. M., & Hole, G. J. (2000). Featural and configural processes in the recognition of faces of different familiarity. *Perception, 29,* 893–909.

Coltheart, M. (1985). Cognitive neuropsychology and the study of reading. In M. I. Posner & O. S. M. Marin (Eds.), *Attention & performance XI* (pp. 3–37). Hillsdale, NJ: Erlbaum.

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

Dunn, J. C., & James, R. N. (2003). Signed difference analysis: Theory and application. *Journal of Mathematical Psychology, 47,* 389–416.

Dunn, J. C., & Kirsner, K. (1988). Discovering functionally independent mental processes: The principle of reversed association. *Psychological Review, 95,* 91–101.

Ellis, H. D., & Shepherd, J. W. (1975). Recognition of upright and inverted faces presented in the left and right visual fields. *Cortex, 11,* 3–7.

Farah, M. J., Tanaka, J. N., & Drain, M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 628–634.

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

Garner, W. R. (1974). *The processing of structure and information.* New York: Wiley.

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.

Hancock, P. J. B., Bruce, V., & Burton, A. M. (2000). Recognition of unfamiliar faces. *Trends in Cognitive Science, 4,* 330–337.

Harley, E. M., Dillon, A., & Loftus, G. R. (2004). Why is it difficult to see in the fog? How contrast affects visual perception and visual memory. *Psychonomic Bulletin & Review, 11,* 197-231.

Hays, W. (1973). *Statistics for the social sciences* (2nd ed.). New York: Holt.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition, 68,* B1–B11.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement.* New York: Academic Press.

Krantz, D. H., & Tversky, A. (1971). Conjoint measurement analysis of composition rules in psychology. *Psychological Review, 78,* 151–169.

Kruskal, J. B. (1964). Multidimensional scaling: A numerical method. *Psychometrica, 29,* 115–129.

Leehey, S., Carey, S., Diamond, R., & Cahn, A. (1978). Upright and inverted faces: The right hemisphere knows the difference. *Cortex, 14,* 411–419.

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition, 6,* 312–319.

Loftus, G. R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 396–405.

Loftus, G. R. (2002). Analysis, interpretation, and visual presentation of data. In H. Pashler (Ed.) & J. Wixted (Vol. Ed.), *Stevens' handbook of experimental psychology* (3rd ed., Vol. 4, pp. 339–390). New York: Wiley.

Loftus, G. R., & Bamber, D. (1990). Learning–forgetting independence: Unidimensional memory models and feature models: Comment on Bogartz (1990). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 916–926.

Loftus, G. R., Busey, T. A., & Senders, J. W. (1993). Providing a sensory basis for models of visual information acquisition. *Perception & Psychophysics, 54,* 535–554.

Loftus, G. R., & Irwin, D. E. (1998). On the relations among different measures of visible and informational persistence. *Cognitive Psychology, 35,* 135–199.

Massaro, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review, 77,* 557–567.

Palmer, J. C. (1986a). Mechanisms of displacement discrimination with a visual reference. *Vision Research, 26,* 1939–1947.

Palmer, J. C. (1986b). Mechanisms of displacement discrimination with and without perceived movement. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 411–421.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Phelps, M. T., & Roberts, W. A. (1994). Memory for pictures of upright and inverted primate faces in humans (*Homo sapiens*), squirrel monkeys (*Saimiri sciureus*), and pigeons (*Columba livia*). *Journal of Comparative Psychology, 108,* 114–125.

Phillips, R. J., & Rawles, R. E. (1979). Recognition of upright and inverted faces: A correlational study. *Perception, 8,* 577–583.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science, 14,* 302–308.

Rhodes, G. (1988). Looking at faces: First-order and second-order features as determinants of facial appearance. *Perception, 17,* 43–63.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Rumelhart, D. E. (1970). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology, 7,* 191–218.

Sargent, J. (1984). An investigation into component and configural processes underlying face perception. *British Journal of Psychology, 75,* 221–242.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrica, 27,* 125–140.

Shibuya, H., & Bundesen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 591–600.

Tversky, A., & Russo, J. E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology, 6,* 1–12.

Valentine, T. (1988). Upside-down faces: A review of the effect of inversion upon face recognition. *British Journal of Psychology, 79,* 471–491.

Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica, 61,* 259–273.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141–145.

Yin, R. K. (1970). Face recognition by brain-injured patients: A dissociable ability? *Neuropsychologia, 8,* 395–402.

Appendix A

Slopes and Intercepts of the Linear State-Trace Plots Implied by Equations 12–14

This appendix concerns the quantitative theory specified by Text Equations 12–14. Our goals are (a) to prove that by this theory, the Upright and Inverted state-trace functions relating House performance to Face performance are linear and (b) to derive the slopes and intercepts of these functions. We prove linearity by deriving the expression for the slope and demonstrating that slope does not depend on duration.

Assume two duration values, d_1 and $d_2 > d_1$. Text Equation 14 provides predicted performance,

$$p_{ijk} = [SF_{ij} + (1 - SF_{ij})C_{jk}]Y_k = Y_k SF_{ij}(1 - C_{jk}) + Y_k C_k, \quad (A1)$$

where, as usual, $j \in \{F, I\}$ and $k \in \{F, H\}$. Here i is restricted to $i \in \{1, 2\}$ for the two durations under consideration. Note that SF_{ij} and C_{jk} are obtained from Text Equations 12 and 13.

The slope is the ratio $(p_{2jH} - p_{1jH})/(p_{2jF} - p_{1jF})$. Substituting the right side of Equation 1 for the p_{ijk}s,

$$\text{Slope} = \frac{[Y_H SF_{2j}(1 - C_{jH}) + Y_H C_{jH}] - [Y_H SF_{1j}(1 - C_{jH}) + Y_H C_{jH}]}{[Y_F SF_{2j}(1 - C_{jF}) + Y_F C_{jF}] - [Y_F SF_{1j}(1 - C_{jF}) + Y_F C_{jF}]}$$

$$= \frac{Y_H(1 - C_{jH})(SF_{2j} - SF_{1j})}{Y_F(1 - C_{jF})(SF_{2j} - SF_{1j})}$$

or

$$\text{Slope} = \frac{Y_H(1 - C_{jH})}{Y_F(1 - C_{jF})}, \quad (A2)$$

which is the expression indicated in the text. Because the Equation A2 expression for slope contains no terms involving d_i—it depends only on parameters relevant to orientation, *j,* and stimulus type, *k*—the slope will be the same no matter what values of d_1 and d_2 are chosen, and the state-trace functions must therefore be linear for both orientation values.

The intercept may be derived by picking performance values for Houses, p_{ijH}, and Faces, p_{ijF}, and noting that because the functions are linear,

$$\text{Intercept} = p_{ijH} - \text{Slope} \times p_{ijF}; \quad (A3)$$

or, substituting Equations A1 and A2 into Equation A3,

$$\text{Intercept} = [Y_H SF_{ij}(1 - C_{jH}) + Y_H C_{jH}]$$

$$- \left[\frac{Y_H(1 - C_{jH})}{Y_F(1 - C_{jF})}\right][Y_F SF_{ij}(1 - C_{jF}) + Y_F C_{jF}]$$

$$= [Y_H SF_{ij}(1 - C_{jH}) + Y_H C_{jH}] + Y_H C_{jH}]$$

$$- \left[\frac{Y_{H(1 - C_{jH})} Y_F SF_{ij}(1 - C_{jF})}{Y_F(1 - C_{jF})}\right] - \left[\frac{Y_H(1 - {jH}) C_{jF} Y_F}{Y_F(1 - C_{jF})}\right]$$

$$= [Y_H SF_{ij}(1 - C_{jH}) + Y_H C_{jH}] - Y_H SF_{ij}(1 - C_{jH})$$

$$- \left[\frac{Y_H(1 - C_{jH}) C_{jF}}{(1 - C_{jF})}\right]$$

$$= Y_H C_{jH} - \frac{Y_H C_{jF}(1 - C_{jH})}{(1 - C_{jF})}$$

$$= \frac{(1 - C_{jF})Y_H C_{jH} - Y_H C_{jF}(1 - C_{jH})}{(1 - C_{jF})};$$

or, finally,

$$\text{Intercept} = \frac{Y_H(C_{jF} - C_{jH})}{(1 - C_{jF})},$$

which is the expression for intercept indicated in the text.

Appendix B

On the Relations Between Linear Theory and Dimensional Theory

Our goal here is to define relations between linear theory and dimensional theory. For convenience, we again assume no statistical error.

Linear theory is defined by

$$p_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \kappa_{jk} + \lambda_{ij} + \nu_{ik} + \xi_{ijk}, \quad \text{(B1)}$$

where i, j, and k index duration, orientation, and stimulus type, respectively.

To define unidimensional theory, we slightly rewrite Text Equation 4 to obtain Strength, S_{ij}, in duration $\times$ orientation condition ij,

$$S_{ij} = f(d_i, O_j), \quad \text{(B2)}$$

and we similarly rewrite Text Equation 5 to obtain performance,

$$p_{ijk} = m_k(S_{ij}), \quad \text{(B3)}$$

where $k \in \{F, H\}$. Thus, performance depends on both Strength, S_{ij}, and stimulus type as instantiated in the two monotonic functions, m_k.

We now attempt to recast linear theory as a unidimensional theory. We first define Strength to be

$$S_{ij} = \mu + \alpha_i + \beta_j + \lambda_{ij}. \quad \text{(B4)}$$

Thus Strength defined by linear theory (Equation B4) is a special case of Strength defined by unidimensional theory (Equation B2). To predict performance, we substitute Equation 4 into Equation B1 to obtain

$$p_{ijk} = S_{ij} + \gamma_k + \kappa_{jk} + \nu_{ik} + \xi_{ijk} = m_k(S_{ij}, d_i, O_j). \quad \text{(B5)}$$

By Equation B5, therefore, performance depends not only on Strength, S_{ij}, and stimulus type, γ_k, but also on κ_{jk}, ν_{ik}, and ξ_{ijk}, that is, on duration and orientation. Therefore, unless κ_{jk}, ν_{ik}, and ξ_{ijk} are all zero, linear theory cannot be a unidimensional theory; it must be multidimensional.

In the most general case in which all interaction terms are nonzero, linear theory is a three-dimensional theory. We define the first dimension, Strength, by Equation B4. Two additional dimensions, which we call F and H both depend only on duration and orientation. They are defined as

$$F_{ij} = \kappa_{j1} + \nu_{i1} + \xi_{ij1} \quad \text{(B6)}$$

and

$$H_{ij} = \kappa_{j2} + \nu_{i2} + \xi_{ij2}. \quad \text{(B7)}$$

Thus, F is determined only by terms relevant when faces are shown ($k = 1$), and H is determined only by terms relevant when houses are shown ($k = 2$). Finally, there are separate performance equations for faces and houses:

$$\text{Faces: } p_{ij1} = S_{ij} + F_{ij} \quad \text{(B8)}$$

and

$$\text{Houses: } p_{ij2} = S_{ij} + H_{ij}. \quad \text{(B9)}$$