Psychonomic Bulletin & Review
2009, 16 (5), 778-797
doi:10.3758/PBR.16.5.778

Why does picture-plane inversion sometimes dissociate perception of features and spacing in faces, and sometimes not? Toward a new theory of holistic processing

Elinor McKone
Australian National University, Canberra, ACT, Australia

and

Galit Yovel
Tel Aviv University, Tel Aviv, Israel

Classically, it has been presumed that picture-plane inversion primarily reduces sensitivity to spacing/configural information in faces (distance between location of the major features) and has little effect on sensitivity to local feature information (e.g., eye shape or color). Here, we review 22 published studies relevant to this claim. Data show that the feature inversion effect varied substantially across studies as a function of the following factors: whether the feature change was shape only or included color/brightness, the number of faces in the stimulus set, and whether the feature was in facial context. For shape-only changes in facial context, feature inversion effects were as large as typical spacing inversion effects. Small feature inversion effects occurred only when a task could be efficiently solved by visual-processing areas outside whole-face coding. The results argue that holistic/configural processing for upright faces integrates exact feature shape and spacing between blobs. We describe two plausible approaches to this process.

Face recognition is critical for our daily social interactions. The ease with which we discriminate among the thousands of faces that we encounter throughout our life is remarkable, given the very similar structure and features that all faces share. An important contribution to this ability, which is found for upright but not inverted faces, is presumed to be a style of perceptual processing commonly referred to as *holistic* (Tanaka & Farah, 1993), *configural* (Maurer, Le Grand, & Mondloch, 2002), or *second-order relational* (Diamond & Carey, 1986). This type of processing has classically been thought to include information about the spacing among the major face features but to exclude information about the shape and/or color of these features. The goal of this review is to examine whether published data support this prevalent claim. We first will describe the theoretical question and present the basic terminology used in studies of holistic/configural face processing. We then will critically examine the results of 22 studies that investigated the effect of picture-plane inversion on the processing of spacing and/or features in faces. At the end, we will integrate these findings with the results of studies in which various other methodologies were used (e.g., childhood development, neuropsychology, neuroimaging, individual differences) and will suggest an alternative view about the nature of holistic/configural processing of faces.

The Aspects of Facial Information Included in Holistic Processing for Upright Faces

It is well established that holistic/configural/second-order relational processing operates only for upright faces and not for inverted (upside-down) faces. Classic evidence includes the *composite effect* (Young, Hellawell, & Hay, 1987), in which identifying a top half-face suffers interference from alignment of a competing-identity bottom half, as compared with a misaligned condition, and the *part–whole effect* (Tanaka & Farah, 1993), in which memory for a face part is much better in the context of the studied whole face than when presented alone. Both the composite and part–whole effects are obtained for upright faces but not inverted faces, and these and many other paradigms confirm a qualitative difference between the perceptual processing of faces in the upright and inverted orientations (for recent reviews, see McKone, in press; Rossion, 2008).

An open theoretical question concerns the aspect or aspects of facial information that are included within the mental holistic/configural/second-order relational repre-

E. McKone, elinor.mckone@anu.edu.au

sentation of an upright face. Some early theorists (Tanaka & Farah, 1993) did not exclude any types of information, proposing an undecomposed whole that would necessarily include information about both spacing between major features (e.g., distance between the eyes or between the nose and the mouth) and the exact shape of those features. Others wavered, even within a single article, with Diamond and Carey (1986) sometimes using *second-order relational information* to refer only to spacing between major features (e.g., p. 110, para. 5), and sometimes to implicitly refer to any deviation, in an individual face, from the basic configuration common to all faces (which would include deviation in feature shape; e.g., p. 115, para. 1; p. 108, beginning of para. 3). Currently, however, perhaps the most common view is that *spacing* or *configural* information in faces (e.g., interocular distance, mouth–nose distance) is coded differently from local feature information (e.g., nose shape, nose size, eye color) and that *only the former is part of configural/holistic/second-order relational processing for faces*.

Classically, this view has been supported by findings that although sensitivity to spacing changes is substantially impaired by inversion of the face, inversion has little or no effect on the perception of feature changes. Our aim in the present article is to present new conclusions arising from a full review of the evidence relevant to this classical claim. Our conclusions, in turn, are relevant to the theoretical question of whether feature information (particularly feature shape) should be included in, or excluded from, the notion of configural/holistic/second-order relational processing.

Definitions

We will begin with some important definitions. There is much variability in the use of terminology in this area. As one relevant example, *configural* is sometimes used to refer to a *type of information* on the face image (and a corresponding stimulus manipulation—i.e., a spacing change such as in interocular distance; e.g., Leder & Carbon, 2006) and sometimes to refer to a *style of mental processing* (e.g., Maurer et al., 2002; McKone, 2008). Moreover, where *configural* is used to refer to mental processing, it is sometimes taken as being different from holistic processing (e.g., as an overarching term including holistic processing as a subtype; Maurer et al., 2002) and sometimes used completely interchangeably with holistic processing (e.g., McKone, 2008).

For the present purposes, however, our primary interest is in the definitions researchers have used for two terms: *spacing/configural/metric distance information* on faces and *feature/local/part information*. In general, researchers have defined these terms by example, rather than explicitly. To illustrate the usage typical in the field over the last 10 years, we take the following quotes from Rossion (2008).

> Metric-distance/spacing/configural information: "metric distances between features (e.g., interocular distance, mouth–nose distance . . .)" (p. 276, ellipsis in original)
>
> Feature/local/part information: "a facial feature (e.g., eyes, nose, mouth, . . .)" (p. 275); "local cues (e.g., shape of the mouth, color of the eyes . . .)" (p. 276); "Rhodes et al. (2006) made featural changes by simultaneously altering eye and lip brightness and making the end of the nose more bulbous. This latter modification of feature shape does not change its distance from the eyes and mouth." (p. 279)

There are several important points to note here. First, in the typical usage, the *features* between which metric distances are computed are presumed to be the large nameable facial components: eyes, nose, mouth, eyebrows. Second, the *shape* of these features (e.g., the bulbous nose) is explicitly stated to be *local* information, and *changing this shape is excluded from affecting metric distance information*. Taking these two points together, we can see that the implicit idea is that metric-distance/spacing/configural information is limited to the distances between the major features *treated as blobs* and presumably, therefore, between the locations of the *centers* of these features.

The classic claim, then, is that inversion dissociates perception of local feature information from perception of configural/spacing/metric-distance information *by these definitions* and, specifically, that perception of local feature information shows no or weak inversion effects.

> The "classic" information type × inversion interaction: "numerous behavioral studies have shown that the perception of metric distances between features (e.g., interocular distance, mouth–nose distance . . .) is more affected by inversion than the perception of local cues (e.g., shape of the mouth, color of the eyes . . .)." (Rossion, 2008, p. 276)
>
> "Adults use [spacing of internal facial features] effectively to discriminate upright but not inverted faces, and, with most stimulus sets, the degradation with inversion is far greater for spacing changes than for featural changes." (Maurer et al., 2007, p. 1439)

Aims and Overview

The aim of this article is to evaluate the status of the claim that inversion effects are much smaller for perception of featural information in faces than for perception of distances between the major features. For convenience, we will refer to *spacing/configural/metric-distance* information, as traditionally defined, simply as *spacing*. We will use *spacing inversion effect* as shorthand for the effect of inversion on the perception of changes to spacing information in faces and *feature inversion effect* as shorthand for the effect of inversion on the perception of changes to feature information (e.g., eyes, nose, mouth) in faces.

In reevaluating the status of the classic information type × inversion interaction, it is important to note that we have no disagreement with the standard view that inversion effects for spacing changes in faces are large. The size of spacing inversion effects can vary somewhat—for example, with the particular features between which distances are changed (interocular vs. nose–mouth; e.g., Barton, Deepak, & Malik, 2003) or with another variable variously argued to be direction (vertical vs. horizontal; Goffaux & Rossion, 2007) or range of the changes (short vs. long; Sekunova & Barton, 2008). However, as long as performance in the spacing condition does not approach ceiling or floor, the consistent finding across all the studies we reviewed was of large spacing inversion effects: The average across the 17 independent studies we examined was 26% of the maximum inversion effect possible on the response scale (see the Method section for definition and Table 1 for results). Our focus instead is on the size of the feature inversion effect and, particularly, whether this is always small, both in absolute terms (i.e., relative to the response scale) and relative to the spacing inversion effect.

Our review and discussion in the following sections will lead to several major conclusions, which can be summarized as follows. First, feature inversion effects can be very large, and most such findings cannot be attributed to methodological problems and, so, must be treated as real and in need of explanation. Second, a great deal of the very substantial variability in feature inversion effects across studies can be accounted for by one simple factor: the extent to which the feature change manipulation involves change in the color of the facial feature versus change in shape. Third, small feature inversion effects occur only when the stimulus manipulation or other aspects of the experimental task setup encourage subjects to attend strategically to information derived from levels of the visual-processing stream other than the whole-face representation. This proposal is supported by findings that (1) feature inversion effects are negligible with extreme or large color changes; (2) when the color change included is moderate, stimuli comprising a small set of faces repeated multiple times produce weaker inversion effects than do larger unrepeated sets; (3) large feature inversion effects occur for changes mostly or entirely in shape; and (4) these can be reduced to zero simply by showing the feature in isolation (i.e., removing it from the facial context).

Theoretically, we will argue that the existence of large inversion effects for shape-only feature changes implies that holistic/configural/second-order relational processing for upright faces *must encompass detailed local shape information*. We will thus propose that holistic face processing includes far more information about facial shape than merely distances between the major face blobs. In terms of how holistic processing should instead be conceptualized, we will then raise two plausible speculations: Holistic processing may derive from calculations of distances between multiple very local *landmark points* on faces, or it may derive directly from an image unprocessed for *key points* or *parts* at any level of scale (e.g., pixel-by-pixel intensity coding).

Method

Our inversion review was based on 22 articles in total. All the data came from normal young adults. We were interested only in own-race face perception: All the data reviewed came from Caucasian subjects tested on Caucasian faces. We included all tasks, including bizarreness ratings, distinctiveness ratings, recognition memory, naming a previously learned face, and sequential same–different decision.

We located 17 articles that tested the inversion effects for both local feature changes and spacing changes. In these articles, the basic design of each experiment was as follows. Manipulations were made within subjects. Stimuli comprised original faces, feature-changed faces (e.g., with new eyes/mouth pasted in from a different individual, or with eye/mouth color/brightness edited), and spacing-changed faces (e.g., eyes moved out and mouth down; see examples in Figure 1). Orientation conditions were *upright* and *inverted*; orientation of the learning and probe trials was always the same (e.g., both inverted) if the procedure used separate learning and probe trials (i.e., same–different decision, recognition memory). The dependent measure varied, depending on the task, and included accuracy of detecting or remembering a stimulus change (same–different tasks, delayed matching-to-sample from three alternatives, old–new recognition memory), accuracy of remembering an assigned name for the face, or change in (specifically, increase in) distinctiveness or bizarreness rating for changed faces, as compared with original faces (where the stimulus change had been designed to make the face more distinctive or bizarre than the original; e.g., by blackening the teeth).

From these 17 articles, we employed several criteria in deciding which particular experiments to include in the review and how to present their data, as follows.

1. Where multiple experiments or conditions were available in the same article, we treated these as independent data points if the feature manipulations were fundamentally different in nature across the experiments/conditions (e.g., the feature appeared in isolation in one case and in the context of the face in another). Where the feature manipulations were the same or conceptually similar (e.g., the same type of manipulation was made, merely to a different particular feature or to a different level of strength), we averaged across the experiments/conditions, subject to the following caveat.

2. If an experiment produced very poor matching of upright performance levels between the feature condition and the spacing condition and there was another experiment in the same article with better matching, we used the results from the experiment with the better matching; this was done on the grounds that it is most reasonable to compare the size of the inversion effect across conditions if some *baseline* performance level (e.g., upright) is matched.

3. In some studies, an additional variable was manipulated simultaneously with the feature versus spacing and inversion manipulations (e.g., presentation duration, degree of physical change); in these cases, we took the data from the condition(s) that produced closest matching of upright performance across feature and spacing, subject to the following caveat.

4. We excluded any conditions that produced ceiling or floor effects. In binary decision tasks (e.g., same–different), for example, we defined this as average performance across upright and inverted either $>90\%$ or $<60\%$.

Table 1
Results of All 22 Studies Included in the Review, Showing Scores for Feature Change Condition Upright (F Up), Feature Change Condition Inverted (F Inv), Feature Inversion Effect As a Percentage of Maximum Possible on the Scale (F Inv Eff As % Scale Range), the Same Information for Spacing-Change Condition, and the Feature Inversion Effect As a Percentage of the Spacing Inversion Effect (F Inv Eff As % S Inv Eff)

Study	Task	Dependent Variable and [Scale Range]	Type of Feature Change	Set Size	F Up	F Inv	F Inv Eff As % Scale Range	S Up	S Inv	S Inv Eff As % Scale Range	F Inv Eff As % S Inv Eff
Rhodes, Brake, & Atkinson (1993)	Recognition memory	%correct [50–100]	Paraphernalia	L	84.7	85.3	−1.2	82.8	76.1	13.4	−9.0
			Isolated feature	L	71.5	77.0	−10.9	74.2	58.9	30.6	−35.6
Leder & Bruce (2000)	Recognition memory	%correct [16.7–100]	Hue/bright (extreme)	S	68.2	65.6	3.1	70.3	47.4	27.5	11.4
Leder & Carbon (2006)	Recognition memory	%correct [16.7–100]	Hue/bright (extreme)	S	92.6	90.7	2.3	83.3	53.7	35.5	6.4
Searcy & Bartlett (1996)	Rate bizarreness	rating altered [1–7] − rating original [1–7]	Bright (large)	L	3.9	4.0	−0.7	4.5	2.4	34.4	−1.9
Murray, Yong, & Rhodes (2000)	Rate bizarreness	rating altered [1–7] − rating original [1–7]	Bright (large)	S	3.4	3.1	5.2	3.8	2.9	15.0	34.4
Barton, Keenan, & Bass (2001)	Odd one out	%correct [33.3–100]	Bright (mod)	S	75.2	64.9	15.4	79.8	50.5	43.7	35.3
Freire, Lee, & Symons (2000)	Sequential match	%correct [50–100]	Bright (mod) + shape	S	86.0	86.0	0.0	74.0	57.0	34.0	0.0
Le Grand, Mondloch, Maurer, & Brent (2001)	Sequential same–different	%correct [50–100]	Bright (mod) + shape	S	80.0	81.0	−2.0	80.0	63.0	34.0	−5.9
Yovel & Duchaine (2006)	Sequential same–different	%correct [50–100]	Bright (mod) + shape	S	84.2	84.0	0.4	77.8	62.9	29.8	1.3
Leder & Bruce (1998)	Rate distinctiveness	rating altered [1–9] − rating original [1–9]	Bright (mod) + shape	L	1.2	0.4	9.1	1.7	0.2	18.3	49.6
Gilchrist & McKone (2003)	Rate distinctiveness	rating altered [1–9] − rating original [1–9]	Bright (mod) + shape	L	0.9	1.0	−1.3	1.3	0.5	9.8	−12.8
Rhodes, Hayward, & Winkler (2006)	Sequential match	%correct [50–100]	Bright (mod) + shape	L	72.9	65.4	15.0	75.1	61.9	26.4	56.8
McKone & Boyer (2006)	Rate distinctiveness	rating altered [1–9] − rating original [1–9]	Bright (mod) + shape	L	1.2	0.8	4.8	1.2	0.6	7.6	62.3
Malcolm, Leung, & Barton (2004)	Odd one out	%correct [33.3–100]	Bright (small) + shape	L	83.9	57.3	39.7	83.8	45.1	57.8	68.7
Goffaux & Rossion (2007)	Sequential same–different	*d′* [0–approx. 4]	Bright (small) + shape	L	3.2	2.5	18.8	3.1	2.0	29.0	64.7
Leder & Carbon (2006)	Recognition memory	%correct [16.7–100]	Bright (small) + shape	S	88.9	71.3	21.1	83.3	53.7	35.5	59.5
Rhodes, Brake, & Atkinson (1993)	Recognition memory	%correct [50–100]	Shape	L	72.7	54.3	36.7	78.5	67.5	22.0	229.8
Riesenhuber, Jarudi, Gilad, & Sinha (2004)	Sequential same–different	%correct [50–100]	Shape	S	83.0	70.0	26.0	73.5	63.0	21.0	122.5
Yovel & Kanwisher (2004)	Sequential same–different	%correct [50–100]	Shape	S	77.0	61.0	32.0	79.0	71.0	16.0	200.0
Yovel & Duchaine (2006)	Sequential same–different	%correct [50–100]	Shape	S	79.1	62.0	34.2	77.8	62.9	29.8	114.8
Pellicano & Rhodes (2003)	Sequential matching	%correct [50–100]	Bright (small) + shape	S	88.5	65.0	47.0	–	–	–	–
Pellicano, Rhodes, & Peters (2006)	Sequential matching	%correct [50–100]	Bright (small) + shape	S	95.0	77.0	36.0	–	–	–	–
Boutet & Faubert (2006)	Sequential matching	%correct [50–100]	Bright (small) + shape	M	91.0	74.0	34.0	–	–	–	–
Tanaka & Farah (1993)	Recognition memory	%correct [50–100]	Shape	S	74.5	65.0	19.0	–	–	–	–
Tanaka & Sengco (1997)	Recognition memory	%correct [50–100]	Shape	S	87.0	69.0	36.0	–	–	–	–

Note—L = large set—that is, faces of at least 20 different individuals, each original edited to have a feature-changed version (and a spacing-changed version); S = small set—that is, either one original face edited to have up to five feature-changed versions (and five spacing-changed versions), or up to four originals each with one feature-changed version; M = medium. Odd-one-out task presented three faces simultaneously, two identical and one different (e.g., by a feature change) and required subjects to choose the different one. Sequential matching tasks presented a target face, followed by two choice faces. "Recognition memory" includes both old–new recognition and naming a previously learned face. Bright = brightness; mod = moderate.

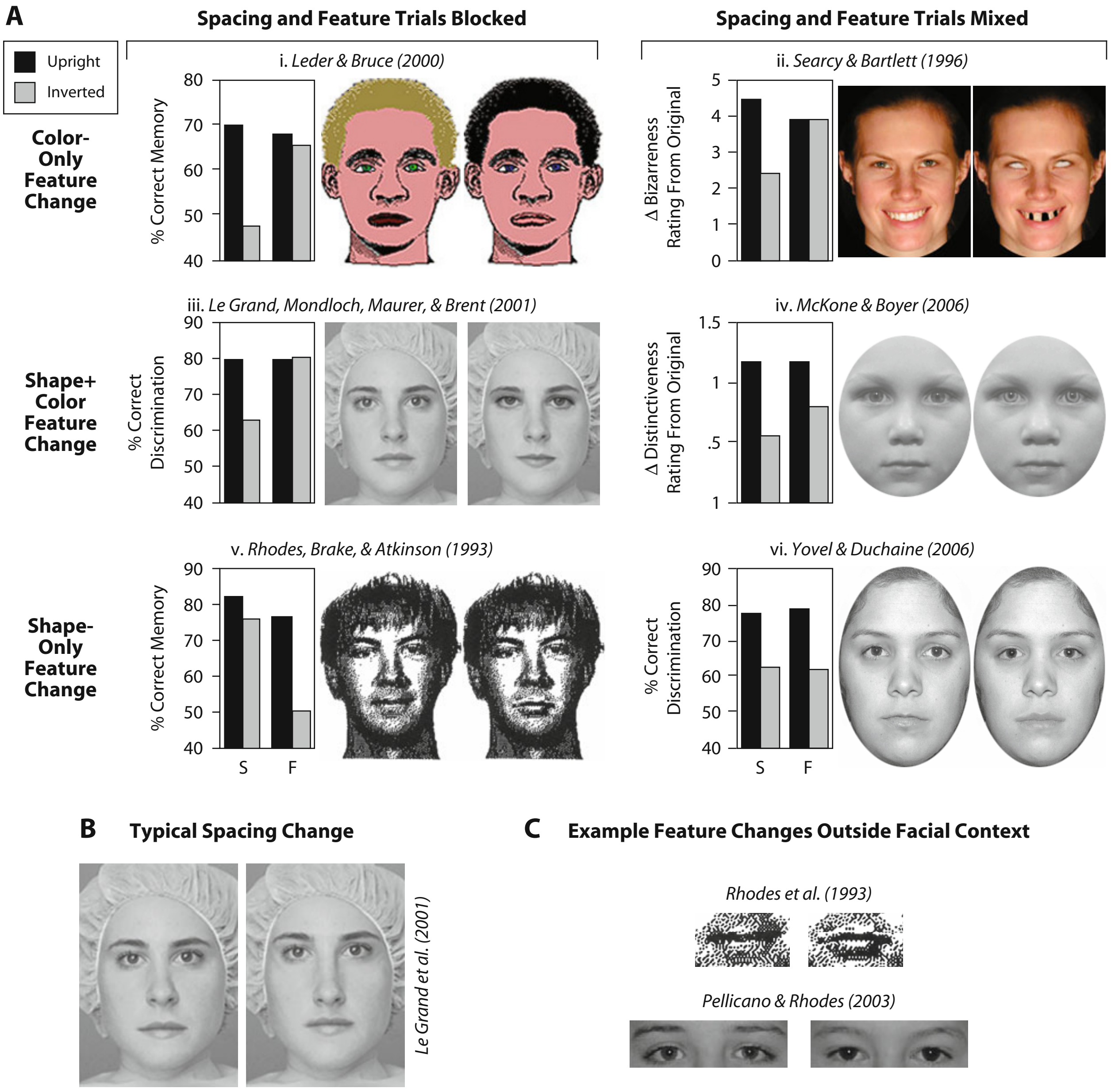

Figure 1. (A) Illustrative examples of stimuli and results (from single experiments) for studies in which the feature changes were in color only (examples are of extreme and large color changes), shape+color (examples are of moderate color changes), or shape only; note that "color" includes brightness in grayscale stimuli. Results show that the inversion effect for feature changes is smallest in color-only cases and largest in shape-only. The examples chosen also illustrate that patterns cannot be explained by ideas that spacing–feature dissociations are an artifact of failure to match performance for the two change types in the upright condition (Yovel & Kanwisher, 2004) or of presentation in separate blocks (Riesenhuber, Jarudi, Gilad, & Sinha, 2004). Patterns are also independent of task. S, spacing change condition; F, feature change condition. Stimuli for Searcy and Bartlett (1996) are re-creations: The stimuli were in color but published in black and white, and copies of the originals were not available. (B) Illustration of the typical type of spacing change: eyes up/down in/out and mouth up/down. (C) Example nonface feature change stimuli. We recommend viewing the full color version of this figure available in the online version of the article.

It was never necessary to exclude an entire experiment because of ceiling or floor effects: These arose mostly where a factor such as degree of physical change was varied across several levels. For example, Barton, Keenan, and Bass (2001) varied eye luminance change in five steps from 3% (giving performance at chance both upright and inverted) to 15% (giving performance close to perfect both upright and inverted), and so we considered data only from the three intermediate luminance change levels.

5. For the spacing conditions, manipulations in most studies involved both vertical and horizontal displacements and affected both eyes and mouth. Where these

factors were varied separately, eyes/mouth/vertical/horizontal were averaged to allow fair comparison across the multiple studies.

We also reviewed results from a paradigm that is not usually discussed in feature-versus-spacing articles but is still informative with respect to the effect of inversion on the processing of *features* in faces. This is the classic part–whole paradigm (Tanaka & Farah, 1993). In the *whole* condition of that paradigm, the method is logically identical to the feature change conditions in the feature-versus-spacing studies: In particular, subjects learn a face and then later discriminate between that face and a face with an alteration to a single major facial feature (e.g., a different nose). Thus, inversion effects for the *whole* condition of the part–whole task are directly relevant to evaluating the size of feature inversion effects. We located 5 relevant studies, which had reported *whole*-condition data for upright and inverted faces. Although no equivalent spacing condition was tested in these articles,[1] we were able to compare the size of the feature inversion effect with that of the typical spacing inversion effect from the 17 feature-versus-spacing studies.

Regarding data analysis, we calculated the size of the inversion effect for features in two ways, both of which were designed to allow comparison across very different tasks and response scales. The first was the feature inversion effect relative to the maximum inversion effect possible on the scale. This was defined as (upright %correct − inverted %correct)/(ceiling − chance) on accuracy measures or (upright rating − inverted rating)/(maximum possible rating − minimum possible rating) on rating measures. For example, a same–different task has chance = 50, ceiling = 100, so if upright = 80 and inverted = 70, the score is (80 − 70)/(100 − 50); similarly, for a distinctiveness rating scale of 1 to 9, the maximum inversion effect for the divisor is 8. This first measure has the advantage of being independent of the size of the spacing inversion effect, which in itself is not a constant (Goffaux & Rossion, 2007; Sekunova & Barton, 2008), thus giving the purest measure of the size of the *feature inversion effect in its own right*.

Where possible, we also computed the feature inversion effect relative to the spacing inversion effect in the same study—for example, as (feature upright %correct − feature inverted %correct)/(spacing upright %correct − spacing inverted %correct). This method has the advantage of telling us to what extent a study produced the classic pattern of inversion × information type interaction: A score of 0 indicates no feature inversion effect and, so, a maximally large interaction (given that spacing change inversion effects were present in all cases), whereas a score of 100% indicates a feature inversion effect as large as the spacing inversion effect in the particular task and, thus, no interaction.

Some readers might question why we did not use meta-analytic techniques based on percentage-of-variance-explained measures (statistical *effect size*). This was primarily because it was not possible to calculate the relevant input data in many studies. To calculate percentage-of-variance-explained for the feature inversion effect requires the original article to have reported either the *SEM* of the upright–inverted difference scores or a direct *t* test on upright versus inverted. This information was not always available, even within just a single experiment, and could never be calculated in cases in which we wished to combine results from across two experiments in the same article (which was important to ensure that the particular stimulus manipulation in that article was not given undue weighting in the review results).[2]

Results

The results of our review are given in Table 1, and key findings are illustrated in Figures 1–3. Figure 1 provides upright and inverted scores separately and sample stimuli, for 6 illustrative studies that included both feature and spacing changes. Figure 2 provides the complete review of the 17 feature-versus-spacing studies, presenting our two summary measures of *feature inversion effect relative to scale* (Figure 2A) and *feature inversion effect relative to spacing inversion effect* (Figure 2B). Figure 3 presents the 5 studies that included only a feature manipulation, plotting feature inversion effect relative to scale. Considering all figures together, five key results are apparent.

First, feature changes often produce very large inversion effects. For comparison, recall that the average spacing inversion effect across 17 studies is 26%, relative to scale range. In 5 of the 17 feature-versus-spacing studies (Figure 2), the feature inversion effect exceeds 30% of the scale and/or 100% of the spacing inversion effect in the same study (Malcolm, Leung, & Barton, 2004; Rhodes, Brake, & Atkinson, 1993; Riesenhuber, Jarudi, Gilad, & Sinha, 2004; Yovel & Duchaine, 2006; Yovel & Kanwisher, 2004). Furthermore, in 4 of the 5 studies testing only the feature manipulation (Figure 3), the feature inversion effect again exceeds 30% of the scale (Boutet & Faubert, 2006; Pellicano & Rhodes, 2003; Pellicano, Rhodes, & Peters, 2006; Tanaka & Sengco, 1997). We also note that, across Figures 2 and 3, 6 additional studies show a moderately large feature inversion effect of at least 20% of the scale and/or 50% of the spacing inversion effect in the study (Goffaux & Rossion, 2007; Leder & Bruce, 1998; Leder & Carbon, 2006; McKone & Boyer, 2006; Rhodes, Hayward, & Winkler, 2006; Tanaka & Farah, 1993). Overall, the results show that findings of large inversion effects for feature changes in faces cannot simply be considered outliers or statistical anomalies.

Second, the classic pattern of essentially no inversion effect for feature changes occurs in only approximately 35% of the reported tests (8/22). Taking this result together with the striking variability across the other studies, it is clear that the classic claim—that inversion affects perception of featural information far less than it affects perception of spacing information—is not an accurate summary of the evidence.

Third, it is possible to order the studies (e.g., from left to right in Figure 2) on a principled basis in a manner that accounts rather well for the continuous variation in the size of the feature inversion effect—which ranges from nothing to as large as or larger than typical spacing inversion effects. Specifically, feature inversion effects are largest

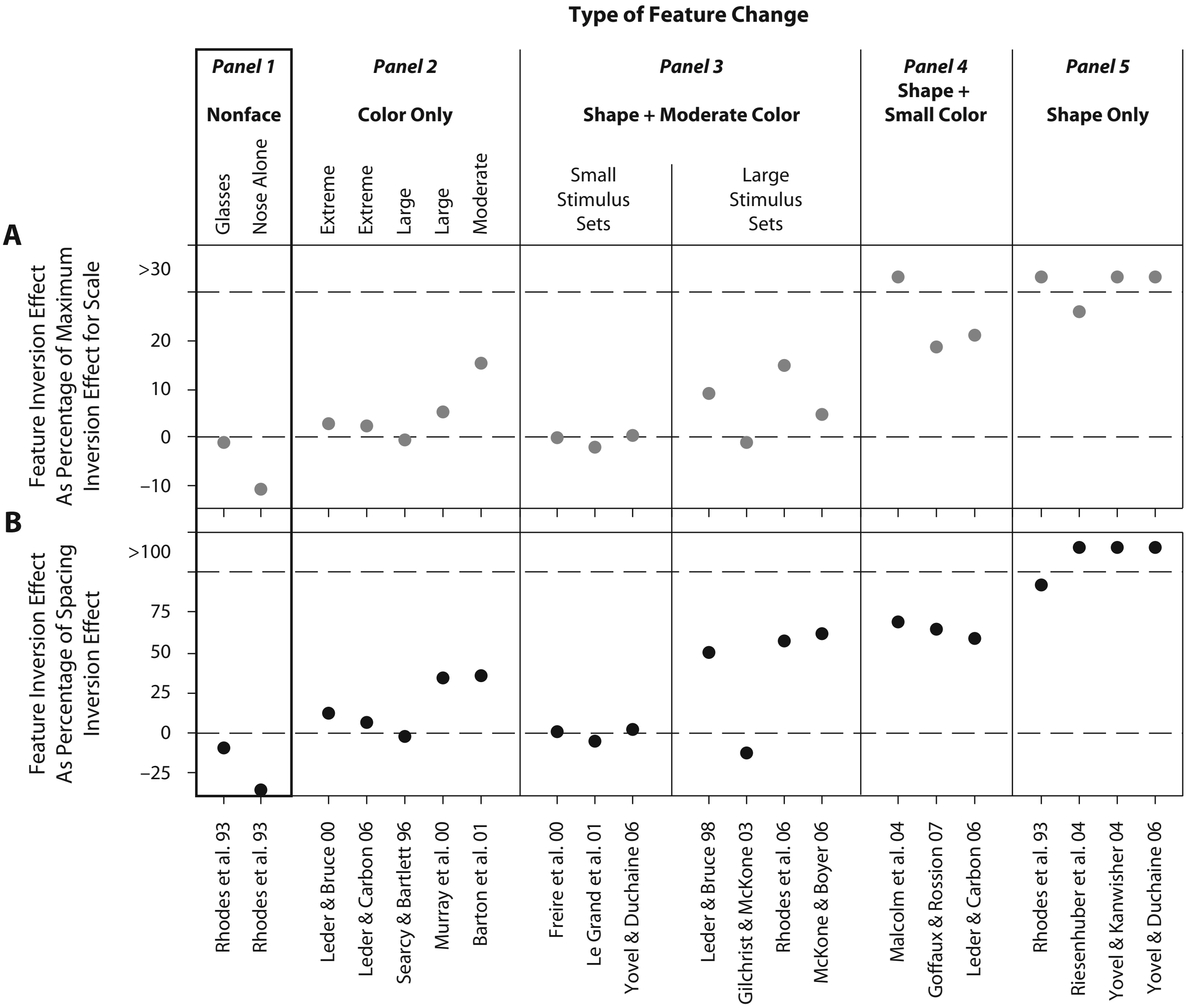

Figure 2. Variability in inversion effects for different types of feature changes in faces: Data from a full review of 17 studies that included both feature manipulation and spacing manipulation conditions. Two measures were used to allow comparison across different tasks. (A) The size of the feature inversion effect as a percentage of the maximum possible inversion effect on the scale—for example, (*upright %correct − inverted %correct*)/(*ceiling − chance*) or (*upright rating − inverted rating*)/(*max possible rating − min possible rating*). This measure is independent of spacing changes. (B) The size of the feature inversion effect as a percentage of the spacing inversion effect. Results of both measures confirm a strong relationship between the extent to which feature changes are in color/brightness versus shape only and the size of the inversion effect (panels 2 to 5; note that within each subpanel of panel 3 and within panels 4 and 5, the studies are ordered by date of publication, rather than on any theoretical basis). The results also show that when the "features" are not *face* features in a *face context* (panel 1), there is no inversion effect for changes in paraphernalia on the face (e.g., glasses) or for isolated face features. Exact values for scores plotted as >30 or >100 can be found in Table 1.

for changes mostly or entirely in shape (panels 4 and 5 of Figure 2, plus Figure 3), are smaller if the manipulation additionally includes a moderate color change (panel 3 of Figure 2), and disappear altogether for dramatic color/brightness changes outside the normal range (panel 2 of Figure 2); in addition, within the band of *shape + moderate color* changes, inversion effects tend to be smaller where the stimulus set includes a small number of items repeated multiple times than where it includes a large number of items without repetition; finally, inversion effects disappear where the feature is presented in isolation (e.g., nose alone) or is not a face feature (e.g., spectacles; panel 1 of Figure 2). We will discuss these findings in more detail in a subsequent section.

Fourth, the results of individual studies, and particularly the ordering of feature inversion effects in terms of size across studies, remain very similar regardless of whether the feature inversion effect is measured relative to scale (Figure 2A) or relative to the spacing inversion effect (Figure 2B). Thus, type of *spacing* change included in a particular study is not relevant to explaining variation in the size of the *feature* inversion effect (contrary to the suggestions of Maurer et al., 2007, and Rossion, 2008). This observation was also demonstrated by a lack of cor-

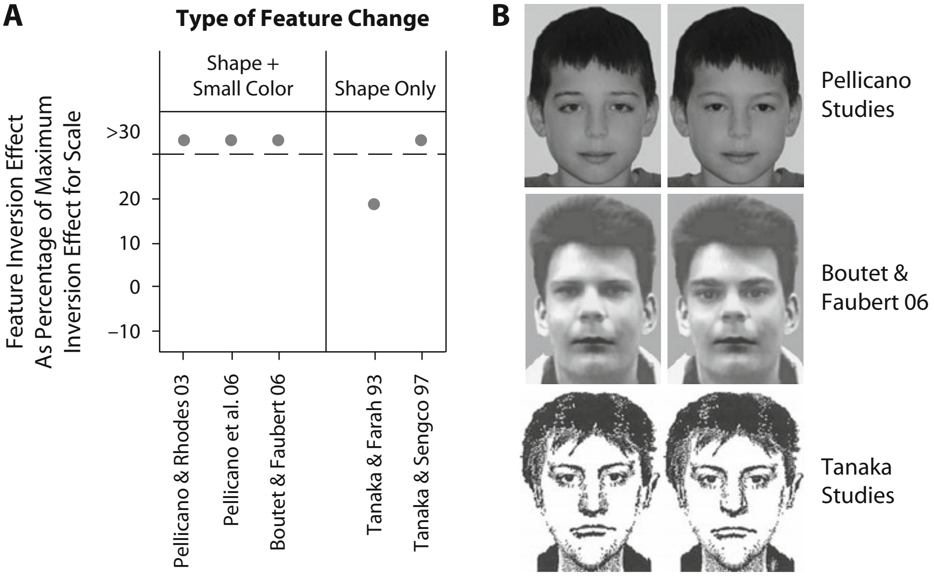

Figure 3. (A) Inversion effects for feature changes in studies that did not also test a spacing condition (specifically, the *whole* condition from part–whole studies). (B) Sample stimuli from these studies. As in Figure 2, the conclusion is that feature changes entirely or mostly in *shape* produce large inversion effects.

relation across studies between the feature inversion effect (“F Inv Eff as % Scale” column in Table 1) and the corresponding spacing inversion effect from the same experiment (“S Inv Eff As % Scale” column in Table 1) [$r(20) = .23$, $p > .3$]. In particular, Table 1 confirms that it is not the case that conditions with very large feature inversion effects in Figure 2B obtained this result merely because the study in question happened to produce an atypically small inversion effect for spacing.

Fifth, the size of the feature inversion effect was not obviously influenced by task (see Table 1). Long-term memory tasks (e.g., old–new recognition, naming a previously learned face) sometimes produced no feature inversion effect (e.g., Leder & Bruce, 2000) and sometimes a very large effect (Rhodes et al., 1993; Tanaka & Sengco, 1997). Sequential same–different tasks produced feature inversion effects ranging in size anywhere from nothing (e.g., Le Grand, Mondloch, Maurer, & Brent, 2001) to very large (e.g., Yovel & Duchaine, 2006). Tasks rating bizarreness or distinctiveness produced feature inversion effects ranging in size from nothing (e.g., Searcy & Bartlett, 1996) to moderate (e.g., Leder & Bruce, 1998).

Can Large Feature Inversion Effects Be Explained Away?

Turning to interpretation of the results, one approach in the previous literature (e.g., Maurer et al., 2007; Rossion, 2008) has been to suggest that large feature inversion effects are invalid and reflect something wrong with the study in question. Clearly, if this were justified, we could place no theoretical weight on the findings of large feature inversion effects. We thus begin our interpretation by addressing—and refuting—these previous criticisms.

With respect to Riesenhuber et al. (2004), Rossion (2008) pointed out that the feature change in that study involved replacing the entire eye–eyebrow region of one individual with that of another and that this noticeably altered the spacing between the centers of two major facial features—namely, the eyes and the eyebrows. Rossion thus argued that Riesenhuber et al.’s “feature” change condition should in fact be described as a feature + spacing change condition and that the inclusion of the spacing change was responsible for the large inversion effect. We agree that Riesenhuber et al.’s manipulation could have produced a large “feature” inversion effect for this reason, although only for the one of two stimulus pairs to which this problem applied. Similarly, we note that the same criticism could be applied to up to one third of the trials in Boutet and Faubert (2006; the eye trials; see Figure 3). Importantly, however, this criticism cannot be applied to other studies that have produced very large or moderate feature inversion effects. Yovel and Duchaine (2006), Yovel and Kanwisher (2004), Malcolm et al. (2004), and Rhodes et al. (2006) all replaced or changed each *single* feature (e.g., just the mouth or just the eyes without the eyebrows) with a new feature *centered in the same position*; thus, there was no change in metric distances between the locations of the major facial features (nor does there appear to be in Pellicano & Rhodes, 2003, or Pellicano et al., 2006; see Figure 3). So, according to the standard definitions, these studies manipulated featural information independently of spacing between features.

Yovel and Kanwisher’s (2004) study has been criticized on two grounds. Maurer et al. (2007, p. 1440) noted that although the feature-changed faces in the Yovel and Kanwisher (2004) study were natural in appearance, the spacing-changed faces appeared somewhat unnatural due to spacing changes outside the normal physical range. Given that inversion effects can vary with facial distinctiveness (Valentine, 1991), we agree that it is possible that the equal-sized inversion effect for feature and spacing changes in the Yovel and Kanwisher (2004) study could have been an artifact of distinctiveness differences.[3] Crucially, however, Maurer et al.'s (2007) criticism is relevant *only to the comparison of feature inversion effects with spacing inversion effects*; it cannot explain why Yovel and Kanwisher's (2004) inversion effect for features *itself*, considered completely independently of the fact that a spacing condition was also tested, was so large (see Figure 2A). In a second criticism, Rossion (2008) suggested that Yovel and Kanwisher's (2004) result might have been due to speed–accuracy trade-offs, given that (as in most studies) only accuracy was reported. However, this was not the case: Although reaction times were shorter for inverted than for upright faces, the inversion effect in that data set was not significantly different for feature changes (upright − inverted = −38 msec) and spacing changes (−24 msec); also, in a second study in which the upright and inverted faces were presented interleaved (Yovel & Kanwisher, 2004, supplementary), reaction times were longer for inverted than for upright faces, and the inversion effect was again of similar magnitude for spacing (31 msec) and features (44 msec). Thus, we see no reason to discount the large feature inversion effect in Yovel and Kanwisher (2004).

Overall, a valid argument has been made that might partially explain away the large feature inversion effects of two studies (Boutet & Faubert, 2006; Riesenhuber et al., 2004). However, there is no reason to discount the large feature inversion effects of seven other articles (including Yovel & Kanwisher, 2004). We are also unaware of any reason to discount the several studies producing medium-sized feature inversion effects (see Figures 2 and 3). We conclude, therefore, that rather than attempting to explain away findings of large feature inversion effects as wrong or atypical, we should treat the existence of large feature inversion effects as genuine and turn our attention to understanding the factors that contribute to the enormous variability in feature inversion effect across studies.

Two Explanations of the Variability That Do Not Work

Prior to 2006, two explanations of variability across studies had been proposed. Importantly, both focused on the variability in the *information type × inversion interaction*, rather than on the more relevant feature inversion effect per se. We will deal with these proposals only briefly, because both are easily refuted by our review data.

First, Riesenhuber et al. (2004) proposed that feature inversion effects were large, relative to spacing effects, when the two change types were intermixed, but not when they were blocked. Figure 1 illustrates clear counterexamples to this proposal: There exist both (1) studies with intermixed stimuli that produced small feature inversion effects and (2) studies with blocked stimuli that produced large feature inversion effects. Rossion (2008) also described in detail how even Riesenhuber et al.'s own data did not, in fact, support their hypothesis.

Second, Yovel and Kanwisher (2004) proposed that feature inversion effects were large, relative to spacing effects, when performance for the two change types was matched upright, but not when the spacing condition was harder than the feature condition. Figure 1 again illustrates clear counterexamples to this proposal: All the studies have well-matched performance for feature and spacing changes in the upright orientation, yet dramatic variation in the relative size of the feature inversion effect is apparent.[4]

Our Proposal: Three Contributing Factors

Given the failure of earlier attempts to explain—or explain away—the variability in feature inversion effects across studies, it is clear that a different approach is needed. On the basis of the results in Figures 2 and 3, we will now argue that three factors affect the size of the feature inversion effect. Note that one of these factors—the extent to which the feature change is shape only or also includes brightness or color changes—has previously been proposed to be important by Leder and Carbon (2006) and Yovel and Duchaine (2006), in two studies in which this variable was experimentally manipulated. However, although both of those studies individually produced data supporting a role for shape versus color, neither article attempted to provide a complete review of the extent to which it could account for the range of findings across previous studies, as we do here.

Factor 1: The extent to which the feature change is shape only or also includes color and/or brightness. We use *color* as shorthand for changes in both brightness in grayscale stimuli and color in color stimuli. To examine the effects of this variable, we grouped the strength of color changes into several categories by examining the original stimuli. Examples of *extreme color changes* are given in Figure 1A(i). Leder and Bruce (2000; see also Leder & Carbon, 2006) used line drawings with unnatural coloring of skin, eyes, lips, and hair, and feature changes involved alteration from one set of striking colors to another. Figure 1A(ii) shows an example of a *large color change*: Searcy and Bartlett (1996; see also Murray, Yong, & Rhodes, 2000) started from natural face images and made these bizarre by whitening the eye region and blackening some of the teeth. The second row of Figure 1 shows examples of *moderate color changes*: Here, all colors were within or close-to-within the normal range, and changes involved, for example, swapping the light eyes of one individual with the dark eyes of another [Figure 1A(iii), the "Jane" faces; Le Grand et al., 2001], swapping the bare lips of one woman with those of another wearing lipstick [Figure 1A(iii)], or lightening the eye color quite substantially by editing the face photograph [Figure 1A(iv); McKone & Boyer, 2006]. *Small color changes* were, then, those where feature replacements were made and less noticeable color changes were introduced and, in some cases, involved an attempt to equate brightness that was reasonably, but not completely, successful. For example, Goffaux and Rossion (2007) pasted in different eyes, and the example they gave shows some brightness/contrast difference in irises/sclera; and Leder and Carbon had small brightness changes arising from differences in eyebrow thickness (see their Figure 1, p. 21). Finally, stimuli with *"no" color change* had variations in local color that were barely detectable and were the closest to genuinely no color change within the studies we reviewed: Figure 1A(v) shows stimuli from Rhodes et al. (1993), Figure 1(vi) from Yovel and Duchaine (2006).

At the same time, we were also interested in whether the feature changes had varied shape. Shape changes were always to major internal features only (eyes, nose, mouth, eyebrows). Example manipulations were replacing one woman's straighter lips with the curvier lips of another or editing to make the nose broader. Shape changes were kept within the approximately normal range and, from the example stimuli provided in the articles, appeared similar in strength across studies. Thus, we defined only two levels of shape change: *no shape change* [e.g., Figures 1A(i) and 1A(ii)] and *shape change* [e.g., Figures 1A(iii)–1A(vi)].

Turning to the review results, panel 2 of Figure 2 shows findings for color-only feature changes. As can be seen, the four studies with dramatic color changes (extreme or large) produced zero or very small inversion effects, whereas the one study with more moderate color changes showed some suggestion of a larger inversion effect. Panel 3 shows the results for manipulations of shape made simultaneously with moderate manipulations of color (shape + moderate color). Here, the results are quite variable, with feature inversion effects ranging from nothing to 50% of the spacing inversion effect. Panel 4 shows tests that used changes of shape simultaneously with small color changes (shape + small color). Here, feature inversion effects become quite large. Finally, panel 5 shows tests using shape-only changes (i.e., shape + "no" color). Here, feature inversion effects reach their maximum, being at least as strong as spacing inversion effects. Turning to Figure 3, studies testing only the feature change condition confirm that feature inversion effects are large or very large when the feature manipulation is shape only or shape + small color.

We thus conclude that perception of feature changes in faces is sensitive to inversion in proportion to the extent that the feature change is in shape only, rather than in color. In the extreme cases, dramatic color-only changes produce no feature inversion effect, and changes entirely or mostly in shape produce very large feature inversion effects. Combinations of manipulations between these two extremes produce intermediate feature inversion effects.

Factor 2: Stimulus set size and item repetition. Within the category of *shape + moderate color* changes, panel 3 of Figure 2 shows that the feature inversion effect still varies noticeably across studies. What explains this variability? A relevant factor appears to be the size of the stimulus set (and/or the confounded variable of stimulus repetition). Feature inversion effects are moderate in size where many different faces are used and not repeated across trials within the experiment. Feature inversion effects disappear, however, when the stimulus set comprises only a few items that are repeated multiple times.

This proposal can account for results of six of the seven studies in panel 3. The three studies in the "small stimulus sets" subpanel (Freire, Lee, & Symons, 2000; Le Grand et al., 2001; Yovel & Duchaine, 2006) all used a feature stimulus set comprising only five to eight images (an original face plus four to seven feature-changed versions), and the task on each trial was to make same–different discriminations among pairs of these images. Across trials, each image was repeated 28 times (Yovel & Duchaine, 2006), 28 times (Freire et al., 2000), or approximately 24 times (Le Grand et al., 2001). Under these circumstances of small sets of items repeated multiple times, shape + moderate color feature changes produced no inversion effect at all for all three studies.

In contrast, the studies in the "large stimulus sets" subpanel used much larger stimulus sets. In Rhodes et al. (2006), the task on each trial was similar to the procedure of the "small sets" studies (i.e., a version of sequential discrimination). However, the experiment used 24 different target faces and 24 corresponding feature-altered distractors, instead of five to eight total images. The other three studies in this subpanel all used a distinctiveness rating task. Leder and Bruce (1998) used 30 original faces, each with a feature-changed version; McKone and Boyer (2006) used 60 original faces, each with a feature-changed version; and Gilchrist and McKone (2003) used 20 faces comprising 10 individuals in the original version and a different 10 individuals shown in a feature-changed distinctiveness-enhanced version of their originals. In all four studies, the stimuli were not repeated across trials. Under these circumstances of large stimulus set sizes and no item repetition, three of the four studies (the outlier being Gilchrist & McKone, 2003) produced noticeable inversion effects for feature changes, despite the inclusion of moderate color change as part of the feature manipulations. These feature inversion effects were 50%–62% of the spacing inversion effect (Figure 2B).[5]

A final point is that the effect of set size does not outweigh the effect of type of manipulation change. For shape-only changes or shape + small color changes (panels 4 and 5 of Figure 2, plus Figure 3), large feature inversion effects were obtained despite many of the studies using small set sizes (Table 1).

Factor 3: The role of facial context. Our final important factor is whether the face feature is shown in isolation or in the context of the face. This is illustrated most dramatically by the results from Rhodes et al. (1993). This study showed that when a feature change was made in a *face* feature on a stimulus showing the *whole face* (i.e., the standard procedure), a very large inversion effect for the shape-only feature changes emerged (panel 5 of Figure 2; see also Figure 1). In contrast, there was a complete lack of inversion effect for exactly the same feature changes when the face feature was presented out of the context of a face (e.g., the mouth in isolation) or when the "feature" was nonface information (e.g., change in spectacles on the face; see panel 1 of Figure 2).[6]

Results from other paradigms are consistent with the view that removal of features from facial context destroys large feature inversion effects. The classic part–whole procedure produces substantial inversion effects for discriminating whole faces differing only in a single feature (Figure 3). However, there is no inversion effect when exactly the same feature change is made but the feature is presented alone: 1% of the maximum inversion effect possible on the scale in Tanaka and Farah (1993), 6% in Tanaka and Sengco (1997), −5% in Pellicano and Rhodes (2003), and 6% in Pellicano et al. (2006), as compared with an average of 34.5% across these same studies for features in a facial context. Rakover and Teucher (1997) also found only small inversion effects for features presented in isolation, although with some suggestion that the effect was larger for features showing larger proportions of the face (eyes + eyebrows = 14% of maximum inversion effect possible on scale) than for smaller features (mouth = 4%). Finally, in categorical perception in noise and peripheral inversion tasks, whole faces produce large inversion effects, but there is no inversion effect at all when the most discriminating feature of the same faces is shown in isolation (McKone, 2004; McKone, Martini, & Nakayama, 2001). Inversion effects are also negligible for scrambled faces (e.g., Martini, McKone, & Nakayama, 2006).

Taken together, these results make a strong case that when large inversion effects on feature perception occur, they occur only if the feature is presented in the context of proper facial structure.

Summary. Integrating the results for all three factors, our review leads to the following conclusions. Feature inversion effects are absent or small—that is, the classic finding is obtained—only if (1) the feature is presented out of context, meaning that the stimulus is not a face; (2) the feature manipulation includes dramatic color change, far outside the normal range; or (3) the feature manipulation includes moderate (normal range or close-to-normal range) color changes in conjunction with a small stimulus set in which items are repeated multiple times and subjects make a simple same–different decision to sequential pairs of faces. Note that in all these cases, the stimuli and/or task requirements differ quite noticeably from those that occur in real-world face recognition settings. In contrast, (4) for shape + moderate color manipulations, feature inversion effects become larger under conditions that more closely mimic naturalistic face perception requirements (i.e., the experiment includes more face stimuli without repetition), and (5) regardless of repetition, feature inversion effects are large where the feature change is mostly or entirely in *shape*.

Variability in Feature Inversion Effects Reflects the Extent to Which the Experiment Encourages Reference to Information Outside a Holistic Face Representation

How should these results be best accounted for theoretically? We will propose two key ideas.

Holistic processing includes feature shape as much as spacing. First, we argue that holistic/configural/second-order relational processing of upright faces encompasses not only spacing between the major features, but the detailed shape of those features as well. That is, holistic processing codes *all shape-related information* in the face.

This is supported by the finding that, as long as the whole face is presented, feature changes that are mostly or entirely in shape produce large feature inversion effects (panels 4 and 5 of Figure 2; Figure 3). Moreover, our review shows that feature shape does not play merely a minor or secondary role. Nine independent studies have reported feature inversion effects for shape changes that *are larger than the typical inversion effect for spacing changes*. This indicates a primary role for local feature shape information in holistic/configural/second-order relational processing.

Attention to holistic processing versus other levels of representation. Second, we propose that large inversion effects for feature changes occur to the extent that the stimuli and procedure of a particular experimental task encourage observers to base their responses on the holistic face representation, as opposed to strategically using information available from other stages of the visual-processing stream (e.g., low-level color processing). That is, large feature inversion effects will be obtained whenever the subject's task performance is driven by reliance on holistic processing. Small feature inversion effects, in contrast, will arise when other cues become easily usable for performing the task and subjects find it more efficient to ignore the holistic processing stage and respond on the basis of information obtained from other stages of the visual-processing stream.

These ideas explain the variation in feature inversion effect across studies as follows. (We assume for the moment that the feature is always in a facial context.) For extreme color or large color changes (i.e., most of the studies in panel 2 of Figure 2), there is no need for the subject to wait until full face information has become available: Color information available earlier in the visual-processing stream will provide simple information that the subject can use to drive a rapid and accurate response. Because color processing in early vision does not depend on the orientation of the object structure on which the color is superimposed, these responses will be insensitive to face inversion. At the other extreme, for shape-only changes (in a facial context; panel 5 of Figure 2; Figure 3), there are no speed or accuracy advantages to the subject to be gained by choosing to base responses on information outside the holistic face representation. Technically, this could probably be done (e.g., subjects could attempt to respond on the basis of monitoring low-level contour information from the outer left eyebrow, ignoring the rest of the face), but in most tasks, such a strategy is likely to be more difficult than relying on true face perception, particularly when the location and exact nature of the possible feature changes varies from trial to trial. Thus, for shape-only changes, subjects will typically rely on the holistic face representation. Because holistic processing occurs only for upright faces, this will produce large inversion effects. Turning to *intermediate types of feature changes* (panel 3), we suggest that a plausible explanation of the intermediate inversion effects is that subjects base their response on whole-face information on some trials of the experiment (tending to give a large inversion effect) and on information outside the face system on other trials (tending to give no inversion effect). The extent to which the balance is shifted between these two types of responses—and thus, the particular size of the inversion effect in a given experiment—will depend on how often the subject attends primarily to the non-whole-face information. Where color changes are included as part of the feature manipulation, this will, in turn, depend on (1) how obvious the color changes in the stimuli are and (2) whether the subjects' attention to the location of the color changes is facilitated by repetition of items within a small stimulus set.

The idea that small inversion effects can result from attention to processing stages outside a whole-face representation is supported by a number of other findings. Barton, Zhao, and Keenan (2003) observed that an 18.5% (relative to scale) inversion effect for changes to spacing information dropped by more than half (to 8.0%) when the spacing change was still present but a color change to a feature was made simultaneously. Barton et al. (2001) observed that inversion effects for a spacing change (mouth–nose distance) *disappeared altogether* when the subject simply knew that the only change in the trial would be to this location. Similarly, Barton, Deepak, and Malik (2003) found that precuing the subject to the location of a color-only change in *featural* information (to either eyes or mouth) completely removed a moderate inversion effect that was present for this type of change without full cuing. All these results make a strong case that even when the stimulus is a whole face, subjects can change their strategies to ignore whole-face-level processing and can respond on the basis of other information in the stimulus, when other information is easily available for use.

What happens when features are shown out of a facial context? For isolated face features, inversion effects are weak or absent, even when the feature change is mostly or entirely in shape. We propose that this arises because a stimulus comprising only an isolated face feature does not activate the holistic face-processing stage. Instead, these stimuli activate only other stages of the visual-processing stream that have little sensitivity to inversion. Exactly what these stages might be is an open question. Perception of isolated face features could depend on general object recognition systems (e.g., as suggested by McKone, 2004, and Moscovitch, Winocur, & Behrmann, 1997). There could also be a stage of the *face-selective* processing system that performs part decomposition and thus is not holistic in nature (e.g., the occipital face area [OFA], as suggested by Pitcher, Walsh, Yovel, & Duchaine, 2007; see further discussion in our later section, A Twofold Contribution of Face Parts to Face Recognition?).

Summary and an implication. In summary, we have proposed a rather simple answer to the apparently difficult question of why picture-plane inversion sometimes dissociates perception of features and spacing in faces and sometimes does not. That is, we have argued that the dramatic variability in the feature inversion effect across studies depends on the extent to which the subjects in a particular experiment find that the feature task is more efficiently solved by attention to the output of a holistic face representation, versus strategic attention to information available from other visual-processing areas.

One important implication of this conclusion is that it reverses an idea common in some previous studies—namely, that color changes are a "better" manipulation of featural information than are changes also including feature shape. The rationale for that idea was that "colour changes do not alter spatial relationships between features, whereas concern has been expressed that changes in feature shape may have subtle secondary effects on second-order relations" (Barton et al., 2001, p. 531). The fact that feature shape changes have (unsubtle primary) effects on second-order relational processing we do not consider problematic; this finding simply supports one particular theory of what aspects of facial information are included in the holistic/configural/second-order relational representation of an upright face (i.e., shape as well as spacing). Instead, the results of the present review imply that it is color changes that are a poor choice, because these may not properly engage the face-processing system—in particular, when those color changes are far outside the range normal in faces or when strategic attention to color information outside the face system is otherwise encouraged by repetition within small item sets.

Testing the Theory: Data From Methods Other Than Inversion

Our review of inversion effects has led us to propose that, at some level of the visual-processing stream, there exists an integrated holistic representation of an upright face, which includes information about *all* types of shape-related information in faces. What about data from other paradigms? Do any of these refute the proposal of a holistic representation that includes both shape of major facial features and spacing between the location of those features? We argue not.

We now will review the effects of several additional variables on feature versus spacing sensitivity. Note that, in all cases, feature changes were made in a facial context.

Race and features versus spacing. Rhodes et al. (2006) used the other-race manipulation. Using shape + moderate color feature changes in conjunction with a large stimulus set size, the results showed that the reduction in memory for own-race, as compared with other-race, upright faces (i.e., the size of the other-race effect) was as large for the feature condition as for the spacing condition. Thus, feature shape and spacing were associated, not dissociated, by the effects of race.

Prosopagnosia and features versus spacing. Yovel and Duchaine (2006) examined the effects of prosopagnosia. Using shape-only feature changes, a group of 13 developmental prosopagnosics showed, for upright faces, a mean deficit, relative to controls, that was as large for discriminating feature changes as for discriminating spacing changes. Again, this indicates association, not dissociation. Dissociation was obtained only when the feature changes involved shape + moderate color with a *small* stimulus set (the "Jane" faces; Figure 1C). For this type of feature change, both Yovel and Duchaine (2006), and Le Grand et al. (2006) found weaker mean deficits in prosopagnosics than for spacing changes, a result we interpret as likely to be due to strategic attention to the location of color in the feature changes, which facilitates performance in prosopagnosics (just as it does for normal subjects on inverted faces). Overall, these results again argue for association between spacing between features and feature shape.

Individual differences and features versus spacing. A third approach has been to examine individual differences within the normal population. Yovel and Kanwisher (2008) found a strong correlation between individuals' performance for shape-only feature changes and for spacing changes, with $r = .55$ for one stimulus set (note that the upper bound on the correlation determined from the split-half reliability scores of the two separate measures was .79) and $r = .75$ on another stimulus set. Importantly, the correlation was obtained only for upright faces, and not for inverted faces (rs $< .17$). Regarding correlations with real-world face recognition ability, Rotshtein, Geng, Driver, and Dolan (2007) found that individuals' recognition of famous faces correlated as strongly with discrimination of feature changes ($r = .44$) as with discrimination of spacing changes ($r = .44$).[7] These results support association, not dissociation, between spacing and feature shape.

Development and features versus spacing. A fourth approach has been to examine childhood development. There have been no developmental studies using shape-only feature changes. Using shape + moderate color with a large stimulus set, 4-year-olds showed equal reductions, relative to adults, for both feature and spacing conditions, in the only study to test both conditions in this age group (McKone & Boyer, 2006). Also, using shape + moderate color with a large stimulus set, 6- to 7-year-olds' improvement in memory from a spacing-change enhancement of distinctiveness was as strong as adults' (the same was true for a feature change enhancement of distinctiveness); this result was obtained using a procedure in which the effects of general cognitive development were removed by equating performance for the original unchanged faces across the two age groups (Gilchrist & McKone, 2003). These two studies argue that there is no developmental dissociation. In particular, they contradict the earlier suggestion (Mondloch, Le Grand, & Maurer, 2002) that development of the ability to perceive spacing information lags behind development of the ability to perceive feature information. Also consistent with a lack of delay for spacing sensitivity are recent findings that faces differing only in spacing information can be discriminated even by human infants (Hayden, Bhatt, Reed, Corbly, & Joseph, 2007; Thompson, Madrid, Westbrook, & Johnston, 2001) and by monkeys deprived of all visual experience with faces from birth until the time of testing (Sugita, 2008). The only results suggesting a specific delay for spacing come from studies in which shape + moderate color was used with a small stimulus set.[8] Both the "Jane" faces (e.g., Mondloch et al., 2002) and another, similar set (Freire & Lee, 2001) produce later maturity of spacing sensitivity than of feature sensitivity. However, this could occur merely because it is easier for subjects to work out simple task strategies for the particular feature sets than for the spacing sets, and so children are able to discover and apply these task-specific strategies at an earlier age. In addition, it is probably relevant that neither study matched the difficulty of the feature and spacing tasks in adults (the feature changes were easier; see McKone & Boyer, 2006, for a discussion).

Congenital cataract and features versus spacing. Le Grand et al. (2001) found that congenital cataract patients show deficits, relative to controls, on detecting spacing changes but no deficits on detecting shape + moderate color feature changes in small stimulus sets (the "Jane" faces). This dissociation, however, could again be attributed to greater opportunities for nonface strategies to lift performance in the "Jane" feature set than in the spacing set. Thus, although this result does not support our hypothesis, it does not refute it either. No published data are available for congenital cataract patients tested with larger stimulus sets, shape-only feature changes, or feature and spacing tasks matched for difficulty in controls.

Spatial frequency filtering and features versus spacing. Although a common idea is that low spatial frequency (LSF) filtering (i.e., blurring) selectively removes featural information and leaves intact *configural* (i.e., spacing) information, we could locate only two studies that have directly examined the effects of spatial frequency filtering on sensitivity to feature versus spacing changes. Both used shape + small color feature changes (for which our theory would predict association, not dissociation, between features and spacing). Unfortunately, the results of the two studies are directly contradictory. The results of Boutet, Collin, and Faubert (2003) favor our theory: The spatial frequency manipulations had exactly parallel effects on feature and spacing conditions (both showed a strong deficit relative to unfiltered baseline for LSF faces, no deficit for medium spatial frequency faces, and a small deficit for high spatial frequency [HSF] faces). In contrast, Goffaux, Hault, Michel, Vuong, and Rossion (2005) found a dissociation: Detection of feature changes was reduced more strongly in LSF faces than in HSF faces, whereas detection of spacing changes was reduced slightly more in HSF than in LSF faces.

More studies are needed to resolve the contradiction, but we are inclined to place the most faith in Boutet et al. (2003). That study had good baseline matching of performance for feature and spacing conditions when faces were unfiltered (i.e., normal full spectrum), whereas the Goffaux et al. (2005) study suffered from a dramatic difference in baseline accuracy for the unfiltered faces (feature condition 95% correct, spacing condition 65% correct). Given this, the Goffaux et al. results do not provide convincing evidence of a dissociation between spacing and feature shape.

Interim summary: Behavioral studies. Overall, results from all these behavioral studies are consistent with our hypothesis that there exists an integrated holistic representation of feature shape and spacing information in faces and that weak effects of any variable on sensitivity to feature changes derive from subjects' strategically attending to information from other stages of the visual-processing stream outside the holistic representation. Specifically, as with inversion, no dissociations between feature and spacing sensitivity occur for race, prosopagnosia, individual differences, or development when the feature changes are shape-only or are shape + moderate color in large stimulus sets. Dissociations do occur when feature changes are shape + moderate color in small stimulus sets, consistent with our idea that this situation encourages attention to nonface information.

Brain imaging and features versus spacing. A final approach to studying feature versus spacing changes uses fMRI. Here, results are much more preliminary because, unfortunately, no studies have used the ideal procedure: a standard fMR adaptation or repetition priming procedure (to test directly whether a brain region can tell apart a pair of faces differing in spacing or in features), combined with localization of the brain regions of most likely interest (e.g., the fusiform face area [FFA] or the OFA). However, we suggest that findings are again consistent with the idea that there exists an integrated holistic representation at one stage within the visual system (the FFA), whereas other stages are responsible for dissociations.

Regarding the FFA, evidence suggests that this area is a plausible location for an integrated representation of features and spacing. The FFA codes individual face identity (for a review, see Kanwisher & Yovel, 2006) and shows holistic processing (the *composite effect*; Schiltz & Rossion, 2006). Regarding feature versus spacing change manipulations, Yovel and Kanwisher (2004) localized the FFA and reported total BOLD response while subjects were performing a same–different task, for shape-only feature sets and for spacing sets. FFA response was equal during both feature and spacing tasks. The FFA response was also reduced by inversion equally for each change type (consistent with the behavioral findings for the same stimuli). In two other fMRI studies that performed exploratory whole-brain analyses, responses from regions approximating the location of the FFA were again suggestive of association, rather than dissociation, between spacing and features. Maurer et al. (2007) revealed no difference between the total BOLD response during spacing and feature change tasks, even for low threshold values. Rotshtein et al. (2007) revealed evidence consistent with processing of both feature information (specifically, repetition effects for feature changes) and spacing information (specifically, correlation between BOLD response and behavioral accuracy) and concluded that information about spacing and features may converge in an area in the fusiform gyrus that overlapped with face-selective voxels (i.e., most plausibly, the FFA).

Outside the FFA, in contrast, multiple regions show clear dissociations between features and spacing. Regions responding more strongly during spacing tasks than during feature tasks include right frontal and parietal regions and a region of the fusiform adjacent to, but not overlapping, the FFA (Maurer et al., 2007). Regions showing the reverse pattern of preference for feature tasks include an object-general area in the occipital temporal cortex (lateral occipital complex; Yovel & Kanwisher, 2004), plus left frontal regions and temporal regions outside the FFA (Maurer et al., 2007). The OFA can also demonstrate sensitivity to face parts but not to spacing among parts, as indicated by a recent transcortical magnetic stimulation study (Pitcher et al., 2007).

Summary. Results from other methods agree with those from inversion. All argue that processing of information about spacing between blobs in faces is *not* necessarily dissociated from processing of local feature information. As long as the feature change is mostly or entirely in shape, the two are in fact strongly associated empirically. Spacing and features become dissociated only when the experimental stimuli or setup make it efficient for subjects to perform a behavioral task by reference to stimulus information or processing stages other than holistic face processing, or when fMRI is used to tap the responses of other stages directly (i.e., to examine brain regions outside the fusiform face area, the most likely location of holistic processing).

What Is the Nature of Shape Processing in the Holistic Face Representation?

Theoretically, these results force a reconsideration of the aspects of facial information that are included in the “special” processing for upright faces. They also imply redefinition of what should be meant by both *features* and *metric distances* in faces.

Metric distance information is based on elements far more local than the major nameable parts. Regarding the type of information included in holistic/configural/second-order relational processing, we emphasize again that, for faces, large inversion effects are a hallmark of this style of processing. Indeed, the basis for the claim that spacing information is included in the holistic/configural representation of an upright face is that spacing changes produce large inversion effects (much larger than those found for nonface objects). Following this logic, the equally large inversion effects for feature shape changes indicate that local feature shape is also a core aspect of the facial information included in the holistic/configural representation of an upright face. Thus, holistic/configural/second-order relational processing must encompass *all* shape-related information in faces, regardless of whether this comprises distances between the centers of the major features treated as blobs (spacing between blobs), or detailed shape of individual regions of the face (e.g., shape of the major features). This view is similar to that of Tanaka and Farah (1993) and differs from the assumption currently popular (e.g., Freire et al., 2000; Maurer et al., 2007) that spacing between major features has special status in terms of its contribution to holistic/configural processing.

Regarding the role of metric distances in holistic face perception, our theory is in *agreement* with the classic idea (e.g., Diamond & Carey, 1986; Rhodes, 1988) that coding of metric distance information in faces is central to the special processing for upright faces. We also hold to the standard view that holistic processing creates an integrated perceptual representation by representing such distance information simultaneously between multiple regions of the face (e.g., Rossion, 2008). Where we differ is that we argue that the definitions of both *metric distance* information, and the *features* between which these are computed, must be changed from the current typical usage. Over the last 10 years, the term *features* has become associated with only the *major nameable features* (eyes, nose, mouth, eyebrows), and *distance information* with distances only between these major features treated as blobs, so that only their *overall location* is relevant (see the Definitions section and the illustration in Figure 4A). However, we know of no evidence either (1) that the elements between which metric distances are calculated are the commonly named facial features (eyes, nose, mouth) or (2) that only distances between their central locations are computed. The results of the present review instead imply that holistic processing must represent distance information between elements far more local than the major nameable parts.

Two speculations on exactly what holistic processing might involve. If holistic processing derives from elements more local than the major parts, exactly what are the elements, and exactly how might the process of constructing a holistic representation work? This remains an open question. We suggest two possibilities as starting points for debate.

One idea is that the features might be better thought of as multiple *landmark points* (cf. Rhodes, Brennan, & Carey, 1987), of the type that, for example, researchers would mark on faces when making face morphs. As is illustrated in Figure 4C, landmark points mark locations on the face, such as the outer corner of the left eye, the location where the groove running down below the nose hits the top lip, the rightmost point on the right nostril, the point of forward-most projection of the cheek bone (particularly important for 3-D morphs), and so on. Our suggestion is, then, that the special form of perceptual integration that constitutes holistic processing might be based on simultaneous processing of *multiple* distances computed between these *multiple* landmark points. This might include, for example, the distance between the outer corner of the left eye and the inner corner of the left eye, between the outer corner of the left eye and the tip of the chin, between the outer corner of the left eye and the forward-most projection of the cheek bone, and so on (and so on). Critically for explaining the results of the present review, computation of such distances would allow the holistic representation to code full details of the shape of the major nameable features (e.g., nose shape), thus accounting for the large inversion effects that arise from shape-only manipulations that do not affect spacing between blobs.

Within this model, we hold no strong views on which particular distances might then form the basis for holistic processing, beyond the fact that it is unlikely that this is merely a *small set* of *short-range distances* within the face. Sekunova and Barton (2008) showed that a single short-range change in spacing (e.g., the eyes moved down while the eyebrow location remains fixed) produced a smaller inversion effect than did a change lacking local spatial references (e.g., eyes and eyebrows moved down together). This argues that holistic processing of upright faces allows an accurate representation of distance information over a longer range than is possible for inverted faces. This might imply that the metric information used to compute a holistic representation is long range (e.g., corner of eye to tip of chin). Alternatively, effective coding of long-range information could potentially be achieved by summing the outputs of multiple short-range distances.

Our second speculation is that holistic processing could derive from a coding in which there is no decomposition of the image at shape-related boundaries at all.[9] For example, the input to holistic processing might be a list of pixel intensities covering the whole face, rather than a set of parts or a set of landmark points. If holistic processing were derived from this type of input, the representation of metric information within the face (e.g., distance between the corner of the left eye and the tip of the nose) would be purely implicit, not explicit.

A final important point to note is that, under either the landmark point or the no-decomposition theory, the holistic mechanism must be restricted to upright faces, in order to explain the large inversion effects and the evidence from other paradigms (e.g., the composite effect; McKone, 2008; Young et al., 1987) that holistic processing is limited to upright faces. Of course, there is nothing in either of these theories that intrinsically explains *why* holistic processing is limited to upright faces, but we note that the same is true for the spacing-between-blobs theory, which again simply states that holistic processing occurs only for upright faces, without explaining why.[10]

Links with the computational face recognition literature. Readers may note that the specific speculations we have made here bear some relationship to models in the computational face recognition literature. Computational models have been proposed that take as their input either landmark points (e.g., Rajapakse & Guo, 2001) or lists of pixel intensities (e.g., principal component analysis approach; Hancock, Burton, & Bruce, 1996). In general, however, computational models make no reference to holistic processing; typically, their aim is to build an engineering system that recognizes faces, rather than to understand how the human brain solves this problem. Correspondingly, most models fail to explain inversion effects. For example, both the landmark-points-based model of Rajapakse and Guo and the principal component analysis approach work equally as well for inverted faces as for upright faces. We are aware of only one computational model that specifically claims to model holistic processing (Schwaninger, Wallraven, & Bülthoff, 2004); unfortunately, it too appears to use a computational procedure that would work just as well for inverted faces as for upright faces. Inversion effects have, however, been produced by the model of Jiang et al. (2006).

These observations indicate that the computational implementation of holistic processing is a topic ripe for investigation.

General advantages of our theory. Although many questions remain about specifics, we emphasize that our theory—that holistic processing is based on elements

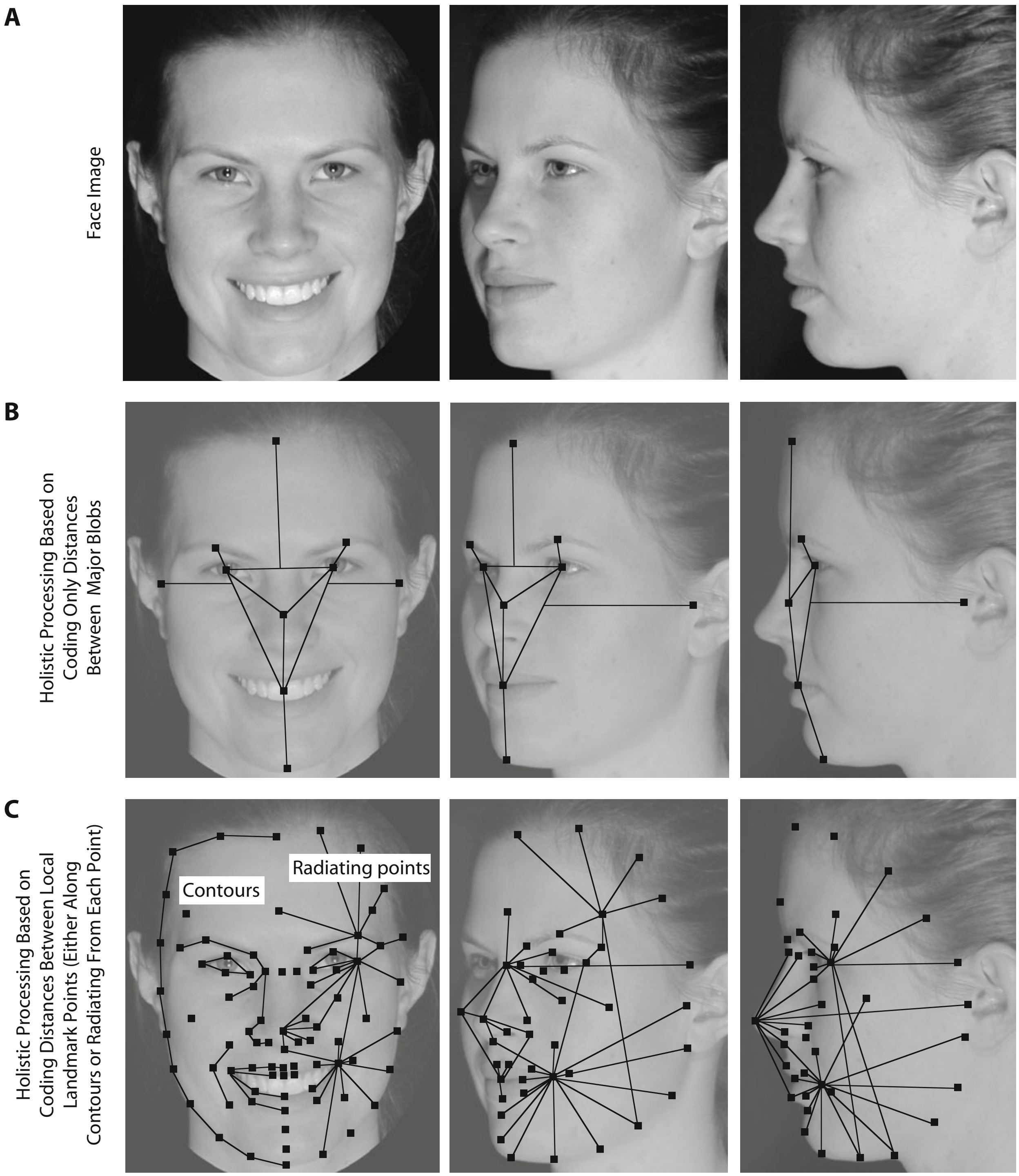

Figure 4. Theoretical ideas regarding the facial information encompassed by holistic/configural/second-order relational processing. (A) Face image. (B) Holistic processing based on spacing between the centers of major nameable facial features (interocular distance, nose–mouth distances, plus overall position of features in head), excluding details of local feature shape. (C) A *landmark points* approach, which assumes that "features" far more local than the major nameable parts are the key locations between which metric distances are computed to form the holistic representation. Note the much richer shape information coded by holistic processing in this second theory.

far more local than the major nameable parts—is necessary to explain the large inversion effects for feature shape changes. Furthermore, our theory has, at a general level, several other potential advantages over the spacing-between-blobs theory.

First, it assumes that all shape information (at all spatial scales) that is useful for individuating faces is retained in the holistic representation. In contrast, the traditional distances-between-blobs theory (Figure 4B) implies that a vast amount of potentially useful information about face shape is thrown away at this level.

Second, it allows that holistic/configural/second-order relational processing includes not only metric distances between *internal* landmark points on the face, but also information about distances between regions in the face interior and the face outline (thus allowing, e.g., representation of chin and cheek shape). A need to include face outline as part of holistic face processing is indicated by evidence that combinations of the internal features of one person with the external features of another are perceived as a new identity, upright but not inverted (Cohen & Cashon, 2001; Young et al., 1987).

Third, our approach appears more suitable for understanding how holistic processing might operate for nonfrontal views of faces. Holistic processing is as strong for profile views of faces as for front views (McKone, 2008). Under the spacing-between-blobs view, this should be impossible, because the profile view occludes almost all of the information that the theory implies would be necessary to compute a holistic representation: Specifically, missing information includes the interocular distance and, for the occluded eye, the eye–mouth, eye–nose, and eye–eyebrow distances. Although our landmark points or image-derived proposals in no way solve the difficult task of understanding how faces are recognized across view changes, they do at least allow that there will be enough information left in a profile view to compute a holistic representation. Indeed, there will be as much information available in a profile view as in a front view, it will merely be different information (e.g., the forward projection of the nose becomes available, to make up for losing information about nose width).

Overall, there would seem to be many theoretical advantages to an approach that allows holistic processing to code the full face structure, rather than merely spacing between a few face blobs.

Is Information Other Than Shape Included in the Holistic Representation?

Because of the particular manipulations in the experimental studies available, we have focused in the present article on the idea that the holistic representation of an upright face includes all aspects of facial *shape*. We do not rule out the idea that other important information about facial appearance is also included. This might include texture (related to age). Indeed, it might also include color-related information, such as general skin tone (e.g., olive, chocolate, pinkish, freckles), eye color, or skin–lip contrast. Our focus on lack of inversion effects when color changes are included in the stimulus manipulation should not be taken as evidence that facial color is not represented in the holistic face representation—merely that it is difficult to design behavioral procedures that directly tap *face*-level color processing, rather than color processing in low-level vision.[11]

A Twofold Contribution of Face Parts to Face Recognition?

We have presented strong evidence that local feature shape contributes to the holistic representation of faces. We also note that there may be an additional contribution of local feature shape to total face recognition, via an additional (possibly parallel) route. There is good evidence that, in addition to faces being processed holistically, they can also be decomposed into parts (e.g., leading to above-chance recognition of scrambled faces [Tanaka & Farah, 1993] or supporting eyebrow-matching strategies in prosopagnosia and for inverted faces [Duchaine & Weidenfeld, 2003; Robbins & McKone, 2003]). Several authors have therefore proposed that total face recognition is based on the summed outputs of *holistic* and *part-based* routes (e.g., McKone, 2004; Moscovitch et al., 1997; Schwaninger et al., 2004). An interesting open question is whether this part decomposition contribution derives from general object recognition mechanisms or from a part decomposition module or stage specifically within the face recognition system. A recent finding that the other-race effect (poorer memory for other- than for own-race faces) occurs for isolated position-scrambled face parts (Hayward, Rhodes, & Schwaninger, 2008) is suggestive that part decomposition might occur within the face system (although note that the result could also derive from a general familiarity effect within the object recognition system).

Implications for Terminology

There have been longstanding discrepancies in terminology used to refer to the "special" form of processing for upright faces. Of the terms *holistic*, *configural*, and *second-order relational*, we suggest, on the basis of the present results, that *holistic* is the most appropriate, for two reasons. First, *holistic* is the only term that avoids any implication of some special status for information about spacing between the centers of major facial features; both *configural* and *second-order relations* have been used in many articles as synonymous with processing of, information about, or changes to spacing between blobs. Second, *holistic* is the term that best captures the idea of the *integration* of *all* the shape-related information in an individual's face.

Our other suggestion for terminology change is from *second-order relations* to *second-order deviations*. We agree with the standard view (derived from Diamond & Carey, 1986) that holistic processing codes the second-order ways in which an individual face deviates from the first-order structure shared by all faces (eyes above nose above mouth in an oval-ish outline). However, we have argued that holistic processing codes how an individual deviates from the average in *any* type of face shape information, not merely in spacing between the centers of major features. But because *relations* and *second-order relations* have been used in many articles to refer specifically to the latter type of spacing information, we suggest that rather than attempting to alter this common usage, it is more feasible to introduce the new term *second-order deviations* to capture the idea of all deviations being coded.

Conclusion

We have presented a comprehensive review of the literature relevant to the size of the inversion effect for *feature* changes in faces and the related question of whether this is smaller than the inversion effect for *spacing* changes. The results clearly reject the widespread view that feature inversion effects are typically only small, and much smaller than spacing inversion effects. Instead, feature inversion effects varied dramatically across studies, with only 35% of the studies reviewed (8/22) finding the classical pattern of no feature inversion effect. Another 40% (9/22) showed very large feature inversion effects—as large as or larger than typical spacing inversion effects—and the remainder showed intermediate feature inversion effects. Where features are presented in a facial context (the standard procedure), we argued that most of this variability could be attributed to the extent to which the feature change was in shape only (large inversion effects) versus color (small inversion effects), with a subsidiary effect within shape + moderate color changes of whether the stimulus set included a large number of items that were not repeated (moderate inversion effects) or a small number of items repeated multiple times (small inversion effects). We also noted evidence that manipulations that produce large feature inversion effects do so only if the feature is shown as part of a face, and not if shown in isolation. Overall, the results showed that the shape of features and the spacing among them are not dissociated in upright whole faces. Importantly, this conclusion is consistent with the findings of studies in which other methodologies were used, including development, neuroimaging, individual differences, and neuropsychology.

Theoretically, we have argued that these results mean that there exists, at some level of the visual-processing stream, a truly *holistic* representation of upright faces that integrates all details of the shape-related information for an individual face. Correspondingly, we have argued that, at least from the perspective of holistic/configural/second-order relational processing, there is no theoretical value in continuing to make distinctions between the processing of spacing between blobs and the processing of detailed shape of the major nameable facial features. Instead, we have argued that the metric distance information central to holistic processing is based on far more local elements of the face than merely a few major "blobs."

AUTHOR NOTE

This study was supported by grants from the Australian Research Council (DP0450636) to E.M. and by Marie Curie-IRG 046448 to G.Y. We thank Daphne Maurer, Helmut Leder, and Max Riesenhuber for kindly providing copies of their stimuli; Gillian Rhodes for helpful comments on an early draft; and Kate Crookes for assistance preparing the manuscript. Correspondence concerning this article should be addressed to E. McKone, Department of Psychology, Australian National University, Canberra, ACT 0200, Australia (e-mail: elinor.mckone@anu.edu.au).

REFERENCES

Barton, J. J. S., Deepak, S., & Malik, N. (2003). Attending to faces: Change detection, familiarization, and inversion effects. *Perception*, **32**, 15-28.

Barton, J. J. S., Keenan, J. P., & Bass, T. (2001). Discrimination of spatial relations and features in faces: Effects of inversion and viewing duration. *British Journal of Psychology*, **92**, 527-549.

Barton, J. J. S., Zhao, J., & Keenan, J. P. (2003). Perception of global facial geometry in the inversion effect and prosopagnosia. *Neuropsychologia*, **41**, 1703-1711.

Boutet, I., Collin, C., & Faubert, J. (2003). Configural face encoding and spatial frequency information. *Perception & Psychophysics*, **65**, 1078-1093.

Boutet, I., & Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory & Cognition*, **34**, 854-864.

Cohen, L. B., & Cashon, C. H. (2001). Do 7-month-old infants process independent features of facial configurations? *Infant & Child Development*, **10**, 83-92.

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

Duchaine, B. C., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia*, **41**, 713-720.

Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year-olds: Processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology*, **80**, 347-371.

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, **29**, 159-170.

Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: Local and relational distinctiveness effects in 7-year-olds. *Visual Cognition*, **10**, 769-793.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, **34**, 77-86.

Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception & Performance*, **33**, 995-1001.

Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: Human perception and principal components analysis. *Memory & Cognition*, **24**, 26-40.

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology*, **97**, 85-98.

Hayward, W. G., Rhodes, G., & Schwaninger, A. (2008). An own-race advantage for components as well as configurations in face recognition. *Cognition*, **106**, 1017-1027.

Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., & Riesenhuber, M. (2006). Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron*, **50**, 159-172.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Proceedings of the Royal Society B*, **361**, 2109-2128.

Leder, H., & Bruce, V. (1998). Local and relational aspects of face distinctiveness. *Quarterly Journal of Experimental Psychology*, **51A**, 449-473.

Leder, H., & Bruce, V. (2000). When inverted faces are recognized: The role of configural information in face recognition. *Quarterly Journal of Experimental Psychology*, **53A**, 513-536.

Leder, H., & Carbon, C.-C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology*, **97**, 19-29.

Le Grand, R., Cooper, P. A., Mondloch, C. J., Lewis, T. L., Sagiv, N., de Gelder, B., & Maurer, D. (2006). What aspects of face processing are impaired in developmental prosopagnosia? *Brain & Cognition*, **61**, 139-158.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, **410**, 890.

Malcolm, G. M., Leung, C., & Barton, J. J. S. (2004). Regional variation in the inversion effect for faces: Differential effects for feature shape, feature configuration, and external contour. *Perception*, **33**, 1221-1231.

Martini, P., McKone, E., & Nakayama, K. (2006). Orientation tuning of human face processing estimated by contrast matching on transparency displays. *Vision Research*, **46**, 2102-2109.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, **6**, 255-260.

Maurer, D., O'Craven, K. M., Le Grand, R., Mondloch, C. J., Springer, M. V., Lewis, T. L., & Grady, C. L. (2007). Neural correlates of processing facial identity based on features versus their spacing. *Neuropsychologia*, **45**, 1438-1451.

McKone, E. (2004). Isolating the special component of face recognition: Peripheral identification and a Mooney face. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 181-197.

McKone, E. (2008). Configural processing and face viewpoint. *Journal of Experimental Psychology: Human Perception & Performance*, **34**, 310-327.

McKone, E. (in press). Face and object recognition: How do they differ? In V. Coltheart (Ed.), *Tutorials in visual cognition*. Hove, U.K.: Psychology Press.

McKone, E., & Boyer, B. L. (2006). Sensitivity of 4-year-olds to featural and second-order relational changes in face distinctiveness. *Journal of Experimental Child Psychology*, **94**, 134-162.

McKone, E., Crookes, K., & Kanwisher, N. (in press). The cognitive and neural development of face recognition in humans. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences IV* (4th Ed.). Cambridge, MA: MIT Press, Bradford Books.

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 573-599.

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, **31**, 553-566.

Mondloch, C. J., Maurer, D., & Ahola, S. (2006). Becoming a face expert. *Psychological Science*, **17**, 930-934.

Mondloch, C. J., & Thomson, K. (2008). Limitations in four-year-old children's sensitivity to the spacing among facial features. *Child Development*, **79**, 1514-1524.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, **9**, 555-604.

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

Pellicano, E., & Rhodes, G. (2003). Holistic processing of faces in preschool children and adults. *Psychological Science*, **14**, 618-622.

Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science*, **9**, 270-277.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. C. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology*, **17**, 1568-1573.

Rajapakse, M., & Guo, Y. (2001). Multiple landmark feature point mapping for robust face recognition. In J. Bigün & F. Smeraldi (Eds.), *Proceedings of Audio- and Video-Based Biometric Person Authentication, Third International Conference, AVBPA 2001* (pp. 95-101). Berlin: Springer.

Rakover, S. S., & Teucher, B. (1997). Facial inversion effects: Parts and whole relationship. *Perception & Psychophysics*, **59**, 752-761.

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

Rhodes, G., Brake, S., & Atkinson, A. P. (1993). What's lost in inverted faces? *Cognition*, **47**, 25-57.

Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures: Implications for mental representations of faces. *Cognitive Psychology*, **19**, 473-497.

Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, **13**, 499-505.

Riesenhuber, M., Jarudi, I., Gilad, S., & Sinha, P. (2004). Face processing in humans is compatible with a simple shape-based model of vision. *Proceedings of the Royal Society B*, **271**, S448-S450.

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

Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, **128**, 274-289.

Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: Behavioral and functional magnetic resonance imaging data. *Journal of Cognitive Neuroscience*, **19**, 1435-1452.

Russell, R., Biederman, I., Nederhouser, M., & Sinha, P. (2007). The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research*, **47**, 157-165.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, **32**, 1385-1394.

Schwaninger, A., Wallraven, C., & Bülthoff, H. H. (2004). Computational modeling of face recognition based on psychophysical experiments. *Swiss Journal of Psychology*, **63**, 207-215.

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 904-915.

Sekunova, A., & Barton, J. J. S. (2008). The effects of face inversion on the perception of long-range and local spatial relations in eye and mouth configuration. *Journal of Experimental Psychology: Human Perception & Performance*, **34**, 1129-1135.

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences*, **105**, 394-398.

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

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

Thompson, L. A., Madrid, V., Westbrook, S., & Johnston, V. (2001). Infants attend to second-order relational properties of faces. *Psychonomic Bulletin & Review*, **8**, 769-777.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology*, **43A**, 161-204.

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

Yovel, G., & Duchaine, B. C. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, **18**, 580-593.

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, **44**, 889-898.

Yovel, G., & Kanwisher, N. (2008). The representations of spacing and part-based information are associated for upright faces but dissociated for objects: Evidence from individual differences. *Psychonomic Bulletin & Review*, **15**, 933-939.

NOTES

1. Tanaka and Sengco (1997) introduced a variant that added a spacing change between features. However, we could not include the results of that procedure with the 17 features-versus-spacing studies, because the Tanaka and Sengco manipulation changes both a feature *and* spacing (i.e., it is feature + spacing, not spacing only).

2. We also have a more theoretical problem with meta-analysis, which is that in basing comparison across studies on percentage of variance explained, it assumes that the variance available to be explained is constant across studies. However, this is not always true: For example, a study containing 60 trials per condition per subject will produce lower

variance across subjects than will a study containing 10 trials per condition per subject, and *statistical effect size* measures do not take this into account.

3. However, we note that, with the exception of Leder and Bruce (1998) and McKone and Boyer (2006), no other studies have equated spacing- and feature-altered faces on distinctiveness either.

4. This should not be taken to mean that it is of no value to match performance for upright faces across feature and spacing conditions. Logically, it is always more straightforward to compare the size of effects when baseline performance in the two conditions is matched.

5. For two of the studies, the feature inversion effect does not appear so large on the relative-to-scale measure (Figure 2A). However, we have preferred the relative-to-spacing measure as more valid for these particular studies, because they used a distinctiveness rating task and, across our review, spacing inversion effects, relative to scale, were smaller for distinctiveness rating tasks ($M = 11.9\%$) than for other tasks (all other tasks, $M = 29.3\%$).

6. It is worth noting that these results from Rhodes et al. (1993) have been almost universally miscited. They have regularly been described as showing a strong feature-versus-spacing × inversion interaction and a lack of feature inversion effect, without noting that this pattern occurred only for features in isolation, and not for features in a facial context.

7. For recognition of novel faces in the same study, the positive correlation with feature discrimination did not reach significance. However, this cannot be taken as reliable evidence of a *lack* of relationship, given that the sample size was small (subjects were also scanned using fMRI). Note that Yovel and Kanwisher's (2008) strong correlation between features and spacing was obtained using unfamiliar faces.

8. Several other studies from Cathy Mondloch's laboratory have indicated poor performance on spacing tasks in young children (e.g., 4 or 6 years old). These studies are not reviewed here, because they did not include a feature change control condition and, without such a control, it remains unclear to what extent the late development of task performance reflects spacing-specific perceptual delay, as opposed to general cognitive limitations in young children that affect demanding perceptual discriminations (e.g., concentration failures). Indeed, after 8 years at least, there is evidence directly suggesting that development of performance on spacing tasks is due to general cognitive development, not development of perceptual coding related to face identification: The rate of development for spacing sensitivity is the same for monkey faces (for which humans show poor identification) and for human faces (Mondloch, Maurer, & Ahola, 2006).

Mondloch and Thomson (2008) have also criticized studies showing early spacing sensitivity (e.g., in infants and preschoolers) on the grounds that the spacing changes were "out of the normal range." This criticism is valid, however, only if it is clear that the feature changes that the preschoolers did respond to were not also equally far from the norm. There has been only one attempt to match spacing and feature changes for where they fall with respect to the normal range (based on adult perceptions). McKone and Boyer (2006) equated this by matching adult-rated distinctiveness of the feature-changed and spacing-changed versions of the same set of original faces. The results showed equally accurate spacing and feature discrimination in 4-year-olds.

9. We thank William Hayward for drawing our attention to this possibility.

10. One possible explanation is that holistic processing is somehow driven by an innate representation of faces. Recent evidence has made a strong case for innate representation of upright face structure (e.g., Sugita, 2008; for a review, see McKone, Crookes, & Kanwisher, in press).

11. Possibly consistent with a role for facial color in holistic processing, Russell, Biederman, Nederhouser, and Sinha (2007) found large inversion effects for a *reflectance* change in the whole face. However, this result is difficult to interpret, given that the particular method used to manipulate reflectance also caused quite substantial changes in part shape.

(Manuscript received November 3, 2008;
revision accepted for publication April 30, 2009.)