Journal of Experimental Psychology:
Human Perception and Performance
2000, Vol. 26, No. 2, 527–551

0096-1523/00/$5.00 DOI: 10.1037//0096-1523.26.2.527

Configural Information in Facial Expression Perception

Andrew J. Calder
Medical Research Council Cognition
and Brain Sciences Unit

Andrew W. Young
University of York

Jill Keane
Medical Research Council Cognition
and Brain Sciences Unit

Michael Dean
University of Sheffield

Composite facial expressions were prepared by aligning the top half of one expression (e.g., anger) with the bottom half of another (e.g., happiness). Experiment 1 shows that participants are slower to identify the expression in either half of these composite images relative to a "noncomposite" control condition in which the 2 halves are misaligned. This parallels the composite effect for facial identity (A. W. Young, D. Hellawell, & D. C. Hay, 1987), and like its identity counterpart, the effect is disrupted by inverting the stimuli (Experiment 2). Experiment 3 shows that no composite effect is found when the top and bottom sections contain different models' faces posing the same expression; this serves to exclude many nonconfigural interpretations of the composite effect (e.g., that composites are more "attention-grabbing" than noncomposites). Finally, Experiment 4 demonstrates that the composite effects for identity and expression operate independently of one another.

Bruce and Young's (1986) functional model of face recognition postulates separate parallel routes for the processing of facial identity (who the person is) and facial expression (what they are feeling). Over the years, this dissociation has been investigated by a number of studies using a range of different methodologies. These include cognitive studies of neurologically normal participants (Campbell, Brooks, de Haan, & Roberts, 1996; Young, McWeeny, Hay, & Ellis, 1986), double dissociations in brain-injured participants (Parry, Young, Saul, & Moss, 1991; Young, Newcombe, de Haan, Small, & Hay, 1993), single-cell recording in nonhuman primates (Hasselmo, Rolls, & Baylis, 1989), and, in more recent years, functional imaging studies of brain activation (George et al., 1993; Sergent, Ohta, MacDonald, & Zuck, 1994). Together, these studies provide substantial support for the idea that facial identity and facial expression recognition are dissociable cognitive functions, and this is perhaps one of the reasons why these two facial attributes have so often been the topics of separate examination. But their isolated investigation is possibly less to do with their proposed functional independence and more to do with the fact that traditionally, facial identity and facial expression processing have been studied within separate domains of psychology.

In general, facial expression recognition has been studied within a social psychology framework, where research has focused on the communicative value of signals of facial affect rather than their perceptual representation. Studies of facial identity processing, however, have been heavily influenced by research in cognitive psychology, and consequently, a firm emphasis has been placed on understanding the perceptual mechanisms involved. In the last 20 years, then, there has been an enrichment in our understanding of the perceptual representation of facial identity, whereas the perceptual mechanisms underlying facial expression recognition have not been so extensively investigated. Hence, although the work of Ekman and his colleagues has greatly enhanced our understanding of the anatomy used to produce facial expressions, knowledge of the perceptual processes needed to decode them remains scant.

Recent research has aimed to redress this imbalance (Calder, Young, Perrett, Etcoff, & Rowland, 1996; Calder, Young, Rowland, & Perrett, 1997; Ellison & Massaro, 1997; Etcoff & Magee, 1992; Young et al., 1997) by taking two approaches. First, these studies have built on the strong knowledge base provided by the social psychology literature, and second, they have applied perceptual paradigms developed within other areas of psychology to the study of facial affect processing. This latter approach has the added advantage of using tried and tested methods, and for the reasons outlined above, the facial identity literature provides a particularly good source of perceptual paradigms. Examples of these include the following: effects of stimulus orientation (Diamond & Carey, 1986; Farah, Tanaka, & Drain, 1995; Valentine, 1988), feature displacement (Haig, 1984), distinctiveness effects (Rhodes, Brennan, & Carey,

Andrew J. Calder and Jill Keane, Medical Research Council Cognition and Brain Sciences Unit, Cambridge, England; Andrew W. Young, Department of Psychology, University of York, Heslington, England; Michael Dean, Department of Psychology, University of Sheffield, Sheffield, England.

We are grateful to P. Ekman for giving us permission to use pictures from the Ekman and Friesen (1976) Pictures of Facial Affect series. We would also like to thank Brian Cox and Gary Jobe for their assistance in preparing figures.

Correspondence concerning this article should be addressed to Andrew J. Calder, Medical Research Council Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 2EF, England. Electronic mail may be sent to andy.calder@mrc-cbu.cam.ac.uk.

1987; Valentine, 1991), and image negation (Bruce & Langton, 1994; Hill & Bruce, 1996), all of which have provided valuable clues to how facial identity is coded. But perhaps the most consistent result to emerge from the facial identity literature is the important role of configural information in face recognition (Bruce, Doyle, Dench, & Burton, 1991; Carey & Diamond, 1977; Rhodes, 1988; Tanaka & Farah, 1993; Young, Hellawell, & Hay, 1987). It is highly pertinent, then, for us to investigate what role, if any, configural information may play in facial expression recognition.

Carey and Diamond (1977) introduced the term *configural information* to mean the interrelationship between different facial features (e.g., the relative shape and positioning of the mouth in relation to the shape and positioning of the nose, eyes, etc.); this type of facial information is seen as distinct from the structure and shape of individual features (e.g., eye, nose, mouth shape, etc.). Diamond and Carey (1986) identified two forms of configural information that they referred to as *first-order* and *second-order relational properties.* The former type refers to the raw inter feature relationships that are common to all normal faces—two horizontally positioned eyes, above a central nose, above a central mouth, etc.; effectively the spatial information that makes a face a face. Second-order relational properties are substantially more subtle and are what are more generally referred to as simply *configural features.* These features are the interrelationships between different feature positions and shapes that help distinguish one facial identity from all others (e.g., the distance between the eyes, position and shape of the nose in relation to the position and shape of the mouth, etc.).

The current consensus in facial identity research is that configural features are particularly important for face recognition; however, individual features may also contribute to some extent. We refer to this view as the configural model. Here, we investigate its applicability to the perception of facial signals of emotion. It worth mentioning that Tanaka and Farah (1993) have distinguished the configural model from their holistic model of face processing. For this latter model, it is proposed that faces are coded as Gestalt representations in which the constituent parts (eyes, nose, mouth, etc.) are not "explicitly represented." In support of their model, Tanaka and Farah showed that a single facial feature (eyes, nose, or mouth) is more readily identified as belonging to a particular person's face when it is shown in the context of the whole face, than when shown in isolation. The same was not shown to be true, however, of scrambled faces, inverted faces, or a set of structurally homogeneous houses (made up of doors and windows in place of facial features).

Recently, Ellison and Massaro (1997) have shown that Tanaka and Farah's (1993) holistic model is not applicable to facial affect recognition (see below). Instead, they suggest that their data are consistent with the antithesis of this model, one in which facial expressions are represented and identified in terms of their individual parts, or features (e.g., eye, nose, and mouth shape, etc.)—what we refer to as the part-based model.

Ellison and Massaro (1997) used facial expressions displayed on a synthetic (computer-generated) face in which just two facial features, the eyebrows and the corners of the mouth, were manipulated. The stimuli were produced by combining five levels of eyebrow displacement (ranging between eyebrows raised and eyebrows flattened) and five levels of mouth displacement (ranging between corners of the mouth turned up, and corners of the mouth turned down). *Prototype* expressions of happiness and anger were defined as eyebrows maximally raised with mouth corners maximally curled up, and eyebrows maximally flattened with mouth corners maximally curled down, respectively. All other combinations of the five mouth and five eyebrow displacements were generated to give a total of 25 full-face images. In addition, the five levels of eyebrow and five levels of mouth features were presented individually in the context of the upper and lower sections of the face, respectively. The participants' task was to decide whether each image signaled a happy or an angry expression.

By modeling their data using Massaro and colleague's fuzzy logical model of perception (FLMP; Massaro, 1998; Massaro & Cohen, 1990), Ellison and Massaro (1997) showed that participants' responses to the whole-face images could be reliably predicted from their responses to the half-face images. Consequently, they argued that their results were inconsistent with the holistic model (as defined by Tanaka and Farah, 1993). However, they pointed out that although their results provided no direct support for the configural model, they did not rule out the idea of configural encoding of facial affect altogether. Instead, they suggested that if configural features are used in the representation and recognition of facial expressions, their results demonstrated that they are unlikely to involve the spatial relationships between the features manipulated in their stimuli (eyebrows and mouth corners). But it is also worth considering that Ellison and Massaro may have failed to find evidence of configural processing because of the particular design and stimuli they used.

For example, Ellison and Massaro (1997) used facial expressions that were generated on a single synthetic face in which only the eyebrows and mouth corners were manipulated. Under these circumstances, the participants may have been able to treat these two altered features as separate objects, basing their decisions on their individual shapes rather than a more global impression of the face. It is also worth noting that manipulating one facial feature in a human face can often have secondary consequences for other features. For instance, changing the positions of the eyebrows can cause the brow to become wrinkled or furrowed, and manipulating the shape of the mouth can affect the shape of the cheeks. The fact that these more global changes were not present in the synthetic expressions used by Ellison and Massaro may also have served to minimize the configural encoding of these images.

In addition, the idea that configural information is important for facial expression recognition is not completely unfounded. In an investigation of the Thatcher illusion, Parks, Coss, and Coss (1985) found that the judged pleasantness of upright and inverted smiling mouths was affected by two factors: (a) the location of the eyes in relation to the mouth (above or below), and (b) the distance between the eyes and the mouth; pleasantness ratings of the eyes showed a strikingly parallel pattern. Hence, even though the participants were being asked to rate just one facial feature (eyes or mouth), the configuration of the face influenced their judgment of the feature. In a separate study, Wallbott and Ricci-Bitti (1993) presented participants with single muscular movements (action units) in the context of an otherwise neutral face, and combinations of action units. The participants task was to rate the emotional intensity of the resultant expressions on seven scales (Happiness, Sadness, Anger, Fear, Disgust, Surprise, and Contempt). Wallbott and Ricci-Bitti found that the meaning of most single action units changes when presented in combination with other action units, and only a few action units transmit a specific emotional meaning that is retained across different contexts. Again, these results point to a role of configural processing in facial affect recognition, a role that Ekman and Friesen (1975) also identified, although not empirically, in their book *Unmasking the Face:* "With many facial expressions a change in just one area gives the impression that the rest of the facial features have changed as well" (p. 39).

Given the above observations, we felt that it was possible that evidence of configural processing of emotional facial expressions might be found using a different design to one used by Ellison and Massaro (1997).

The Composite Paradigm

Earlier we mentioned that contemporary facial expression research is in the fortunate position of being able to borrow tried and tested methodologies from the facial identity literature. Consequently, we felt that the most direct method of distinguishing between configural and part-based models of facial expression recognition was to adopt a paradigm that has been described by Bruce (1988) as "[a] compelling illustration of the power of configural processing of faces" (p. 41), the facial composite phenomenon originally shown by Young et al. (1987).

The composite effect shows that when the top half of one face is aligned with the bottom half of another's, the two halves fuse to create a perceptually "new" (composite) face (Figure 1). Consequently, people are significantly slower to name the top or bottom segments of these composite faces relative to a control condition in which the two halves are *misaligned* (noncomposite condition; Figure 1) so that they do not form a face shape. Young et al. (1987) suggested that this effect can be explained in terms of the important role that configural features play in facial identity recognition. In the composite condition, the top and bottom halves of two different faces align to form a novel configuration, and this interferes with the recognition of the identity shown in either of the two halves; that is, the novel configuration does not match the configural information for either the top or bottom identity. Misaligning the two halves, however, means that the image is no longer encoded as a configural whole, and the separate parts of the face can be accessed without interference from an inappropriate configuration. In a second experiment Young et al. (1987) bolstered this interpretation by showing that the composite effect is abolished when the stimuli are inverted (i.e., rotated by 180°; see also Carey & Diamond, 1994). This second finding is consistent with Carey and Diamond's (1977) earlier observation that configural information is more difficult to encode from inverted faces.

The advantage of the composite paradigm is that the same physical features (i.e., the top and bottom sections of the face) are present in both conditions (composite and noncomposite). The only difference between the two conditions is whether the two halves are aligned, to form a face, or misaligned, so that they do not. Consequently, if responses are slower for the composite condition, this demonstrates that the composite images are being processed differently to the noncomposites. In facial identity research, a number of investigators concur with Young et al.'s (1987) idea that

Figure 1. The composite effect shown by Young, Hellawell, and Hay (1987). The top half of one face is aligned with the bottom half of another's to create a "new" facial identity (composite). Young et al. (1987) showed that the top and bottom segments of faces are easier to identify in the misaligned (noncomposite) condition than in the aligned (composite) condition.

slower reaction times (RTs) for the composite condition can be attributed to a disruption of configural encoding (Bruce, 1988; Carey & Diamond, 1994; Endo, Masame, & Maruyama, 1989; Endo, Takahashi, & Maruyama, 1984; Hole, 1994). The composite effect, then, seems a highly appropriate paradigm to distinguish between configural and part-based models of facial expression processing.

Interestingly, historical research shows that Young et al. (1987) were not the first to use composite faces. They had originally been used some 60 years earlier for facial expression research (Dunlap, 1927). Here, they were not used to examine configural processing, however, but rather the relative contribution of the upper and lower face regions in expression recognition. For example, in one experiment, Dunlap presented his participants with frames containing four faces; two of the faces were posing different prototype expressions (selected from the list *natural, amusement, mirth, startle, expectation, pain, disgust, grief, strain,* and *relaxation*), and two were composite facial expressions prepared by combining the top half of one prototype with the bottom half of the other. For each composite expression, the participants were asked to decide which of the two prototype expressions it resembled most. The results showed that on 80% of the trials, participants selected the prototype that corresponded to the bottom half of the composite. Consequently, Dunlap concluded that the bottom region of the face is more important for facial expression recognition.

Since Dunlap (1927), other studies have addressed the issue of upper versus lower face dominance in emotion recognition, and the majority of these were reviewed by Ekman, Friesen, and Ellsworth (1972). Ekman et al. discussed the fact that Dunlap's findings proved difficult to replicate (Coleman, 1949; Frois-Wittmann, 1930), and subsequent investigations of this issue generally have found that the emotion is more readily recognizable from the upper face region for some facial expressions and the lower face region for others (Bassili, 1979; Hanawalt, 1944; Plutchik, 1962). Hence, these studies suggest that there are what we refer to as facial expressions with a *recognizable-top* or *recognizable-bottom* half.

For our own purpose of investigating a composite effect for facial expression, the results of these latter studies are highly relevant. This is because the participants' task in the composite paradigm is to identify the expressions in one half (top or bottom) of the composite or noncomposite images. Hence, by using composites prepared from the top segments of recognizable-top expressions and the bottom sections of recognizable-bottom expressions, we could ensure that the task was readily accomplishable.

The facial expressions used in this study were taken from Ekman and Friesen's (1976) pictures of facial affect series. This stimulus set is especially important because it is well validated, on the basis of exact anatomical criteria, and has been extensively used in other studies. The set contains pictures of facial expressions associated with six basic emotions (happiness, sadness, anger, fear, disgust, and surprise) posed by a number of different models. Ekman and his colleagues have shown that each emotion is associated with distinct facial musculatures that are recognized by a number of cultures throughout the world (Ekman, 1972; Ekman et al., 1987). As far as we are aware, there have been no attempts to determine which of the Ekman and Friesen faces can be identified from their top or bottom sections. Consequently, we conducted a preliminary experiment (described in Experiment 1) that identified that anger, fear, and sadness were more readily recognized from the top half of the face, whereas happiness and disgust were more recognizable from the bottom half; surprise was found to be equally recognizable from both top and bottom sections.

On the basis of this information, composite expressions were prepared for Experiment 1 by aligning the top half of a recognizable-top expression (e.g., anger) with the bottom half of a recognizable-bottom expression (e.g., happiness) posed by the same model. As a comparison condition, we used noncomposite images; these were identical to the composites except that the top and bottom halves were misaligned horizontally. Following Young et al. (1987), we reasoned that support for a configural model of facial expression recognition could be found if participants were slower to identify the top (or bottom) half of an expression when it was shown as part of a composite (face-like) image, relative to when it was presented as part of a noncomposite (non–face-like) image. If, on the other hand, configural information is relatively unimportant for facial expression identification (part-based model), then no significant difference should be found between the composite and noncomposite conditions. This would occur because if facial expression recognition is based largely on the analysis of individual features, then aligning or misaligning the top and bottom face halves should have little affect on the participants' ability to identify the emotion.

In Experiment 2, we studied the effect of stimulus inversion on the composite phenomenon for facial expression. As we have already noted, Young et al. (1987) found that the composite effect for identity was disrupted by inverting the stimuli. It was clearly of interest, then, whether a composite effect for expression would be similarly affected.

Having demonstrated a composite effect for facial expression in Experiments 1 and 2, in Experiment 3 we addressed the criticism that the longer RTs for the composite condition could be attributed to the composite images appearing somehow more "attention grabbing" than the noncomposites. This might occur for a number of reasons; for example, the join between the top and bottom face halves can produce abrupt changes in texture and unnatural contours in the middle of nose and cheeks, causing the face to look slightly unusual in appearance. A method of addressing this issue presented itself during the preparation of the composites.

While making the stimuli, we noted Young et al.'s (1987) original effect that aligning the top and bottom halves of two peoples' faces generates a perceptually new face (see also Hole, 1994). However, we also noted a second interesting phenomenon. When the two face halves are taken from two identities posing the same facial expression (e.g., happiness), the resultant composite expression is also readily identifiable as happiness. This suggested an interesting prediction: that composite faces prepared from two identities posing the same expression (same-expression composites) should not show the composite effect for facial expression. Confirmation of this prediction would demonstrate that the composite effect observed in Experiments 1 and 2 cannot be attributed to some form of inherent quality of composite faces that causes them to produce slower response times (e.g., as a result of abrupt discontinuities in texture, etc.). On the other hand, a significant composite effect for the same-expression composites would question the idea that the composite paradigm taps configural processing. It was important, then, to address this issue.

Finally, Experiment 4 examined whether configural information for facial identity and facial expression recognition can be disrupted independently of one other. This was done by comparing participants' RTs to report the expression or identity shown in bottom half of composite faces containing the same or different expressions and same or different identities in the two facial halves.

Experiment 1

The first section of this experiment aimed to identify which of the expressions in the Ekman and Friesen (1976) series are identifiable from their top or bottom halves. This information was then used to create the composite and noncomposite images for Experiment 1.

Recognition Rates for Top and Bottom Sections of the Ekman and Friesen (1976) Faces

Method

Participants. Eight members of the MRC Cognition and Brain Sciences Unit subject panel (6 women, 2 men) participated in the experiment for payment. The participants were between the ages 21 and 40 years and had normal or corrected-to-normal vision.

Materials. The stimuli were prepared from gray-scale pictures from the Ekman and Friesen (1976) pictures of facial affect. Pictures of 10 people's faces (6 women, 4 men) were used, each posing one example of six facial expressions (happiness, sadness, anger, fear, disgust, and surprise). These 10 models were selected because a reliably recognized example of the six expressions was available for each. Each of these 60 pictures of facial expressions was divided into top and bottom segments. This was done by cutting each face along a horizontal line through the bridge of the nose. Examples of the stimuli are shown in Figure 2.

Design and procedure. One within-subjects factor, stimulus format (whole face, top segment, and bottom segment), was investigated. Participants saw the 60 faces (10 identities posing six

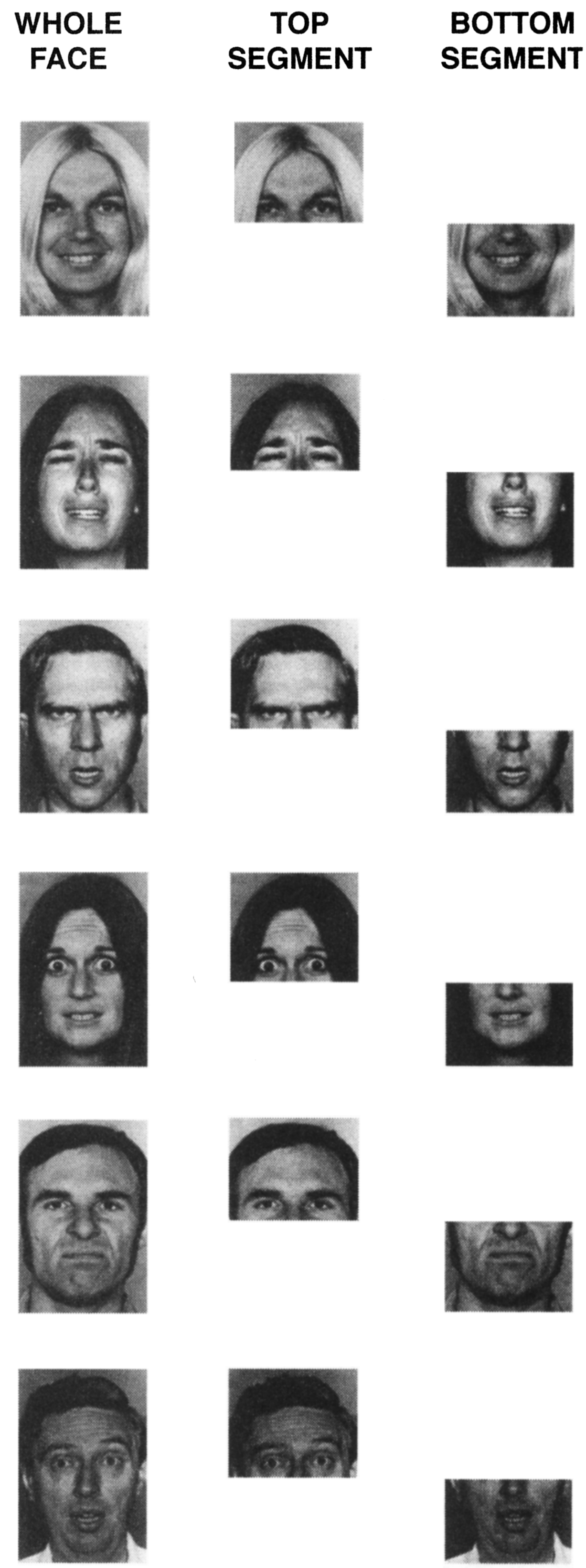

Figure 2. From Experiment 1, examples of the whole-face, top-segment, and bottom-segment stimuli. One example of six facial expressions (happiness, sadness, anger, fear, disgust, and surprise) posed by six models from the Ekman and Friesen (1976) series is shown in each of the three stimulus formats (whole face, top segment, and bottom segment). Images from *Pictures of Facial Affect,* by P. Ekman and W. V. Friesen, 1976. Copyright 1976 by P. Ekman and W. V. Friesen. Adapted with permission.

facial expressions) in each of these three stimulus formats. The 180 different stimuli were presented individually in random order on a 256 gray-scale computer screen. The top-segment images were presented in the location corresponding to the top half of the whole face and bottom-segment images in the corresponding bottom half location (see Figure 2). Each image subtended a horizontal visual angle of approximately 4.6°. The participant was asked to identify the emotion expressed in each image. Responses were made using a box with six labeled buttons (one for each emotion category, happiness, sadness, anger, fear, disgust, and surprise); the position of the emotion labels was counterbalanced across participants. The button box was interfaced with a Macintosh Power PC computer to record the participant's choice of emotion label and decision time.

On each trial, the image remained in view until the participant responded, and consecutive trials were separated by an interval of approximately 2.5 s. Participants were asked to respond quickly and accurately. After all 180 images had been presented, there was a short break, and then the same procedure was repeated in a second block.

To familiarize participants with the experimental format, the experiment began with 12 practice trials. These trials contained pictures of additional models from the Ekman and Friesen (1976) series posing the same six emotional expressions listed above in whole-face, top-segment, and bottom-segment formats. These practice faces were not seen in the main experimental trials.

Results

Participants' mean error proportions and mean correct RTs to identify the emotion displayed in the whole-face, top-segment, and bottom-segment images are listed in Table 1 by emotion category. Standard errors are shown in brackets.

Table 1
Data from Experiment 1

	Face format					
	Whole		Top		Bottom	
Emotion	*M*	*SE*	*M*	*SE*	*M*	*SE*
	Error proportions					
Anger	.22	.08	.28	.06	.49	.09
Fear	.25	.07	.33	.08	.56	.09
Sadness	.09	.03	.19	.05	.34	.08
Happiness	.01	.01	.20	.09	.01	.01
Disgust	.14	.05	.62	.10	.14	.04
Surprise	.21	.07	.21	.06	.33	.07
	Reaction times (in milliseconds)					
Anger	1,910	130	1,963	118	2,380	271
Fear	2,041	132	2,043	125	2,210	237
Sadness	1,742	202	1,803	142	2,400	332
Happiness	1,178	103	1,394	74	1,119	113
Disgust	1,738	258	2,320	206	1,413	124
Surprise	1,748	177	1,847	245	1,949	156

Note. Participants' error proportions and mean correct reaction times to identify the emotion displayed in examples of six facial expressions (happiness, sadness, anger, fear, disgust, and surprise). The faces were presented in three formats: whole face, top segment, and bottom segment.

Error rates. Our principal form of analysis involved error rates in identifying the emotions, because we were interested in determining which of the expressions could be identified accurately from their top halves (recognizable-top expressions), and which could be identified accurately from their bottom halves (recognizable-bottom expressions). Error proportions were arcsin transformed and submitted to two analysis of variance (ANOVAs), one by participants (F_1) the other by items (F_2). Two factors were investigated: face format (whole face, top segment, and bottom segment; repeated measure) and emotion (happiness, sadness, anger, fear, disgust, and surprise; repeated measure). Both analyses showed a significant effect of face format, $F_1(2, 14) = 30.44$, $p < .0001$, and $F_2(2, 18) = 11.94$, $p < .0005$. Post hoc t tests ($p < .05$) of the two analyses showed the same pattern; overall, the emotions were more accurately identified from the whole-face images than from the top or bottom segments, which did not reliably differ. This main effect was qualified by a significant interaction between emotion and face format, $F_1(10, 70) = 11.46$, $p < .0001$, and $F_2(10, 90) = 9.89$, $p < .0001$. Simple effects analyses by participants (F_1) and by items (F_2) showed significant effects of face format for all emotions except surprise. The F values of these simple effects analyses are listed by emotion category in the following section, and where appropriate, a summary of post hoc t tests ($p < .05$) of the simple effect is shown in brackets (note: in each case the post hoc effects were identical for the analyses by participants and by items): anger, $F_1(2, 14) = 10.07$, $p < .005$, and $F_2(2, 18) = 6.43$, $p < .01$ ([whole = top] < bottom); fear, $F_1(2, 14) = 27.10$, $p < .001$, and $F_2(2, 18) = 6.25$, $p < .01$ ([whole = top] < bottom); sadness, $F_1(2, 14) = 7.13$, $p < .01$, and $F_2(2, 18) = 19.28$, $p < .001$ ([whole = top] < bottom); happiness, $F_1(2, 14) = 5.77$, $p < .02$, and $F_2(2, 18) = 10.64$, $p < .001$ ([whole = bottom] < top); disgust, $F_1(2, 14) = 29.77$, $p < .001$, and $F_2(2, 18) = 33.12$, $p < .001$ ([whole = bottom] < top); and surprise, $F_1(2, 14) = 3.41$, $p > .05$, and $F_2(2, 18) = 1.97$, $p > .2$. Finally, both analyses also showed significant effects of emotion, $F_1(5, 35) = 7.94$, $p < .0001$, and $F_2(5, 45) = 7.46$, $p < .0001$. Post hoc t tests ($p < .05$) showed that, overall, happiness was more accurately recognized than the other emotions.

In summary, the results of the error rates analysis show that anger, fear, and sadness were more recognizable from the top half of the face (recognizable-top expressions), whereas happiness and disgust were more recognizable from the bottom half of the face (recognizable-bottom expressions). Surprise was equally recognizable from its top and bottom sections.

These results essentially replicate those of Bassili (1979), who examined the same six facial expressions, although his images were not taken from the Ekman and Friesen (1976) series, and they were animated. The only difference between Bassili's findings and our own is that Bassili's sadness expressions were equally recognizable from their whole, top, and bottom segments, whereas we found that the bottom segments of the Ekman and Friesen sadness expressions were less accurately recognized than their whole or top segments, which did not reliably differ.

RTs. Two subsidiary analyses (one by participants, F_1, the other by items, F_2) were carried out on the RT data to check that the more accurate responses were not accompanied by slower RTs. Again, the factors investigated were face format (whole face, top segment, and bottom segment; repeated measure) and emotion (happiness, sadness, anger, fear, disgust, and surprise; repeated measure). Neither analysis showed a significant effect of face format, but both showed a significant interaction between emotion and face format, $F_1(10, 70) = 3.78$, $p < .0005$, and $F_2(10, 90) = 3.12$, $p < .005$. Simple effects analyses by participants (F_1) and by items (F_2) showed a significant effect of face format for happiness, $F_1(2, 14) = 7.15$, $p < .01$, and $F_2(2, 18) = 6.46$, $p < .01$; and disgust, $F_1(2, 14) = 8.06$, $p < .005$, and $F_2(2, 18) = 5.76$, $p < .05$, only. Post hoc t tests ($p < .05$) showed that participants were significantly slower to identify the happiness and disgust emotions from the top segment of the face; RTs to identify these emotions from the bottom-segment and whole-face images did not reliably differ. Thus, there was no evidence of participants trading accuracy for speed.

In summary, this preliminary study identified that anger, fear, and sadness are readily identified from the top section of the face (recognizable-top expressions), whereas happiness and disgust are readily identifiable from the bottom half of the face (recognizable-bottom expressions). Because surprise could be recognized from either part of the face, we used it as a recognizable-bottom expression to even up the number of expressions in each condition of our design.

In the next section of the experiment, we created composite facial expressions composed of the top halves of the recognizable-top expressions and bottom halves of the recognizable-bottom expressions (e.g., top = anger, bottom = happiness). These images allowed us to test whether a similar phenomenon to the composite effect for facial identity (Young et al., 1987) could be found with facial expressions. Following Young et al.'s reasoning, if configural information is important for facial expression recognition, then participants should be slower to identify the top or bottom half of a facial expression when it is presented as part of a composite image than when it is shown as part of a noncomposite (misaligned) image.

Identifying the Top and Bottom Sections of Composite and Noncomposite Expression Images

Method

Participants. Twelve people (9 women, 3 men) aged between 21 and 40 years and from the same population as the previous section participated in the experiment. None had taken part in the previous section.

Materials. The stimuli were prepared from pictures of four female models from the Ekman and Friesen (1976) series (C, NR, PF, and SW), each posing one example of the expressions happiness, sadness, anger, fear, disgust, and surprise; these pictures were selected from the stimuli used in the previous experiment. Composite and noncomposite stimuli comparable with the facial identity composites used by Young et al. (1987) were then prepared from these facial expressions. Their preparation is described below.

Composites. Composite facial expressions were prepared by aligning the top segment of a recognizable-top expression (e.g., anger) with the bottom segment of a recognizable-bottom expression (e.g., happiness) posed by the same model. For each of the four models, all nine possible combinations of these recognizable-top and recognizable-bottom segments were prepared; these combinations were as follows: anger–happiness, anger–disgust, anger–surprise, fear–happiness, fear–disgust, fear–surprise, sadness–happiness, sadness–disgust, and sadness–surprise (the first emotion of each pair indicates the top half of the composite). This gave a total of 36 composite faces.

Noncomposites. The noncomposite facial expressions were essentially identical to the composites except that the top and bottom segments were misaligned horizontally. This was done by aligning the middle of the nose in the top segment with the edge of the face in the bottom segment. For half of the images, the top segment was shifted to the left of the bottom segment, and for the other half, this positioning was reversed (see Figure 3). Note that when the noncomposites were presented in the center of the computer screen, neither the bottom or top half of the image was centralized in the screen. To allow for this fact, half of the composite stimuli were presented in the same position as the left section of the noncomposites and half in the same location as their right section (see Figure 3); positioning was counterbalanced across stimuli. This method of presentation follows the basic procedure used by Young et al. (1987).

Examples of composite and noncomposite facial expressions prepared from pictures of one of the models used in Experiment 2 are shown in Figure 3.

Design and procedure. Two within-subjects factors were investigated: stimulus type (composite and noncomposite) and task instructions ("identify the top-half expression" and "identify the bottom-half expression"). The experiment began with a block in which each of the 24 whole (prototype) facial expressions (four models, each posing six facial expressions) were presented individually in random order. The participant's task was to identify the emotion displayed in each face by pressing one of six buttons marked with the emotion labels *happiness, sadness, anger, fear, disgust,* and *surprise;* the position of these labels was counterbalanced across participants. Each face was preceded by a fixation cross for 500 ms followed by a blank interval of the same duration. The face remained in view until the participant responded, with their response initiating the next trial after an interval of approximately 2.5 s. All images were displayed on a 22-in. gray-scale computer screen using a Macintosh Power PC. The purpose of this block of trials was to familiarize the labeling task, to ensure that accuracies in the main experimental trials were sufficiently high to allow meaningful measurement of RTs.

Participants then completed two blocks of experimental trials. In one block, they were asked to identify the expression displayed in the top segment of the composite and noncomposite images (top-segment block) and in a second block the expression shown in the bottom segment of these same images (bottom-segment block); half of the participants did the bottom-segment block trials first. The general design of these two blocks was the same, so we only give a detailed description of the bottom-segment block.

The bottom-segment block began with a single presentation of the bottom segment of each of the three facial expressions (happiness, disgust, and surprise) posed by the four models (C, NR, PF, and SW); the presentation times were as for the whole-face

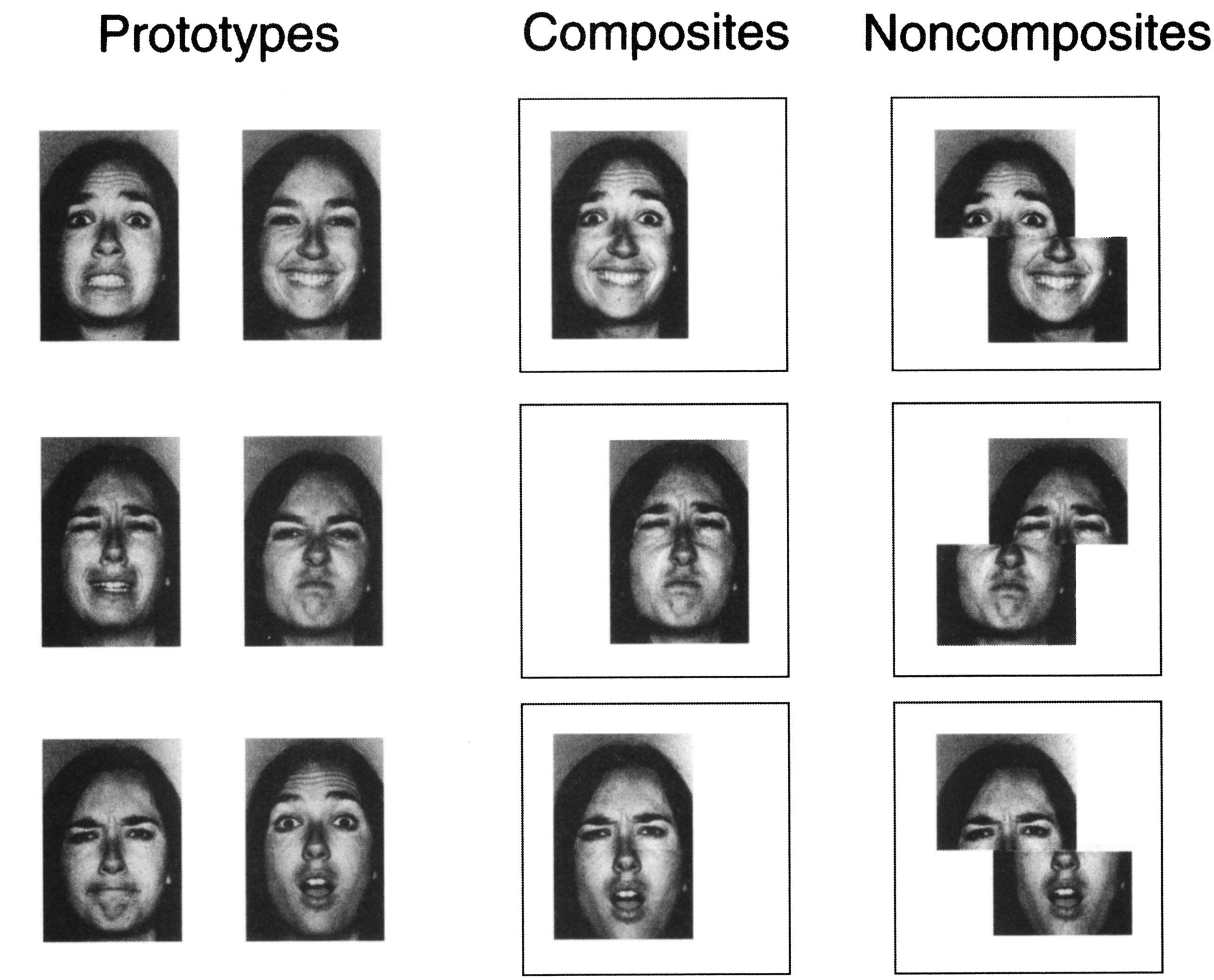

Figure 3. Examples of stimuli used in Experiment 1. The top and bottom segments of recognizable-top and recognizable-bottom prototype expressions (left), respectively, were combined to create composite (middle) and noncomposite (right) stimuli. The two face sections of each composite and noncomposite image were from pictures of the same model (Model C in the example shown). Images from *Pictures of Facial Affect,* by P. Ekman and W. V. Friesen, 1976. Copyright 1976 by P. Ekman and W. V. Friesen. Adapted with permission.

presentations described above. Participants were asked to make an identification decision by pressing one of three buttons labeled *happiness, disgust,* and *surprise.* Following this, the experiment proper began. This included one presentation of each of the 36 composite and 36 noncomposite stimuli described above. The images were presented in random order, and the participant was asked to identify the expression displayed in the bottom segment of the images as quickly and accurately as possible by pressing one of the three labeled keys. Again, the presentation times were identical to the whole-face presentations described earlier. To familiarize the participants with the composite and noncomposite images, the experiment was preceded by 10 practice trials selected at random from the 72 experimental trials. The composite images subtended a horizontal visual angle of approximately 4.6°, and for the noncomposites, a horizontal visual angle was approximately 5.7°; the vertical visual angle for both was approximately 6.3°.

For the top-segment block, the design was virtually identical. However, this time the block began with one presentation of the top segment of the facial expressions anger, fear, and sadness posed by the same four models. The participants were then presented with the same composite and noncomposite images seen in the bottom-segment block, but this time they were asked to identify the expression displayed in the top segment of the face. In both sections of the top-segment block, participants made their response by pressing one of three keys labeled *anger, fear,* and *sadness.*

Results

Participants' mean correct RTs (with standard error bars) to identify the top and bottom halves of the composite and noncomposite facial expressions are shown in the left graph

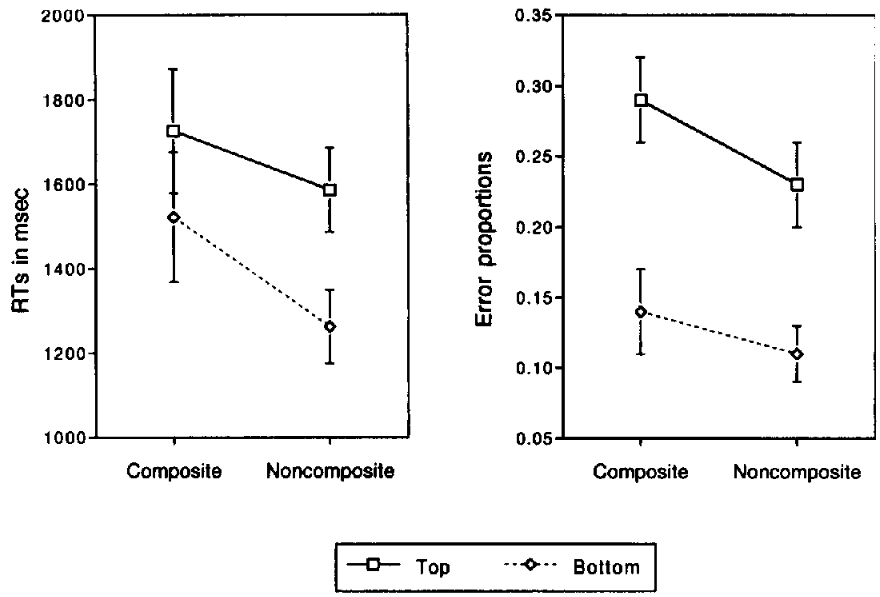

Figure 4. Data from Experiment 1. The left graph shows participants' mean correct reaction times (RTs; with standard error bars) to identify the expression displayed in top and bottom halves of the composite and noncomposite stimuli. The right graph shows participants' mean error proportions (with standard error bars) from the same experiment.

of Figure 4. The right graph shows participants' mean error proportions (with standard error bars) for the same experiment.

RTs. Our principal form of analysis involved RTs for correct responses. These were submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and task instructions ("identify top-half expression" and "identify bottom-half expression"; repeated measure). There was a significant effect of stimulus type, $F(1, 11) = 6.35, p < .05$, indicating that participants found it harder to identify the top and bottom segments of the images in the composite condition. There was also a significant effect of task instructions, $F(1, 11) = 17.26, p < .005$, demonstrating that, overall, participants were faster to recognize the expression shown in the bottom half of the images. There was no significant interaction between these two factors.

Error rates. A subsidiary analysis examined error rates to ensure that the slower RTs to the composite images were not accompanied by increased accuracy. Error proportions were arcsin transformed and submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and task instructions ("identify top-half expression" and "identify bottom-half expression"; repeated measure). This showed a significant main effect of task instructions, $F(1, 11) = 23.02, p < .001$, reflecting that, overall, participants were significantly more accurate at identifying the expressions shown in the bottom half of the images. There was also a marginally significant effect of stimulus type, $F(1, 11) = 4.65, .1 > p > .05$, but no significant interaction between these two factors. Thus, there was no evidence of participants trading accuracy for speed.

Discussion

The results of Experiment 1 support the configural model of facial expression recognition over the part-based model. Participants were significantly slower, and marginally less accurate, at identifying the expression shown in half of the composite images than the noncomposite images. Moreover, the effect was equally strong when they were asked to identify the top half of the images as when they were asked to identify their bottom half. This composite effect for facial expression is all the more striking when we consider that the participants were only asked to discriminate among three different facial expressions in each of the top-segment and bottom-segment blocks. Hence, although strategies were readily available to the participants (e.g., if the mouth is open wide, the bottom-half expression must be surprise, or if the eyes are wide open, the top-half expression must be fear), they did not, or were not able to, make full use of them. In this sense, these findings essentially parallel those found for facial identity (Carey & Diamond, 1994; Young et al., 1987), and a similar explanation can be invoked. Following Young et al.'s reasoning, we suggest that facial expressions are processed in terms of their configural make-up; that is, the shape and position of the mouth in anger may be coded relative to the shape and position of other features in the expression (e.g., furrowed brow, close-set eyebrows, etc.). Hence, when the top and bottom segments of different facial expressions are aligned, they fuse to form a perceptually new facial expression configuration that interferes with the processing of the constituent parts of the top and bottom sections. This effect can be seen in the examples shown in Figure 3. The top row shows the top half of a fear expression combined with the bottom half of a happiness expression. The result is a wild expression that could not really be accurately described as happiness or fear.

It is important to emphasize, however, that a composite effect for facial affect does not mean that the individual features of facial expressions are not also encoded for identification. It simply implies that the configural relationship of the features plays a significant role in the encoding of facial expression.

Recall that Ellison and Massaro (1997) found that their data could be reliably modeled by the FLMP if one assumed that the information in the upper and lower sections of the face were evaluated independently and then integrated to produce an overall degree of support for a particular emotion category (e.g., happiness). Our own data do not concur with this finding. In Experiment 1, the same face halves were present in the composite and noncomposite conditions. Hence, if the face halves were being processed independently of one another, we would predict that the RTs for the two conditions should not significantly differ. However, this was not found: the participants responses were significantly slower for the composite condition. In other words, aligning the face halves to produce a facial image has a significant effect on the speed with which the participants can perform the task. For the present, then, we note there is a disagreement between Ellison and Massaro's results and our own, and in the General Discussion section we address possible explanations.

It is also worth emphasising that our results cannot simply be attributed to a Stroop (1935) interference effect between the different conceptual (or semantic) information conveyed by the top and bottom face halves. This is because the same halves are present in both composite and noncomposite conditions. Hence, although a Stroop effect between emotion concepts may operate in both experimental conditions, it can not be the source of the increased RTs found for the composite condition.

Experiment 2

As we discussed in the introduction, Young et al. (1987) found that the composite effect for facial identity was lost when the stimuli were inverted (see also Carey & Diamond, 1994, and Hole, 1994). This is consistent with Carey and Diamond's (1977) suggestion that configural information for identity is more difficult to process in inverted than upright faces. Therefore, in Experiment 2, we investigated the effect of stimulus inversion on the composite effect for facial expression. We reasoned that if configural processing constitutes the basis of the effect we have observed, then the composite effect for expression should be significantly disrupted when the stimuli are inverted.

In Experiment 1, the participants were asked to identify both top and bottom sections of the composite and noncomposite images, and no difference in the pattern of findings was noted across "identify-top" and "identify-bottom" conditions—both showed an equivalent composite effect. For Experiment 2, therefore, we arbitrarily selected the bottom section of the images for the participants to identify in both upright and inverted conditions.

Method

Participants. Twelve people (6 women, 6 men) aged between 19 and 45 years and from the same population as Experiments 1 and 2 participated in the experiment. All had normal or corrected-to-normal vision, and none had taken part in Experiments 1 and 2.

Materials. The stimuli were identical to those used in Experiment 1.

Design and procedure. In the previous experiment, participants identified the expression shown in the top half of the composite and noncomposite images in one block and the expression shown in the bottom half in a second block. In Experiment 2, participants were only asked to identify the expression shown in the bottom half of these same stimuli, but under two conditions: (a) when the stimuli were presented upright and (b) when the same stimuli were inverted. Hence, in the inverted condition, the bottom half of the face was effectively the top half of the image.

The beginning of the experiment was identical to Experiment 1; participants were presented with the 24 original whole-face images and asked to categorize each with one of six emotion labels (happiness, sadness, anger, fear, disgust, and surprise). Following this, half of the participants were assigned first to the upright condition and half to the inverted condition. The upright condition block was identical to the bottom-segment block described in Experiment 1. Hence, participants were first presented with the bottom segments of the expressions happiness, disgust, and surprise posed by four models and asked to categorize each image with one of three emotion labels (*happiness, disgust,* and *surprise*). In the experiment proper the same composite and noncomposite stimuli used in Experiment 1 were presented individually in random order. The participants' task was to categorize the bottom segment of each image with one of the same three emotion labels as quickly and accurately as possible.

The inverted condition block was essentially identical to the upright condition block, except that all of the stimuli were inverted. In all other respects, the design and procedure of Experiment 2 were the same as for Experiment 1.

Results

Participants' mean correct RTs (with standard error bars) to identify the bottom half of the composite and noncomposite facial expressions in upright and inverted formats are shown in the left graph of Figure 5. The right graph shows participants' mean error proportions (with standard error bars) for the same experiment.

Reaction times. Our principal form of analysis involved RTs for correct responses. These were submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and stimulus orientation (upright and inverted; repeated measure). There was a significant effect of stimulus type, $F(1, 11) = 9.74$, $p < .01$, indicating that participants found it harder to identify the expression shown in the bottom half of the composite images. This was qualified by a significant interaction between stimulus type and stimulus orientation, $F(1, 11) =$

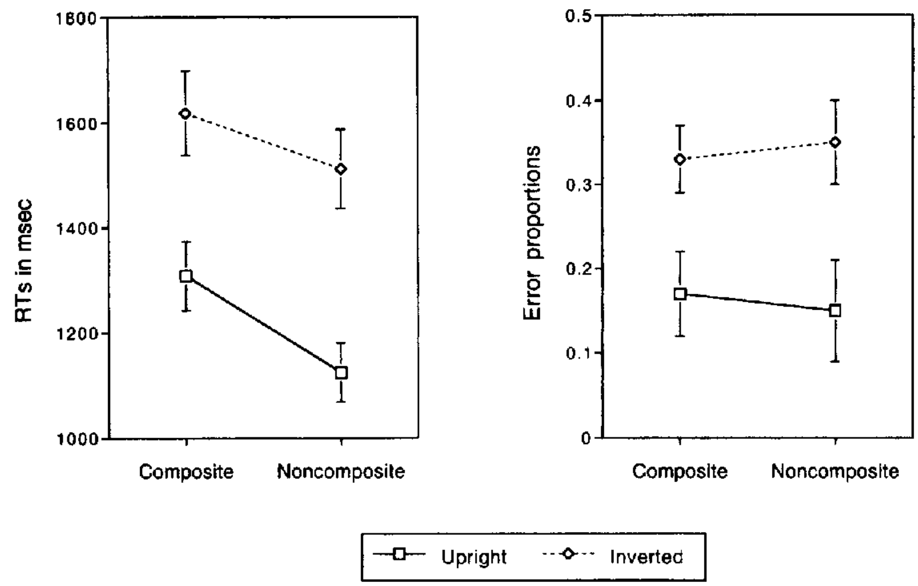

Figure 5. Data from Experiment 2. The left graph shows participants' mean correct reaction times (RTs; with standard error bars) to identify the expression displayed in the bottom half of the composite and noncomposite images presented in upright and inverted formats. The right graph shows participants' mean error proportions (with standard error bars) from the same experiment.

6.62, $p < .05$. Simple effects analyses of the interaction effect showed a significant effect of stimulus type (composite and noncomposite) for the upright condition, $F(1, 11) = 15.14$, $p < .005$, and a borderline, nonsignificant effect for the inverted condition, $F(1, 11) = 4.02$, $.1 > p > .05$. Finally, there was also a significant effect of stimulus orientation, $F(1, 11) = 14.06$, $p < .005$, demonstrating that, overall, participants found the task easier when the stimuli were upright.

Error rates. A subsidiary analysis examined participants' error rates to check that the slower responses in the composite condition were not also accompanied by more accurate performance. Error proportions were arcsin transformed and submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and stimulus orientation (upright and inverted; repeated measure). This showed a significant main effect of stimulus orientation, $F(1, 11) = 11.80$, $p < .01$, reflecting that, overall, participants were significantly more accurate at identifying the expression shown in the upright images. There were no other significant effects (Fs < 1.10). Thus, there was no statistical evidence of participants trading accuracy for speed.

Discussion

Experiment 2 demonstrates three findings. First, the results replicated the findings of Experiment 1. In the upright condition, participants were significantly slower (but no more accurate) to identify the expression shown in the bottom half of the composite images relative to their performance with the noncomposite images. Second, inverting the images significantly disrupted the composite effect for facial expressions; the composite effect was statistically reliable for the upright condition only. This second finding is similar to Young et al.'s (1987) observation that the composite effect for facial identity is lost when the stimuli are inverted (see also Carey & Diamond, 1994; Hole, 1994). Finally, Experiment 2 also showed a significant main effect of stimulus inversion. This indicates that, overall, participants were significantly slower to identify the expression shown in the bottom half of the composite and noncomposite stimuli when they were inverted. This is consistent with McKelvie's (1995) finding that facial expressions are more difficult to recognize in inverted faces (see also Valentine & Bruce, 1988).

It is important to note that the negative effect of inversion on facial identity recognition is usually attributed to the idea that configural features are more difficult to process in inverted faces or, less specifically, that holistic processing of faces is made more difficult by stimulus inversion. The fact that the composite effect for facial expression is also disrupted by inversion converges on the idea that configural features may also be used to encode facial expressions. Hence, the results of Experiment 2 provide further support for the configural model rather than the part-based model of facial affect recognition.

Given that our interpretation of the results of Experiments 1 and 2 has substantial implications for the understanding of facial expression perception, it was important to consider whether there were any alternative interpretations of the composite effect we had found. We considered that one possibility was that the composite stimuli were simply more attention-grabbing than the noncomposites, possibly because facial composites prepared from the top and bottom halves of two different pictures inevitably look like unusual (or distorted) faces, causing participants to look longer at them before deciding on a response. Experiment 3 addressed this alternative explanation.

Experiment 3

Although it seems likely that the composite effect found in Experiments 1 and 2 is attributable to a disruption of configural processing, an alternative explanation may exist. As we have said, it is possible that the participants may have been distracted by the slightly distorted and unusual appearance of the composite stimuli and, hence, slower to make their response. Clearly, it was important to address this alternative explanation, and one means of testing it became evident while we were preparing the stimuli for Experiments 1 and 2.

When creating stimuli, we noted that if the two face halves are taken from different identities posing the *same* facial expression (e.g., happiness), the resultant composite face also looks happy. This result suggested a prediction: If the composite effect we had observed was due to a disruption of configural information for facial expression, then composite faces prepared from two different identities posing the same expression (same-expression composites) should not show the effect. This is because the top and bottom segments of these images contain configural information relating to the same facial affect, meaning that there is no conflict between the configural information for expression in the two halves, even though the identities are different. Alternatively, if the effect we had observed was due to the composite stimuli being more attention-grabbing than the noncomposites (as a result of discontinuities in texture across the two face halves, etc.), then a significant composite effect should be observed for the same-expression composites. This was tested in Experiment 3.

As a comparison condition, we also included composite images prepared from different identities posing different facial expressions (different-expression composites). We predicted that these images should produce the same composite effect found in Experiments 1 and 2, because for the different-expression composites, there is a conflict of configural information for facial expression across the two halves of the image. Hence, our suggestion that the composite effect reflected configural processing of the images would hold true if Experiment 3 showed a significant interaction effect between stimulus type (composite and noncomposite) and top–bottom expression congruency (same expression and different expression).

Method

Participants. Twelve participants (7 women, 5 men) aged between 18 and 40 years and from the same population as Experiments 1 and 2 took part in the experiment. All had normal or corrected-to-normal vision, and none had participated in the previous experiments.

Materials. The stimuli were prepared from pictures of the same four models (C, NR, PF, and SW) used in Experiments 1 and 2 posing the facial expressions happiness, disgust, and surprise. The top and bottom halves of these faces were combined to produce all possible composite expressions in which the two halves were taken from different models' faces. For 36 of these images, the top and bottom halves showed the same expression (same-expression composite; e.g., top = happiness Model C, bottom = happiness Model NR), and for the remaining 72, the two halves showed different expressions (different-expression composite; e.g., top = happiness Model C, bottom = disgust Model NR).

Noncomposite versions of the same stimuli were produced using the method described in Experiment 1. Recall that there are two possible versions of noncomposite stimuli: (a) top half shifted to the right of the bottom half and (b) top half shifted to the left of the bottom half. Given that there were twice as many different-expression composites as same-expression composites, both versions of noncomposite were produced for each of the same-expression images, whereas for the different-expression images, the two versions were counterbalanced across stimuli. Examples of composite and noncomposite images prepared from two of the four models used in Experiment 3 are shown in Figure 6.

Design and procedure. Two within-subjects factors were investigated: stimulus type (composite and noncomposite) and top–bottom expression congruency (same expression and different expression).

All stages of the experiment used the presentation format described in Experiment 2 (i.e., 500-ms fixation, 500-ms blank interstimulus interval followed by the stimulus, which remained in view until the participant responded). The experiment began with a session in which the original 12 whole faces (four models, each posing three facial expressions) used to prepare the composites were presented individually in random order. Each face was shown three times, and the participant identified the emotion displayed by pressing one of three keys marked with the labels *happiness, disgust,* and *surprise;* label positions were counterbalanced across participants. Next, half of the participants were presented with the top segments of these same faces and half with the bottom segments. Again, each image was presented three times, and the participants' task was to identify the facial expression as one of happiness, disgust, or surprise. After this, the participants that had seen the top sections were presented with the bottom sections of the same facial expressions and vice versa. Their task was the same, namely to identify the emotion.

In the experiment proper the participants were presented with equal numbers (36) of same-expression composites, same-expression noncomposites, different-expression composites, and different-expression noncomposites in random order. The stimuli were counterbalanced across two stimulus sets to accommodate the different numbers of same-expression and different-expression images; half of the participants were assigned to one stimulus set and half to the other. Participants were instructed to identify the expression displayed on the bottom half of each image by pressing the appropriate response key (happiness, surprise, or disgust) as quickly and accurately as possible. To familiarize the participants with the composite and noncomposite images, the experiment proper was preceded by 10 practice trials selected at random from the experimental trials. The composite images subtended a horizontal visual angle of approximately 4.6°, and the noncomposite images subtended a horizontal visual angle of approximately 5.7°; the vertical visual angle for both was approximately 6.3°.

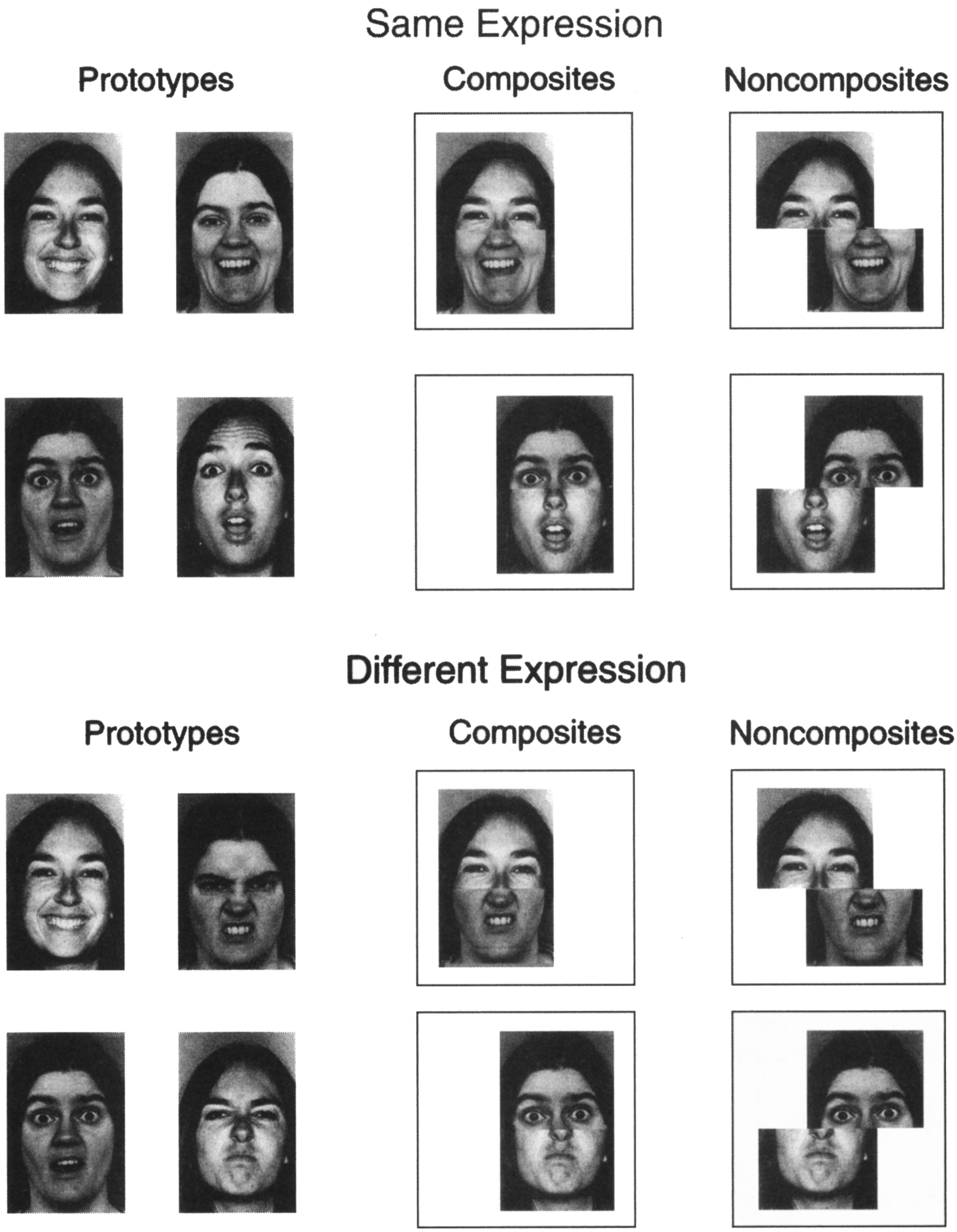

Figure 6. Examples of the stimuli used in Experiment 3. Composite (middle) and noncomposite (right) facial expression stimuli were prepared from the top and bottom halves of two models' faces (left) posing the same expression (same-expression images; top row) or different expressions (different-expression images; bottom row). Images from *Pictures of Facial Affect,* by P. Ekman and W. V. Friesen, 1976. Copyright 1976 by P. Ekman and W. V. Friesen. Adapted with permission.

Results

Participants' mean correct RTs (with standard error bars) to identify the bottom half of the composite and noncomposite facial expressions in same-expression and different-expression conditions are shown in the left graph of Figure 7. The right graph shows participants' mean error proportions (with standard error bars) for the same experiment.

RTs. Our principal form of analysis involved RTs for correct responses. These were submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and top–bottom expression congruency (same expression and different expression; repeated measure). There was a significant effect of stimulus type, $F(1, 11) = 10.10, p < .01$, indicating that, overall, participants were slower to identify the expression shown in the bottom half of the composite images. This was qualified by a significant interaction between stimulus type and top–bottom expression congruency, $F(1, 11) = 12.94, p < .005$. Simple effects analyses showed a significant effect of stimulus type for the different-expression images, $F(1, 11) = 24.00, p < .0001$, but not for the same-expression images ($F < 1.00$). There was also a significant effect of expression congruency, $F(1, 11) = 9.67, p < .01$, demonstrating that, overall, participants were significantly slower to identify the expression in the different-expression images; post hoc t tests ($p < .05$) showed that this held for the composite images (same expression < different expression) but not for the noncomposite images.

Error rates. A subsidiary analysis examined participants' error rates to check that the slower responses were not accompanied by more accurate performance. Error proportions were arcsin transformed and submitted to a two-factor ANOVA investigating stimulus type (composite and noncomposite; repeated measure) and top–bottom expression congruency (same expression and different expression; repeated measure). There was a marginal main effect of stimulus type, $F(1, 11) = 3.86, .1 > p > .05$, reflecting an overall trend toward more errors with the composite stimuli. This was qualified by a significant interaction between stimulus type and top–bottom expression congruency, $F(1, 11) = 5.61, p < .05$. Simple effects analyses showed a significant effect of stimulus type for the different-expression images, $F(1, 11) = 15.88, p < .005$, but not for the same-expression images ($F < 1.00$). There was also a significant main effect of top–bottom expression congruency, $F(1, 11) = 7.23, p < .05$, indicating that, overall, participants made significantly more errors with the different-expression images. Post hoc t tests ($p < .05$) indicated that this effect held for the composite and noncomposite images (same expression < different expression). The results of the error rates analysis, then, show no evidence of participants trading speed for accuracy.

Discussion

The results of Experiment 3 can be summarized as follows. First, a composite effect for facial expressions was found when the images compose the top and bottom segments of different people's faces posing different expressions; this replicates and extends the findings of Experiments 1 and 2. Second, no composite effect was observed when the stimuli were prepared from the top and bottom

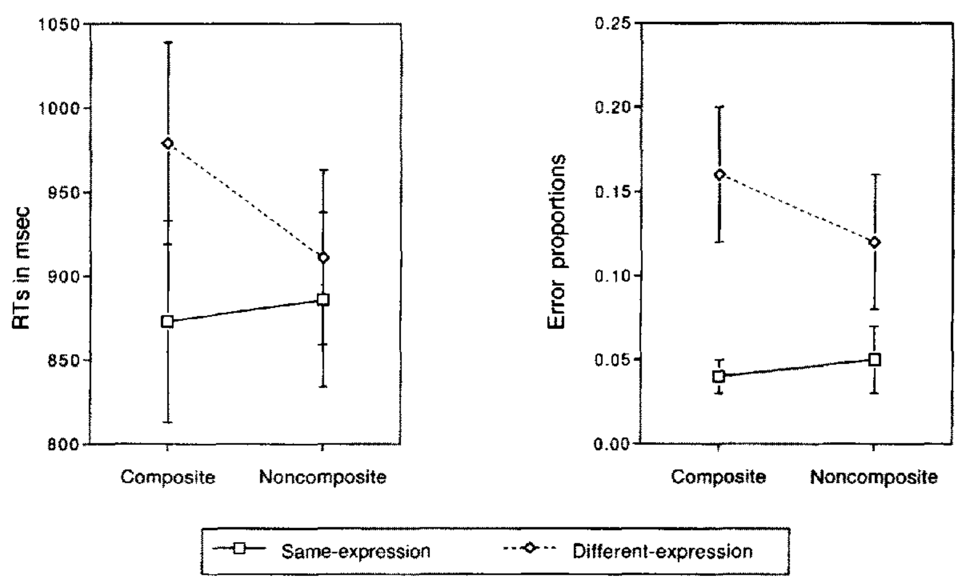

Figure 7. Data from Experiment 3. The left graph shows participants' mean correct reaction times (RTs; with standard error bars) to identify the expression displayed in bottom half of composite and noncomposite images containing the same facial expression (same-expression images) or different facial expressions (different-expression images). The right graph shows participants' mean error proportions (with standard error bars) from the same experiment.

segments of different people's faces posing the same expression.

The significant factor that differentiated these two types of stimuli was that for one stimulus, the expressions shown in the top and bottom segments were different (different-expression composite), whereas for the other stimulus they were the same (same-expression composites). The results of Experiment 3, then, confirm that the composite effect found in Experiments 1 and 2 cannot be attributed to the idea that composite faces are more attention-grabbing than noncomposites. Instead, these findings are consistent with the suggestion that for the different-expression composites, there is a conflict between the configural information in the two face halves, whereas for the same-expression composites there is no such conflict. Once again, then, our data are consistent with the configural model of facial affect processing.

Note that in Experiments 1 and 2, the top and bottom halves of each composite expression were taken from pictures of the same model; hence, it was relatively easy to align the nose, hairline, and so forth, in these images and to avoid abrupt discontinuities of stimulus texture. This was considerably more difficult in Experiment 3 because the two halves belonged to the faces of different models. Experiment 3, then, demonstrated that the composite effect for facial expressions is robust, because the effect was found even when the composite images did not look (on close inspection) like fully credible faces.

Finally, it is worth noting one further point. In Experiment 3, our composite stimuli were prepared from the top and bottom halves of recognizable-bottom facial expressions. However, we still observed a large composite effect in the different-expression condition. This would suggest that the face half that the participant was not instructed to attend to (i.e., the top half of the face in Experiment 3), did not need to display a highly recognizable expression for the effect to occur; the average recognition rates from Experiment 1 for the top halves of these expressions across the four models used were as follows: happiness, 68.5%; disgust, 43.75%; and surprise, 84.75% (chance = 16.67%). Contrast these rates with the recognition rates for the bottom halves of the same expressions (happiness, 100%; disgust, 87.5%; and surprise, 85.75%). Hence, the critical factor for a composite effect to be observed may be that the configural information in the unattended half is inconsistent with that in the attended half, rather than that the unattended half itself contains another readily identifiable facial expression.

Experiment 4

Figure 6 illustrates a point demonstrated in Experiment 3 that the top and bottom halves of two faces of different people posing the same expression (e.g., happiness) can be combined to generate a perceptually new facial identity without disrupting facial expression (i.e., the face still looks happy). In other words, the composite face is a poor match for either of the two original models' faces but a good match for the expression posed by both models. This observation implies that the configural information used to encode facial identity may be different from that used to encode facial expression. We reasoned that support for this hypothesis could be found by showing that the configural processing of facial identity and facial expression can be selectively disrupted.

To demonstrate this, we used three types of composite stimuli prepared from (a) pictures of the same person posing different facial expressions (same-identity–different-expression composites), (b) pictures of different people posing the same facial expression (different-identity–same-expression composites), and (c) pictures of different identities posing different facial expressions (different-identity–different-expression composites). We predicted that if participants could selectively attend to the configural information that specifies either facial identity or facial expression, then we should find different patterns of performance with different task instructions. Hence, when the instructions are to indicate the identity shown in the bottom half of a composite face, participants' responses should be fastest when the top and bottom segments contain the same person's face (same-identity–different-expression composites). However, when the instructions are to identify the expression shown in the bottom half, participants responses should be fastest when the top and bottom segments contain the same facial expression (different-identity–same-expression composites).

The third type of composites (different-identity–different-expression composites) was used for the following reason. If different configural information is used to specify identity and expression, then although there should be a significant cost to RTs when the attribute (identity or expression) that the participants are asked to attend to is incongruent across the two face halves, there should be no additional cost when the two halves are incongruent with respect to both facial attributes. So, for example, participants' RTs to identify the expression shown in the bottom half of the same-identity–different-expression and different-identity–different-expression composites should not differ. Similarly, there should be no reliable difference between their RTs to indicate the identity shown in the bottom half of the different-identity–same-expression and different-identity–different-expression composites.

To test these hypotheses, it was necessary to use either pictures of already familiar faces posing different expressions or faces from the Ekman and Friesen (1976) series, which were made familiar to the participants at the beginning of the experiment. The latter method of making unfamiliar faces familiar in the course of the experiment has been successfully used by Young et al. (1987) and Carey and Diamond (1994) in their investigations of the composite effect for facial identity, and, on balance, we selected it for two reasons. First, this method facilitates comparison with other experiments in this article and, second, full-face pictures of personally familiar people or of celebrities posing different facial expressions are difficult to obtain.

Note that we did not use noncomposite images in Experiment 4 for the following reason. Experiments 1, 2, and 3 were all consistent with the suggestion that facial expression composites are encoded configurally. This effect was observed despite the participants being instructed to attend to only one section of the face. Hence, although participants would have improved their performance for the composite condition by only processing the information in the attended half, they were apparently not able to use this strategy. These results strongly suggest that under these circumstances, the configural encoding of facial expression is an automatic process that is beyond conscious control. Similarly, the results of previous studies examining the composite effect for facial identity (Carey & Diamond, 1994; Young et al., 1987) suggest that the same is true for the perception of configural information relating to facial identity.

Given that Experiment 4 used the same basic task used in the experiments discussed above (i.e., identify the person, or identify the expression shown in the bottom half of the composite), we could see little reason for including a series of noncomposite conditions as a check of an effect that is apparently beyond the participants' control. In addition, Experiment 3 had demonstrated that RTs to identify the expression shown in the bottom section of same-expression–different-identity and different-expression–different-identity composites were significantly different. This was consistent with the idea that the configural information for facial expression was disrupted in one condition (different-expression–different-identity) but not the other (same-expression–different-identity). Hence, the same-expression–different-identity composites essentially served as a control for the expression recognition task (i.e., a similar role to the noncomposite images used in Experiments 1, 2, and 3). Similarly, we used the different-expression–same-identity images as a control for the identity recognition task. This is because we predicted that participants should show significantly less interference with these images compared with the composites in which the top and bottom sections contained different people's faces. As it turned out, our predictions were confirmed.

Method

Participants. Fifteen participants (13 women, 2 men) aged between 18 and 50 years and from the same population as Experiments 1–3 took part in the experiment. All had normal or corrected-to-normal vision, and none had participated in the previous experiments.

Materials. All stimuli in Experiment 4 were prepared from pictures of three models (C, NR, and SW) from the Ekman and Friesen (1976) series, each posing three different facial expressions (happiness, disgust, and surprise). All possible combinations of the top and bottom halves of these nine different pictures were produced to give a total of 72 different composite faces. For 18 of these, the top and bottom halves displayed different expressions posed by the same model (different-expression–same-identity composites; e.g., top = happiness Model C, bottom = disgust Model C); for 36, the top and bottom halves displayed different expressions posed by the different models (different-expression–different-identity composites; e.g., top = happiness Model C, bottom = disgust Model NR) and for the remaining 18 the top and bottom halves displayed the same expression posed by different models (same-expression–different-identity composites; e.g., top = happiness Model C, bottom = happiness Model NR). Note that the original images (literally same expression–same identity) were not used in the experiment proper. Examples of the three image types prepared from two of the models used in Experiment 4 are shown in Figure 8.

Design and procedure. Two within-subjects factors were investigated: composite type (different expression–same identity, same expression–different identity, and different expression–different identity) and task instructions ("identify the person" and "identify the expression"). All stages of the experiment used the presentation format described for Experiment 2 (i.e., 500-ms fixation, 500-ms blank interstimulus interval and then the stimulus that remained in view until the participant responded).

The experiment consisted of two sections corresponding to the two levels of the task instructions factor (identify the person and identify the expression); half of the participants were assigned to the identify-the-person section first and half to the identify-the-expression section. Both sections used the same basic design.

Identify-the-person trials. The section began with a training session. In this training session, the three models' faces were presented with neutral facial expressions (expressionless faces), and each was accompanied by an arbitrarily assigned first name (Model C = Susan, Model NR = Margaret, Model SW = Tracy); these names were printed in uppercase letters and positioned below the face. Each face–name pair was presented five times for 5 s each in random order. The participant was instructed to look at the faces and try to remember the models' names because later they would be tested on them. Following this, pictures of the same models posing the three facial expressions, happiness, disgust, and surprise, were presented individually and without name labels. The participant was asked to identify each model's name by pressing one of three buttons on a box interfaced with the computer; the keys were marked with the names (Susan, Margaret, and Tracy), and their positions were counterbalanced across participants. Each of the nine different pictures (three models × three facial expressions) were presented three times in random order. If participants made an error (e.g., pressed the Susan button in response to Margaret's face), the computer made a "beep" noise, and they were invited to try again until the correct response was made. Participants who each made more than 3 errors (out of a total of 27 trials; maximum total errors = 54; i.e., 2 errors per trial) in this stage of the experiment were excluded from the analysis.

Next, half of the participants were presented with the top halves of the same faces and half of the participants with the bottom halves of the same faces. Again, each image was presented individually, three times in random order, and the participant's task was to indicate each model's name by making a button-press response. Following this, the participants who had seen the top sections were presented with the bottom sections of the same faces and vice versa. Their task was the same: to identify the models' names.

In the experiment proper the participants were presented with equal numbers (18) of the three types of composite faces (different expression–same identity, different expression–different identity, and same expression–different identity) in random order; each image was presented twice. The stimuli were counterbalanced across two different stimulus sets to accommodate the different number of images in the three levels of the composite type condition. Half of the participants were assigned to one stimulus set and half to the other. The composite images subtended a horizontal visual angle of approximately 4.6°, and a vertical visual angle of approximately 6.3°. Participants were instructed to identify the name of the model shown in the bottom half of each composite image by pressing one of the three keys listed above. To familiarize

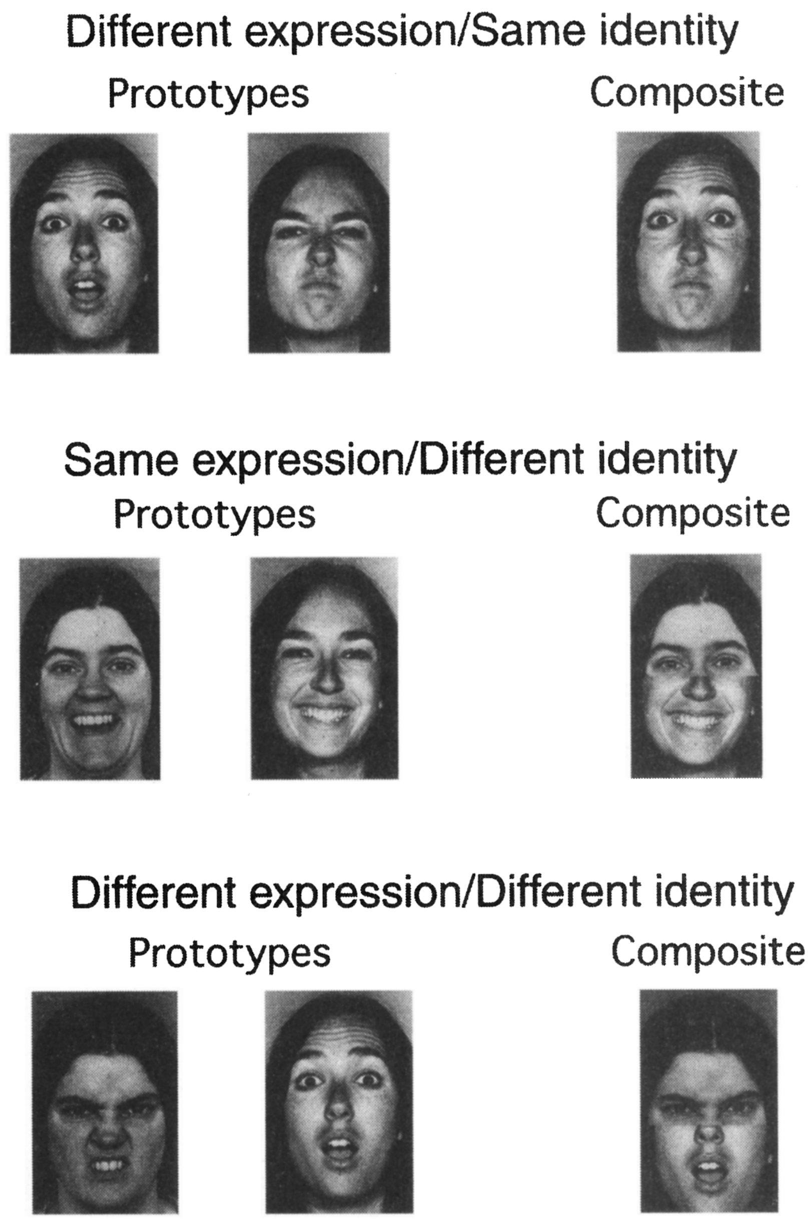

Figure 8. Examples of the stimuli used in Experiment 4. Composite facial expressions were prepared from the top and bottom halves of (a) two different prototype expressions posed by the same model (different expression–same identity; top row), (b) the same prototype expression posed by different models (same expression–different identity; middle row), and (c) two different prototype expressions posed by different models (different expression–different identity; bottom row). Images from *Pictures of Facial Affect,* by P. Ekman and W. V. Friesen, 1976. Copyright 1976 by P. Ekman and W. V. Friesen. Adapted with permission.

the participants with the composite images, the experiment proper was preceded by six practice trials selected at random from the 54 experimental trials.

Identify-the-expression trials. The identify-the-expression section began with exactly the same training session described above, in which each model was presented five times with a neutral expression and their name label. Next, the same three models were presented posing the expressions happiness, surprise, and disgust, and participants were asked to identify their facial expression by pressing one of three keys labeled *happiness, disgust,* and *surprise;* each face was presented three times in random order. If the participant made an error in his or her choice, the computer made a beep noise and they were asked to try again. Any participant who made more than 3 errors out of a total of 54 was again excluded from the analysis. In all other respects, the design of the identify-the-expression trials was the same as the identify-the-person trials described above. The only difference was in the task instructions.

Hence, participants were next presented with the top segments of the stimuli in one block and bottom segments of the same faces in another block; order of presentation of these two blocks was counterbalanced across participants. Their task in each case was to identify the facial expression by pressing one of three keys marked *happiness, disgust,* and *surprise.* In the experiment proper, they identified the expression shown in the bottom half of the composite stimuli.

Results

In the first block of the identify-the-person trials in which the participants were shown the whole faces and asked to indicate the models' name, 3 of the participants made three errors (the criterion number for rejection). In the corresponding section of the identify-the-expression block, the same 3 participants reached or exceeded this criterion error rate. These participants were therefore excluded from the following analysis, leaving data from 12 participants (10 women, 2 men; 18–50 years). The mean number of correct responses made by these 12 participants when identifying the expression and person in the whole-face block in the respective sections were as follows: Identify the person, $M = 26.42$, $SD = 0.90$; and identify the expression, $M = 26.50$, $SD = 0.80$. Clearly, then, these participants had little difficulty in identifying the models' names or their facial expressions.

Participants' mean correct RTs (with standard error bars) to identify the person and expression in the bottom half of the three types of composite images (different expression–same identity, different expression–different identity, and same expression–different identity) are shown in the left graph of Figure 9. The right graph shows participants' mean error proportions (with standard error bars) for the same experiment.

RTs. Our principal form of analysis involved RTs for correct responses. These were submitted to a two-factor ANOVA investigating composite type (different expression–same identity, different expression–different identity, and same expression–different identity; repeated measure) and task instructions (identify the person and identify the expression; repeated measure). There was a significant effect of composite type, $F(2, 22) = 7.39, p < .005$. Post hoc t tests ($p < .05$) indicated that, overall, participants were slowest to identify the bottom segments of the different-identity–different-expression composites; RTs to the different-identity–same-expression and same-identity–different-expression composites did not reliably differ. The main effect of composite type was qualified by a significant interaction between composite type and task instructions, $F(2, 22) = 14.39, p < .0001$. Simple effects analyses showed a significant effect of composite type for both levels of task instructions condition: Identify the person, $F(2, 22) = 12.01, p < .0001$, and identify the expression, $F(2, 22) = 9.82, p < .001$; however, the pattern of the effects in these two conditions was different. Post hoc t tests ($p < .05$) showed that for the identify-the-person condition, RTs to the different-expression–same-identity composites were significantly faster than those to the same-expression–different-identity and different-expression–different-identity compos-

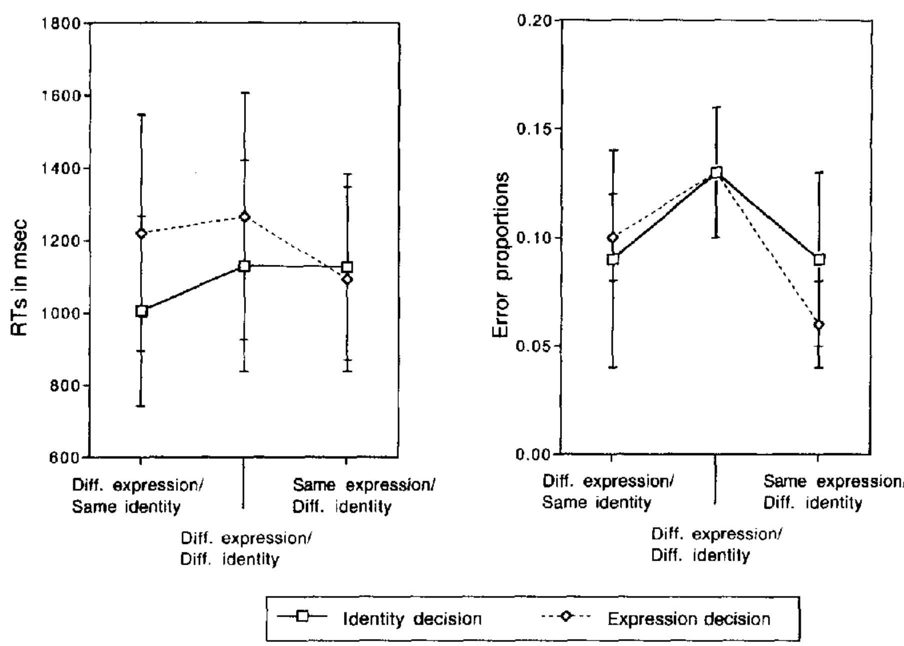

Figure 9. Data from Experiment 4. The left graph shows participants' mean correct reaction times (RTs; with standard error bars) to identify the expression (expression decision) or identity (identity decision) in the bottom segment of three types of composite image (different expression–same identity, different expression–different identity, and same expression–different identity). The right graph shows participants' mean error proportions (with standard error bars) from the same experiment.

ites, which did not reliably differ (different expression–same identity < [same expression–different identity = different expression–different identity]). For the identify-the-expression condition, RTs to the same-expression–different-identity composites were significantly faster than those to the different-expression–same-identity and different-expression–different-identity composites, which did not reliably differ (same expression–different identity < [different expression–same identity = different expression–different identity]). Finally, there was no overall significant effect of task instructions, indicating that participants' RTs to perform the identify-the-person and identify-the-expression tasks did not reliably differ; this shows that the two tasks were of comparable difficulty, as assessed by the RTs measure.

In summary, the results of the RTs analysis demonstrate that participants were significantly slower to perform the task when the attribute (expression or identity) that they were asked to attend to was incongruent across the two face halves. Moreover, there was no additional significant cost when the unattended attribute was also incongruent across the two halves. Note that this result is consistent with the breakdown of the main effect of composite type. This showed that, overall, participants were slowest to classify composites in which the top and bottom halves were different identities and different expressions. This result is to be expected because we predicted that the different-expression–different-identity condition should show slower RTs in both levels of the task instructions condition (identify the person and identify the expression).

Error rates. A subsidiary analysis examined participants' error rates to check that the slower responses were not also accompanied by more accurate performance. Error proportions were arcsin transformed and submitted to a two-factor ANOVA investigating composite type (different expression–same identity, same expression–different identity, and different expression–different identity; repeated measure) and task instructions (identify the person and identify the expression; repeated measure). The only significant effect was the main effect of composite type (different expression–same identity, same expression–different identity, and different expression–different identity), $F(2, 22) = 3.76, p < .05$. Post hoc t tests ($p < .05$) showed that, overall, participants made significantly more errors in the different-expression–different-identity condition compared with the same-expression–different-identity condition. There were no other statistically reliable effects. Thus, there was no evidence of participants trading accuracy for speed. Furthermore, the absence of a significant effect of task instructions indicates that the two tasks (identify the person and identify the expression) were of comparable difficulty, as assessed by error proportions; this is consistent with the findings of the RT analysis.

Discussion

The results of Experiment 4 showed that when viewing the same stimulus set, participants produced different patterns of RTs depending on whether they were asked to perform a facial identity task or a facial expression task.

Three types of composite image were used in which the top and bottom halves were (a) different expressions posed by the same model (different-expression–same-identity composites), (b) the same expression posed by different models (same-expression–different-identity composites), and (c) different expressions posed by different models (different-expression–different-identity composites). When the task instruction was to identify the facial expression shown in the bottom half of these composites, participants' RTs were significantly faster when the top and bottom halves contained the same expression (same-expression–different-identity composites), than when they contained different expressions (different expression–same identity and different expression–different identity). Moreover, for the two conditions in which the top and bottom segments of the composite showed different expressions (different expression––same identity and different expression–different identity), there was no significant additional cost when the two halves contained both different expressions and different identities (different expression–different identity).

When the task was to recognize the identity shown in the bottom half of the composites a different pattern of performance was observed. Here participants' RTs were significantly faster when the top and bottom segments showed the same identity (different-expression–same-identity composites) than when they contained different identities (same expression–different identity and different expression–different identity). And for the two conditions in which the two halves of the composite contained different identities (same-expression–different-identity and different-expression–different-identity composites), there was no additional significant cost when they displayed both different identities and different expressions (different-expression–different-identity composites).

These results are consistent with previous findings showing that participants can selectively attend to information in a face that is relevant to its expression and discard information relevant to its identity, or vice versa (Campbell et al., 1996; Etcoff, 1984; Young et al., 1986). However, the results of Experiment 4 go beyond these previous studies and offer an impressive demonstration of participants' ability to selectively attend to different types of configural information; one relating to the representation of facial identity the other to the representation of facial expression. Experiment 4 also demonstrates that these two forms of configural information can be selectively disrupted. One implication of this finding is that the configural information used for facial identity and facial expression perception is different.

Finally, it is worth noting that these results were obtained using a within-subjects design. This shows that the participants have an impressive ability to shift their attention from processing configural information that is relevant to facial identity in one block to processing configural information that is relevant to facial expression in another without experiencing any substantial interference from the immediately preceding task.

In summary, Experiment 4 suggests that different configural information is used to encode facial identity and facial expression. The nature of these two types of configural features is discussed in the following section.

General Discussion

The results of these experiments demonstrate a number of points. These can be summarized as follows.

1. Facial expressions of the basic emotions can be divided into recognizable-top and recognizable-bottom categories. Experiment 1 found that anger, fear, and sadness showed a significant recognizable-top bias, whereas happiness and disgust showed a significant recognizable-bottom bias. Surprise showed no significant bias and was equally distinguishable from whole-face, top and bottom segments. These results largely replicate the findings of a previous study by Bassili (1979).

2. Composite facial expressions were prepared by aligning the top half of one facial expression (e.g., anger) with the bottom half of another (e.g., happiness). In three separate experiments, we have demonstrated that participants are significantly slower to identify the expression in either half of these composite images relative to a noncomposite control condition in which the two halves are misaligned (Experiments 1, 2, and 3). These results parallel Young et al.'s (1987) earlier finding of a similar effect for facial identity.

3. Young et al. (1987) demonstrated that the composite effect for facial identity is abolished when the stimuli are inverted. The results of Experiment 3 demonstrate that the composite effect for facial expression is also significantly disrupted by stimulus inversion.

4. The composite effect for facial expression is found when the top and bottom segments are taken from pictures of different expressions posed by the same model (Experiments 1 and 2) or two different models (Experiment 3). However, this effect is not found when the two segments are taken from pictures of different models posing the same facial expression (Experiment 3). This result serves to exclude the suggestion that the composite effect is an artefact of stimulus quality—for example, the composite stimuli appearing slightly distorted and unusual as faces and, hence, more attention-grabbing than the noncomposite images.

5. In Experiment 4, participants were presented with three types of composite faces in which the top and bottom segments were (a) different expressions posed by the same model, (b) the same expression posed by different models, and (c) different expressions posed by different models. When participants were asked to name the identity shown in the bottom half of these images, their RTs were significantly slower for those composites prepared from the top and bottom halves of different models' faces (b and c above). However, when they were asked to identify the expression shown in the same half, significantly slower RTs were found for images in which expression was incongruent across the two halves (a and c above). Moreover, no added cost was found when the attribute that participants had not been instructed to report (e.g., expression, in the identity task) was also incongruent across the two face halves. These findings suggest that the composite effects for facial expression and facial identity may disrupt the perception of different configural features.

These results have important implications for the perceptual representations of facial expressions and we deal with each of them in turn.

As we discussed in the introduction, previous studies have shown that some facial expressions are more recognizable from the top half of the face (recognizable-top expressions), whereas others are more readily recognized from the bottom half (recognizable-bottom expressions, Bassili, 1979; Hanawalt, 1944; Plutchik, 1962). Our results confirm these observations and are highly consistent with Bassili's findings that were obtained using animated examples of the same facial expressions from a different image set.

It is worth emphasizing that observations of top–bottom expression dominance are not inconsistent with the idea that configural information is important for facial expression recognition. It is possible for the overall configuration of a facial expression to contribute toward its recognition, despite the sufficient information for accurate recognition of an the emotion being contained largely in the top (or bottom) section of the face. Likewise, the observation that a person's identity is more readily recognized from the eye region than the mouth region (see Shepherd, Davies, & Ellis, 1981, for a review) does not detract from the well-established finding that configural features are important for facial identity recognition. We should also point out that in the preliminary experiment conducted in Experiment 1 (see Table 1), none of the top or bottom sections of the expressions were recognized at chance (chance error proportion = 0.83). This means that for all six expressions, both top and bottom sections of the face contained information that was associated with the emotion.

In the introduction, we outlined a configural model and a part-based model of facial expression recognition. Consistent with a configural model, Experiments 1, 2, and 3 showed that a facial composite effect, similar to the one shown for facial identity by Young et al. (1987), can also be found for facial expression. Young et al. suggested that the composite effect for facial identity reflects a disruption of configural information processing, because when the top and bottom halves of two identities' faces are aligned, they produce a new facial configuration that interferes with one's ability to recognize the identity in the top or bottom part of the face. We think that a similar explanation can be applied to the composite effect for facial expression. That is, the top and bottom halves of the two expressions align to produce a new facial expression configuration. Consequently, this interferes with identifying the emotion shown in either half of the composite expressions. Misaligning the two face halves, however (noncomposite condition), means that the face is no longer encoded as a configural whole, and, hence, the feature information relating to the expression in the top and bottom halves can be accessed faster. The results of our experiments, then, suggest that the composite effects for facial identity and facial expression are somewhat similar. This similarity is further emphasized by our observation that the composite effect for facial expression is disrupted by inverting the stimuli.

Recall that Young et al. (1987) showed that the composite effect for facial identity is only found when the stimuli are presented upright; inverting the stimuli (i.e., 180° rotation) abolished the effect (see also Carey & Diamond, 1994). This finding is consistent with Carey and Diamond's (1977) suggestion that inversion impairs the perception of configural information. Hence, our observation that the composite effect for facial expression is also disrupted by stimulus inversion further supports a configural model of facial expression recognition.

In Experiment 3, we addressed the criticism that the composite effect could be attributed to some inherent quality of composite faces that makes them more attention-grabbing than the noncomposites. However, Experiment 3 discounted this interpretation by showing that a composite effect is not found when the top and bottom sections contain the same facial expression posed by different models (same-expression composites). This finding also suggested a further hypothesis.

From examining the same-expression composites, our intuition was that the configural features used for facial expression recognition were different to those used for facial identity. We noted that although composites composed of the top and bottom sections of two people's faces no longer resemble either of the original faces, if both faces are posing the same expression (e.g., happiness), the composite face also looks happy. Similarly, inspection of the stimuli used in Experiment 1 showed that the opposite was true. That is, composites prepared from the top and bottom sections of different expressions posed by the same model were also highly identifiable as that particular model, although the composite facial expression resembled neither of the two starting expressions. This seemed to suggest that the composite effects for facial identity and facial expression were tapping two different types of configural processing. The results of Experiment 4 were consistent with this intuition.

Experiment 4 showed that either form of configural interference (identity or expression) can be produced from the same set of composite faces depending on whether participants are instructed to attend to the faces' identity or their expression. Thus, when asked to report the identity shown in the bottom half of a composite face, participants were significantly slower if the two halves contained different models' faces. Likewise, participants were slower to report facial expression if the two halves showed different expressions. More important, however, no significant cost was produced if the unattended attribute (e.g., expression in the identity task) was incongruent across the two face halves. Nor was there any additional cost when both attended and unattended attributes were incongruent relative to the condition in which the attended attribute alone was incongruent. These observations suggest that participants were encoding different types of configural information when processing facial identity and facial expression.

Configural Information for Facial Identity and Facial Expression

As we discussed in the introduction section, cues to facial expression and facial identity are generally thought to be processed by separate cognitive routes (Bruce & Young, 1986; Hay & Young, 1982; Young & Bruce, 1991; Young et al., 1993). It seems entirely plausible, then, that these parallel processing routes should use different types of visual information from the same facial image. What is slightly more contentious, however, is the idea that these two routes should process different types of configural information. However, in line with this idea, it is worth remembering that Diamond and Carey (1986) identified two forms of configural features, which they referred to as first-order and second-order relational properties.

The term first-order relational properties refers to the raw interfeature relationships that are common to all normal faces—two horizontally positioned eyes, above a central nose, above a central mouth, and so forth—which is effectively the spatial information that makes up a face. Second-order relational properties are substantially more subtle and are what are more generally referred to as simply configural features. As we discussed earlier, these features are the interrelationships between different feature positions and shapes that help distinguish one facial identity from all others (e.g., the distance between the eyes; position and shape of the nose in relation to the position and shape of the mouth, etc.). Furthermore, it is generally thought to be these second-order features that are disrupted by inversion and by the composite effect for facial identity. At first glance, then, it seems natural to assume that second-order features are also disrupted in the composite effect for facial affect shown here. But, as we have already discussed, this explanation is inconsistent with the observation that configural information for facial identity and facial expression can be selectively disrupted (Experiment 4). Instead, this finding points to the conclusion that the configural cues to these two facial attributes are different. Hence, one possibility is that the composite effect for facial expression may reflect a disruption of a more coarse form of configural information, one more akin to first-order relational properties.

As applied to facial identity by Diamond and Carey (1986), first-order relational properties are the interfeature relationships that are common to all faces (i.e., the average or prototype configuration associated with all faces one has encountered). For facial expressions, we suggest that there are the interfeature relationships that make a surprise expression surprised, or happiness expression happy, etc. In other words, we suggest that each emotional facial expression is associated with its own average configuration. The average configuration could be regarded as a distinct representation that is abstracted from encountered exemplars of each type of facial expression (happiness, sadness, anger, fear, disgust, surprise, etc.). Alternatively, it could be envisaged as the centroid of a cluster of stored exemplar representations, with separate clusters for each emotion category. In other words, it is not necessary for the average to exist as a distinct prototype structure in its own right.

In relation to this discussion of average expression representations, it is worth remembering that Ekman and his colleagues have shown that each emotion category is associated with more than one facial structure. For example, in the case of surprise, the mouth could be anything between wide open and closed, or, in anger, the eyes can be narrowed or wide open. Nonetheless, the different variants of each expression contain enough common features for them to form a cluster around an average or prototype configuration.

We should point out that we do not wish to imply that facial expressions are coded in terms of configural information alone, and we have no fundamental objection to the idea that the individual features of facial expressions are also important for recognition (Ellison & Massaro, 1997). Hence, a facial expression, such as surprise, might be encoded in terms of its individual features (i.e., raised eyebrows, wide open eyes, and a wide open mouth) and in terms of its facial configuration (i.e., a symmetrical arrangement of raised eyebrows, above wide open eyes, above a wide open mouth). Consequently, when the eye and eyebrow regions from a surprise expression are aligned with the bottom half of a face displaying a different expression (e.g., one in which the upper lip is raised to signal disgust), the overall representation no longer resembles the average configuration for surprise (or disgust). Hence, although participants are able to use the individual features in the top (or bottom) half of the face to identify the expression, this process is slowed by the configural mismatch. In the noncomposite condition, however, there is no conflicting configural information because the face halves are misaligned. Hence, the participants can use the information in one face half to identify the emotion without experiencing interference from an unusual facial configuration.

Configural and Part-Based Models of Facial Expression Recognition: Weighing Up the Evidence

So where does the above discussion leave us in relation to Ellison and Massaro's (1997) largely part-based account of facial expression recognition? As we discussed in the introduction, these authors asked participants to identify whole-face expressions in which two different features (eyebrows and mouth corners) were manipulated. Participants were also presented with the individual features shown in the context of the upper and lower sections of the face. For each image, they were asked to make a binary decision response: Is the emotion expressed happiness or anger? By modeling the data using the fuzzy logical model of perception, Ellison and Massaro showed that participants' categorization of whole-facial expressions could be reliably modeled by assuming that the critical features of the face (eyebrow and mouth sections) are evaluated separately.

The composite paradigm is not dissimilar to Ellison and Massaro's (1997) task in that it also uses stimuli prepared by recombining the upper and lower sections of different facial features. Nonetheless, our data do not concur with their findings. It is relevant, then, that we consider why, but in doing so, it is important to recognize two points.

First, the FLMP model does not exclude the possibility that configural features are used to encode facial expressions (Ellison & Massaro, 1997; Massaro, 1998). The only constraint the FLMP makes on the information used to encode facial affect is that each feature must be evaluated as an independent perceptual unit. Hence, if one assumes that configural information can be encoded as one or more independent units, then our data are not at odds with the FLMP.

The second important point to take note of is that our own data do not exclude the possibility that part-based information is used for facial expression recognition. The data simply rule out the idea that configural information is not used for recognizing facial expressions. In actual fact, we think that our data actively support the suggestion that both configural and part-based information are used to decode facial affect; otherwise, the participants would have found it virtually impossible to identify the emotions in the upper or lower parts of the face, because with the exception of Experiment 3 (same-expression composites), the overall configuration did not match the emotion shown in either half of the face.

With these factors in mind, what are the differences between the two studies that may account for the different results? First, perhaps the main difference between the two studies is that Ellison and Massaro (1997) used identification rates and affect ratings as their dependent measures, whereas the composite paradigm uses RTs as the principal measure of interest. In relation to this point, it may be relevant that in our own series of experiments, a significant composite effect was found for the RT measure in all of the experiments, but only Experiment 3 showed a reliable composite effect for the error data, although there was no evidence of a speed–accuracy trade-off in the other experiments. Hence, it is possible that RTs provide a more sensitive measure of the configural contribution.

It is interesting that Ellison and Massaro (1997) reported that participants' RTs were longer for "ambiguous expressions" in their study; these included faces in which the top and bottom sections displayed different emotional signals (e.g., top = anger, bottom = happiness). Without a noncomposite condition, however, it is difficult to determine whether the longer RTs reflect interference between the different emotional concepts expressed in the two face halves or whether (as we have found) there was some additional interference from the inappropriate configuration.

The second point to take note of is that our experiments used a number of human models' faces from the Ekman and Friesen (1976) set with expressions associated with six basic emotions (happiness, sadness, anger, fear, disgust, and surprise). These images are based on an anatomical analysis of facial affect, and their repeated use in psychological studies verifies that the expressions are highly recognizable. A fact that is further substantiated by the preliminary experiment described in Experiment 1, this experiment also demonstrated that the upper and lower face sections used to prepare the composites were reliably identified using a six-way, forced-choice task. Ellison and Massaro's (1997) study, on the other hand, used a single (computer-generated) synthetic face posing two facial expressions, anger and happiness. Hence, there were differences between the number of expressions and number of different examples of the expressions used in the two experiments.

To expand further on the last point, it may also be relevant that we used a three-way decision task, whereas Ellison and Massaro (1997) used a binary response task. One of the problems in using a binary decision task is that one cannot be sure that the participants spontaneously recognize the facial signals used as the intended emotions (e.g., happy and angry). For example, it is possible that participants might use a strategy of classifying the images as "happy" and "not happy." It is interesting that previous studies that have applied the composite effect to the recognition of facial identity have used a vocal response task with at least four response options (Carey & Diamond, 1994; Young et al., 1987). Hence, it is possible that by making the task more demanding (by increasing the number of response options), one can increase the paradigm's sensitivity to configural interference.

We also feel that it worth emphasizing once again that the composite paradigm is particularly suited to differentiating between part-based and configural accounts of facial expression recognition, because the same part-based information is present in both composite and noncomposite conditions. This means that for the two conditions, any interference between the emotional concepts expressed in the two separate halves should be constant. Hence, any difference in the RTs between the composite and noncomposite conditions would appear to reflect the difference in coding a facial (composite) as opposed to a nonfacial (noncomposite) image. In this sense, the data speak for themselves: RTs are slower when the two halves are aligned to form a facial expression configuration (composite condition) than when they are misaligned and there is no facial expression configuration (noncomposite condition). Having demonstrated this finding in a number of experiments, we conclude that these data are consistent with a configural model of facial affect recognition in which both configural and part-based information is used to identify the emotion. The use of part-based information is further substantiated by the findings of Ellison and Massaro (1997), and by our own observations that recognition of expressions of partial faces does not fall to chance level.

As we have already emphasized, this conclusion is not inconsistent with the FLMP, provided that one assumes the configural information can be evaluated as one or more independent perceptual units. To successfully demonstrate that this is the case, however, one would first have to identify the important configural features and then assess their contribution using the sort of extended factorial design that has become associated with FLMP research (Massaro, 1998). Given that 20 years of research into configural coding of facial identity have failed to identify the precise nature of the configural information for facial identity, this seems a tall order for facial expression research.

Finally, it worth noting that our results are also in line with a recent model facial expression production outlined by Smith and Scott (1997). These authors suggested that each emotional facial expression is made up of a number of individual features (or components) and that at least some of these features are in themselves meaningful. However, they also suggested that when the individual features are produced in combination (i.e., in the form of a facial expression of anger, disgust, or sadness, etc.), the overall facial configuration may convey additional information that is not captured by the individual features themselves. In other words, these authors suggested that for facial expression production, the whole is more than the sum of the parts. Similarly, we think that our own results demonstrate that when faced with a facial expression image, the perceptual system not only analyzes cues present in individual features but also the configuration of interrelationships between these features.

References

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37,* 2049–2058.

Bruce, V. (1988). *Recognising faces.* London: Erlbaum.

Bruce, V., Doyle, Y., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition, 38,* 109–144.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identity of faces. *Perception, 1994,* 803–822.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327.

Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996). Categorical perception of morphed facial expressions. *Visual Cognition, 3,* 81–117.

Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society London, 264,* 919–925.

Campbell, R., Brooks, B., de Haan, E., & Roberts, T. (1996). Dissociating face processing skills: Decisions about lip-read speech, expression and identity. *Quarterly Journal of Experimental Psychology, 49A,* 295–314.

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

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition, 1,* 253–274.

Coleman, J. C. (1949). Facial expressions of emotion. *Genetic Psychology Monographs, 63*(1), Whole No. 296.

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

Dunlap, K. (1927). The role of eye-muscles and mouth-muscles in the expression of the emotions. *Genetic Psychology Monographs, 2,* 199–233.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. K. Cole (Ed.), *Nebraska Symposium on Motivation* (pp. 207–283). Lincoln: University of Nebraska Press.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Englewood Cliffs, NJ: Prentice Hall.

Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect.* Palo Alto, California: Consulting Psychologists Press.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1972). *Emotion and the human face: Guidelines for research and an integration of findings.* New York: Pergamon Press.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., Ayhan LeCompte, W., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., & Tomita, M. (1987). Universals and cultural differences in the judgement of facial expressions of emotion. *Journal of Personality and Social Psychology, 53,* 712–717.

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 213–226.

Endo, M., Masame, K., & Maruyama, K. (1989). Interference from configuration of a schematic face onto the recognition of its constituent parts. *Tohoku Psychologica Folia, 48,* 97–106.

Endo, M., Takahashi, K., & Maruyama, K. (1984). Effects of observer's attitude on the familiarity of faces: Using the difference in cue value between central and peripheral facial elements as an index of familiarity. *Tohoku Psychologica Folia, 43,* 23–34.

Etcoff, N. L. (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia, 22,* 281–295.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44,* 227–240.

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

Frois-Wittmann, J. (1930). The judgement of facial expression. *Journal of Experimental Psychology, 13,* 113–151.

George, M. S., Ketter, T. A., Gill, D. S., Haxby, J. V., Ungerleider, L. G., Herscovitch, P., & Post, R. M. (1993). Brain regions involved in recognizing facial emotion or identity: An oxygen-15 PET study. *Journal of Neuropsychiatry, 5,* 384–394.

Haig, N. D. (1984). The effect of feature displacement on face recognition. *Perception, 13,* 505–512.

Hanawalt, N. G. (1944). The role of the upper and lower parts of the face as the basis for judging facial expressions: II. In posed expressions and "candid camera" pictures. *Journal of General Psychology, 31,* 23–36.

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research, 32,* 203–218.

Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 173–202). London: Academic Press.

Hill, H., & Bruce, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 986–1004.

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

Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioural principle.* Cambridge, MA: MIT Press.

Massaro, D. W., & Cohen, M. M. (1990). Perception of synthesized audible and visible speech. *Psychological Science, 1,* 55–63.

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology, 34,* 325–334.

Parks, T. E., Coss, R. G., & Coss, C. S. (1985). Thatcher and the Cheshire cat: Context and the processing of facial features. *Perception, 14,* 747–754.

Parry, F. M., Young, A. W., Saul, J. S. M., & Moss, A. (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical and Experimental Neuropsychology, 13,* 545–558.

Plutchik, R. (1962). *The emotions: Facts theories and a new model.* New York: Random House.

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

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

Sergent, J., Ohta, S., MacDonald, B., & Zuck, E. (1994). Segregated processing of emotional identity and emotion in the human brain: A PET study. *Visual Cognition, 1,* 349–369.

Shepherd, J. W., Davies, G. M., & Ellis, H. D. (1981). Studies of cue saliency. In G. M. Davies, H. D. Ellis, & J. Shepherd (Eds.), *Perceiving and remembering faces* (pp. 105–131). London: Academic Press.

Smith, C. A., & Scott, H. S. (1997). A componential approach to the meaning of facial expressions. In J. A. Russell & J. M. Fernandez-Dols (Eds.), *The psychology of facial expressions* (pp. 229–254). Cambridge, England: Cambridge University Press.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

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

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

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

Valentine, T., & Bruce, V. (1988). Mental rotation of faces: Memory and Cognition. *Memory and Cognition, 16,* 556–566.

Wallbott, H. G., & Ricci-Bitti, P. (1993). Decoders' processing of emotional facial expression—A top-down or bottom-up mechanism. *European Journal of Social Psychology, 23,* 427–443.

Young, A. W., & Bruce, V. (1991). Perceptual categories and the computation of "grandmother." *European Journal of Cognitive Psychology, 3,* 5–49.

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

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. *Psychological Research, 48,* 63–68.

Young, A. W., Newcombe, F., de Haan, E. H. F., Small, M., & Hay, D. C. (1993). Face perception after brain injury: Selective impairments affecting identity and expression. *Brain, 116,* 941–959.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Megamixing facial expressions. *Cognition, 63,* 271–313.

Appendix

Experimental Faces: Identifier in Ekman and Friesen's (1976) Series and Percentage Recognition as This Emotion in Their Norms

Experiment 1: Identification of the Top and Bottom Sections of Ekman and Friesen (1976) Faces

Happiness (M = 99%). 7 C-2-18; 14 EM-4-07; 34 JJ-4-07; 48 MF-1-06; 57 MO-1-04; 66 NR-1-06; 74 PE-2-12; 85 PF-1-06; 93 SW-3-09; 101 WF-2-12.

Surprise (M = 91%). 11 C-1-10; 19 EM-2-11; 39 JJ-4-13; 54 MF-1-09; 63 MO-1-14; 70 NR-1-14; 81 PE-6-02; 90 PF-1-16; 97 SW-1-16; 107 WF-2-16.

Fear (M = 90%). 9 C-1-23; 16 EM-5-21; 37 JJ-5-13; 50 MF-1-26; 59 MO-1-23; 68 NR-1-19; 79 PE-3-21; 88 PF-2-30; 95 SW-2-30; 104 WF-3-16.

Sadness (M = 90%). 8 C-1-18; 15 EM-4-24; 36 JJ-5-05; 49 MF-1-30; 58 MO-1-30; 67 NR-2-15; 75 PE-2-31; 86 PF-2-12; 94 SW-2-16; 102 WF-3-28.

Disgust (M = 93%). 12 C-1-04; 20 EM-4-17; 40 JJ-3-20; 55 MF-2-13; 64 MO-2-18; 71 NR-3-29; 82 PE-4-05; 91 PF-1-24; 98 SW-1-30; 108 WF-3-11.

Anger (M = 90%). 10 C-2-12; 18 EM-5-14; 38 JJ-3-12; 53 MF-2-07; 61 MO-2-11; 69 NR-2-07; 80 PE-2-21; 89 PF-2-04; 96 SW-4-09; 105 WF-3-01.

Experiments 1, 2, and 3

Happiness (M = 98%). 7 C-2-18; NR-1-06; 85 PF-1-06; 93 SW-3-09.

Surprise (M = 92%). 11 C-1-10; 70 NR-1-14; 90 PF-1-16; 97 SW-1-16.

Fear (M = 88%). 9 C-1-23; 68 NR-1-19; 88 PF-2-30; 95 SW-2-30.

Sadness (M = 94%). 8 C-1-18; 67 NR-2-15; 86 PF-2-12; 94 SW-2-16.

Disgust (M = 95%). 12 C-1-04; 71 NR-3-29; 91 PF-1-24; 98 SW-1-30.

Anger (M = 88%). 10 C-2-12; 69 NR-2-07; 89 PF-2-04; 96 SW-4-09.

Experiment 4

Happiness (M = 97%). 7 C-2-18; NR-1-06; 93 SW-3-09.
Surprise (M = 92%). 11 C-1-10; 70 NR-1-14; 97 SW-1-16.
Disgust (M = 94%). 12 C-1-04; 71 NR-3-29; 98 SW-1-30.

Received May 6, 1998
Revision received January 12, 1999
Accepted April 20, 1999 ■