

Postprint Version	1.0
Journal website	http://www.informaworld.com/smpp/content?content=10.1080/13506280444000418
Pubmed link	
DOI	10.1080/13506280444000418

This is a NIVEL certified Post Print, more info at <http://www.nivel.eu>

Configural coding of facial expressions: The impact of inversion and photographic negative

ANDREW J. CALDER^A; JESSE JANSEN^B

^a MRC Cognition and Brain Sciences Unit, Cambridge, UK

^b Netherlands Institute for Health Services Research, Utrecht, The Netherlands

ABSTRACT

In previous research we used the composite paradigm (Young, Hellawell, & Hay, 1987) to demonstrate that configural cues are important for interpreting facial expressions. However, different configural cues in face perception have been identified, including holistic processing (i.e., perception of facial features as a single gestalt) and second-order spatial relations (i.e., the spatial relationship between individual features). Previous research has suggested that the composite effect for *facial identity* operates at the level of holistic encoding. Here we show that the composite effect for facial expression has a similar perceptual basis by using different graphic manipulations (stimulus inversion and photographic negative) in conjunction with the composite paradigm. In relation to Bruce and Young's (1986) functional model of face recognition, a suitable level for the composite effect is a stage of front-end processing referred to as structural encoding, that is common to both facial identity and facial expression perception.

Maurer and colleagues (Maurer, le Grand, & Mondloch, 2002) recently presented a review of research addressing configural processing of faces. Despite the fact that faces convey a whole host of information, including a person's identity, emotional state, sex, age, and physical attractiveness, and so on, Maurer et al.'s paper illustrates that virtually all studies addressing the role of configural facial cues have used just three types of tasks—"is it a face", "are the two faces the same", or "who is it". Thus, in addition to "faceness" decisions our knowledge of configural coding of faces is tied largely to one facial dimension, facial identity.

Research is slowly beginning to address this imbalance, with some recent work exploring the role of configural processing in facial expression recognition.

CONFIGURAL ENCODING OF FACIAL EXPRESSIONS

One might imagine that salient facial features such as a smiling mouth, or raised eyebrows would imply that facial expressions are processed in a feature-based manner. However, Calder, Young, Keane, and Dean (2000) have demonstrated that configural cues play a significant role (see also White, 2000). This was shown using composite paradigm, a procedure that was initially devised to address configural coding of facial identity (Young, Hellawell, & Hay, 1987). For the facial expression version, composite expressions were produced by aligning the top half of one expression (e.g., anger) with the bottom half of another (e.g., happiness) posed by the same individual (Calder et al., 2000). When asked to identify the expression shown in one half, participants' RTs were slower when the top and bottom halves of the composite were aligned (to form a face) relative to a control condition in which the two halves were misaligned (i.e., the nose in the

top half was aligned with the outline of the face in the bottom half) (Calder et al., 2000). Figure 1 provides an illustration of these stimuli in the form of a single trial from the matching paradigm used in this study.

[FIGURE 1]

In a further experiment, Calder et al. (2000) manipulated the relationship between the identity and expression shown in the two face halves; same identity/different expression, same expression/different identity, and different expression/different identity. When participants were instructed to report the expression shown in one face half, a mismatch in the top and bottom expressions had a significant effect on response times relative to a condition in which the identities in the two halves were different but the expressions were the same. Moreover, a mismatch in both expression *and* identity across the two halves produced no further costs in subjects' ability to report expression relative to a mismatch in expression alone; likewise, the corresponding effect for facial identity was observed when subjects were asked to report the identity of one face half. These results demonstrate that distinct configural cues underlie the perception of facial identity and facial expression, and that subjects can attend selectively to these cues.

LOCUS OF THE COMPOSITE EFFECT

Maurer et al. (2002) discuss that three forms of configural information have been identified: (1) Sensitivity to first-order spatial relations (i.e., an appreciation of the basic physical structure that makes a face a face—i.e., two eyes above a nose above a mouth, etc.), (2) holistic processing (i.e., perception of facial features as a single gestalt representation), and (3) second-order spatial relations (i.e., the spatial relationships between individual features).

Studies addressing the configural encoding of faces have generally concentrated on holistic and second-order relations processing, and Carey and Diamond's (1994) definitions of these processes are helpful in this respect (see also Rakover, 2002). These authors suggested that holistic processing corresponds to the perception of a face as a whole, as opposed to as individual features (e.g., nose, mouth) or even clusters of certain features (e.g., the top half of the face). Moreover, they proposed that holistic encoding is perhaps best defined in terms of the idea that representations of whole faces are “more accessible” than representations of separate facial features. In other words, our natural tendency is to encode each face as a single object, as opposed to a constellation of individual features; although Carey and Diamond are careful to point out that individual feature coding is also possible.

Second-order relations features were defined as the fine-grained interrelationships that exist between different facial features, such as the distance between the eyes, the relationship between eyes and nose, or more complex relationships, such as those that exist among the eyes, nose, and the outline of the face, and so on. In addition, Carey and Diamond suggested that proficiency in coding these second-order relations is a significant factor in making us experts in face perception, with improvements in performance being found up until the teenage years.

Carey and Diamond (1994) put forward these ideas in a developmental study of face processing that investigated the effects of the composite paradigm and stimulus inversion (i.e., 180° orientation) in three groups of participants of different ages (6-year-olds, 10-year-olds, and adults). All three groups showed a significant composite effect that was not significantly modulated by age. In addition, inverting the stimuli significantly reduced the composite effect for each age group to the same extent (i.e., the Age group \times Condition (composite and noncomposite) \times Inversion interaction was not significant). In contrast, and consistent with other studies (Carey, 1981, 1992; Goldstein & Chance, 1964), the *overall effect of inversion* (i.e., independent of the composite/noncomposite contrast) became more pronounced with increasing age, as reflected by an Inversion \times Age group interaction. On the basis of these results, Carey and Diamond concluded that the well-known impact of stimulus inversion on face perception reflects the disruption of two distinct forms of configural processing. One form that shows an early developmental onset that is tapped principally by the composite paradigm, which they identified as holistic encoding. And a second form that shows increasing proficiency with age, which they identified as second-order relations processing.

Carey and Diamond's (1994) observation that holistic encoding is relatively unaffected by age is supported by a study by Tanaka and colleagues (Tanaka, Kay, Grinnell, Stansfield, & Szechter, 1998) that addressed holistic encoding using a different procedure, the part-whole paradigm. This procedure examines memory for a particular individuals' facial feature (e.g., Bob's nose) when it is presented in isolation or in the whole face. The difference in performance between these conditions serves as an index of holistic processing.

Tanaka et al. (1998) showed that 6-, 8-, and 10-year-olds produced the same pattern as adults (Tanaka & Farah, 1993), in that memory for face parts from upright faces was superior when tested in the whole-face context than in isolation, and was not modulated by age.

Both Carey and Diamond (1994) and Tanaka et al.'s (1998) findings for holistic encoding contrast with the increasing effect of stimulus inversion with age that is found for most face perception tasks (Carey, 1981, 1992; Goldstein & Chance, 1964). As discussed, however, Carey and Diamond attributed the monotonic relationship between age and the size of the inversion effect to increasing proficiency with second-order relations processing. It is therefore important that subsequent research by Mondloch and colleagues (Mondloch, le Grand, & Maurer, 2002) has supported this hypothesis.

Mondloch et al. (2002) investigated children's (6-, 8-, and 10-year-olds) and adults' ability to detect changes in single facial features (e.g., mouth or eyes replaced), face contours (i.e., the same internal features were shown in different face contours), or second-order relations (e.g., eyes moved closer together, or up/down) when faces were presented upright or inverted (Mondloch et al., 2002). Consistent with Carey and Diamond's (1994) hypothesis that sensitivity to second-order spatial relations becomes more pronounced with age, Mondloch and colleagues (Mondloch et al., 2002) found that 10-year-olds and adults showed a more marked inversion effect for detecting changes in second-order relations than for detecting feature or contour changes, whereas 6- and 8-year-olds showed no difference between the inversion effects for the three change conditions. In addition, Mondloch et al. found that the inversion effect for second-order relations increased with age.

Together, then, these developmental studies concur with the idea that holistic encoding (as tapped by the composite and part-whole paradigms) can be distinguished from second-order relations processing (Carey & Diamond, 1994). Of course these studies do not exclude the possibility that second-order relations make some contribution to the composite effect in adult participants. However, there is no support for the idea that proficient second-order relations processing is *necessary* for the composite effect, because although sensitivity to second-order relations shows a monotonic relationship with age, the composite effect does not.

Like Carey and Diamond (1994), Hole and colleagues (Hole, George, & Dunsmore, 1999) have also attributed the composite effect for facial identity to a holistic processing mechanism. Furthermore, these authors demonstrated that the composite effect was resistant to a photographic negative manipulation (i.e., inverting the greyscale values of the pixels), but not, as already discussed, to stimulus inversion (see also Carey & Diamond, 1994; Young et al., 1987). Note that this finding contrasts with the observation that photographic negative impairs perception of second-order relations (Kemp, McManus, & Pigott, 1990). Thus, it seems reasonable to infer that if a configural processing task is resistant to a photographic negative manipulation, then the configural processing is more likely to reflect holistic encoding rather than second-order relations encoding.

Hole and colleagues (Hole, 1994; Hole et al., 1999) used an adapted version of the composite paradigm in which participants were asked to decide whether the top halves of two simultaneously presented composite faces prepared from unfamiliar faces were the same or different (the bottom halves of the faces were always different). The authors argued that their procedure has an advantage over the standard composite paradigm (Carey & Diamond, 1994; Young et al., 1987) in that it removes the need to match half of the face to a stored memory representation, and instead focuses on whether the *perception* of one face half is affected by the information in the other half. By contrast, the original paradigm used composites prepared from familiar faces and the task was to identify (name) one half of the composite face. Hole (1994) found a composite effect for facial identity with his matching paradigm and therefore concluded that the effect operates at the level of perceptual encoding rather than the access to stored representation of a familiar face.

Overall, then, these studies indicate that the composite effect for facial identity operates at the level of holistic encoding, and is disrupted by stimulus inversion but not photographic negative. In addition, they support the idea that holistic encoding and second-order relations processing constitute different forms of configural processing (Carey & Diamond, 1994).

Drawing on the findings of these different studies, our present study investigated whether the composite effect for *facial expression* operates at a similar level of configural encoding to its identity counterpart (i.e., holistic encoding). In light of Calder et al.'s (2000) observation that different types of configural cues underlie identity and expression perception, the idea that composite effects for expression and identity reflect a similar 'holistic' mechanism cannot be regarded as a foregone conclusion.

Our study used Hole's (1994) composite matching procedure. We reasoned that if both identity and expression forms of the composite effect operate at the same perceptual level, then it follows that like its identity counterpart, the composite effect for facial expression should be disrupted by inversion but relatively unaffected by photographic negative. Moreover, drawing on the findings of the different studies we have reviewed it would seem reasonable to infer that if the aforementioned pattern is confirmed, then the composite effect for facial expression reflects holistic encoding.

In Experiment 1 we investigated whether a composite effect for facial expression could be found with an adapted version of Hole's (1994) paradigm in which subjects were asked to indicate whether the top halves of two composite facial expressions were the same or different (the bottom halves of the composites were always different—see below for a description of the stimuli). Images were presented in upright and inverted orientation, and the index of a composite effect was whether the reaction times to make a same/different decisions were significantly slowed in the upright relative to the inverted condition.

Experiment 2 used the same basic paradigm as Experiment 1, with the exception that the inversion condition was replaced by a photographic negative transform. This enabled us to address whether the composite effect for expression was as resistant to a photographic negative manipulation as its identity counterpart. Evidence of resistance would provide support for the idea that both operate at the same level of perceptual processing. By inference from the results of the studies investigating configural coding of facial identity, this level likely reflects holistic encoding of faces.

Finally, Experiment 3 used a more precise analogue of the original composite paradigm (Young et al., 1987) in which subjects were asked to decide whether the top halves of composite and noncomposite facial expressions were same or different. This provided a clear indication of the magnitude of the composite effect for facial expression, and the extent to which it was influenced by photographic negative.

STIMULUS MATERIALS

All of the experiments used the same basic stimulus set, hence we begin by describing the materials. The stimuli were prepared from greyscale pictures of three female models (C, NR, and PF) from Ekman and Friesen's (1976) pictures of facial affect series. Each model posed one example of six facial expressions: Happiness, sadness, anger, fear, disgust, and surprise. Composite faces were prepared by combining the top half of an expression that can be identified from its top half alone (anger, fear, and sadness) with the bottom half of one of the remaining expressions that can be identified from their bottom face halves (happiness, disgust, and surprise) (Bassili, 1979; Calder et al., 2000); the top and bottom halves of each composite were always posed by the same model. This gave nine possible top/bottom combinations (anger/happiness, anger/disgust, anger/surprise, fear/happiness, fear/disgust, fear/surprise, sadness/happiness, sadness/disgust, sadness/surprise); a total of 27 different composite faces (i.e., 3 models \times 9 combinations). Examples of the stimuli are shown in Figure 1.

Experiment 3 used both composite and noncomposite faces. The noncomposites were essentially identical to the composites, except that the top and bottom halves of the face were horizontally misaligned so that the middle of the nose in the top half was aligned with the contour of the bottom half of the face (Figure 1). For each composite image there were two versions of the corresponding noncomposite image—in one the top half was shifted to the right of the bottom half, in the other the top half was shifted to the left.

All experiments included one of two image manipulations—inversion (i.e., 180° orientation, Experiment 1) or photographic negative (Experiments 2 and 3). Details of experimental designs are provided in the method sections for each experiment.

EXPERIMENT 1

Hole (1994) showed that participants were slower to judge whether the top halves of two simultaneously presented composite faces showed the same individual or not when the faces were presented in upright orientation relative to inverted orientation; each composite always comprised different individuals' top and bottom face halves. This concurs with earlier research showing that upright facial identity composites elicit a form of configural processing that is disrupted by inverting the images (Carey & Diamond, 1994; Young et al., 1987). In Experiment 1, we investigated whether a similar pattern would be found for composites comprising different facial expressions in the two face halves posed by the same identity. Hole found that his effect was evident with a stimulus presentation time of 80 ms, but not 1200 ms. Because two different facial expressions posed by the same individual are generally more perceptually similar than two different individuals' faces, we opted to use a longer presentation time (2000 ms).

METHOD

Participants

Fourteen members of the MRC Cognition and Brain Science Unit (CBU) volunteer panel (8 women and 6 men; mean age = 27.1, $SD = 7.0$) took part in Experiment 1. All participants had normal or corrected-to-normal vision.

Stimulus materials

The stimulus set for Experiment 1 comprised the 27 different composite facial expressions described earlier shown in original and inverted format (i.e., oriented by 180°).

Design and procedure

Each trial began with a 500 ms alerting tone followed by a pair of composite stimuli that were presented simultaneously for a maximum of 2 s. The two composites were positioned on the left and right of the screen in staggered vertical positions (see Figure 1); the positioning of the left and right images was counterbalanced across the stimulus set (i.e., on half of the trials the left image was above the right, and vice versa for the other half). The composite images subtended a horizontal visual angle of approximately 4.9 degrees and were separated by a visual angle of approximately 2 degrees. The participants' task was to decide whether the top halves of the two faces were the same or different by pressing one of two labelled keys as quickly as possible. A response terminated the presentation of one stimulus pair and initiated the next after an interval of 1.5 s. On half of the trials the top halves of the two composites were the same (same-top composite pairs), and for the other half they were different (different-top composite pairs). For both same and different trial formats, the bottom halves of the composite pairs were different. For each trial, the pairs of composite faces were prepared from pictures of the same identity (model).

Two repeated measures factors were investigated: (1) Top-half similarity (same-top composite pairs and different-top composite pairs), and (2) stimulus orientation (upright and inverted). This gave four different trial formats.

- Upright same-top composites: pairs of composite faces with same-top halves presented in upright orientation.
- Upright different-top composites: Pairs of composite faces with different-top halves presented in upright orientation.
- Inverted same-top composites: As described in 1, except that both composites were rotated by 180°.
- Inverted different-top composites: As described in 2, except that both composites were rotated by 180°.

Because there were twice as many possible combinations of different-top composite pairs (54) as same-top composite pairs (27), different-top pairs were counterbalanced between the two stimulus diagonal position formats (left image above right image and vice versa; see Figure 1), whereas all possible combinations of same-top pairs were shown in both diagonal position formats. Each participant completed two blocks of trials—in one block, the composites were shown in their original format; in the other block they were oriented by 180°. Order of block presentation was counterbalanced across participants. Each block contained 108 experimental trials preceded by 10 practice trials selected at random from the experimental trials. Note that in both upright and inverted conditions, the participants' task was the same—to decide whether the top halves of the two faces were same or different; however, the top half of the composites in the inverted condition was effectively the bottom half of the image.

Prior to the start of each block (upright or inverted), the participants were shown a printed example of a trial to ensure that they were aware of the trial format.

RESULTS

Participants' mean correct reaction times and untransformed error proportions are summarized in Figure 2.

[FIGURE 2]

Reaction times

Reaction times greater than 2500 ms (2.96% of data) were excluded from the analyses. A two-factor ANOVA investigating top-half similarity (same-top composites and different-top composites; repeated measure) and orientation (upright and inverted; repeated measure) was performed on the remaining mean correct RTs. There was a significant main effect for orientation, $F(1, 13) = 11.64, p < .005$, qualified by a significant interaction between top-half similarity and orientation, $F(1, 13) = 6.14, p < .05$. Simple effects of the interaction showed a significant effect of orientation for same-top composites, $F(1, 13) = 14.84, p < .005$ (upright RTs > inverted RTs), but not different-top composites ($F < 1$). The main effect of top-half similarity did not reach statistical significance, $F(1, 13) = 2.5, p > .1$.

Error rates

Participants' error proportions were arcsin transformed and submitted to a two-factor repeated measures ANOVA investigating the same two factors defined above (i.e., top-half similarity and orientation). There was no significant effect of orientation ($F < 1$), but a significant main effect of top-half similarity, $F(1, 13) = 17.88, p < .001$ (different-top errors > same-top errors), and a significant interaction between orientation and top-half similarity, $F(1, 13) = 15.78, p < .001$. Simple effects of the interaction showed a significant effect of orientation for both the same-top condition, $F(1, 13) = 6.80, p < .05$, and different-top condition, $F(1, 13) = 15.71, p < .05$. However, whereas the same-top composites generated fewer errors in the inverted condition, the opposite was found for the different-top composites.

DISCUSSION

The analysis of RT data showed that inverting the composites led to faster correct "same-top" decisions relative to when composites were shown in upright orientation. Inversion had no significant effect on RTs to make different decisions; this pattern was confirmed by an interaction between top-half similarity and orientation. The analysis of error proportions showed that inverting the composites led to increased accuracy for same-top decisions relative to when composites were shown in upright orientation; however, the opposite effect was observed for different-top decisions.

Finding a composite effect for facial expression with Hole's (1994) matching paradigm demonstrates that the composite effect for facial expression is not dependent on a particular task or design (i.e., the identification paradigm used by Calder et al., 2000). In addition, it is consistent with the idea that the effect reflects a configural processing mechanism at the level of perceptual encoding rather than matching expressions to their stored representations. The results of Experiment 1 are also consistent with Calder et al.'s observation that that inversion disrupts the composite effect for facial expression when an identification task is used (e.g., classifying the emotion in the top-half of the face as anger, sad, or fear).

Two specific aspects of the data are worthy of further discussion. The first concerns the fact that the composite effect was only evident for the same-top condition. Second, the error data showed opposite effects for same-top (upright errors > inverted errors) and different-top decisions (upright errors < inverted errors).

Turning first to the findings for the same-top condition, it is relevant that researchers have discussed that the composite effect for facial identity operates primarily on holistic encoding (Carey & Diamond, 1994; Hole et al., 1999; Maurer et al., 2002), and that inverting the stimulus disrupts this stage of processing. Thus, participants take longer to decide that the top halves of a pair of *upright* composite faces are the same when they comprise the same top halves and different bottom halves because they have to suppress their natural tendency to encode each composite as a whole face, and processing the faces in this way influences participants' perception of faces' top halves. Inverting the images, however, disrupts holistic encoding and facilitates the participants' ability to focus on the top half of the images alone. This was reflected as decreased RTs and errors in the inverted same-top condition.

For different-top composites, however, the same interpretation does not follow because a decision based on whole face matching would lead to an identical response to a decision based on processing the top halves of the image alone (i.e., "different"). However, although this interpretation is consistent with the

absence of an RT effect for upright and inverted different-top composites, the error data actually showed an increase in error proportions when the stimuli were inverted. One possibility, then, is that the differences between the top halves are enhanced under optimal conditions for whole face encoding (i.e., when the images are shown in upright orientation) because the different bottom halves of the two composite faces augment the perceived differences in the top halves. In other words, as in the case of the same-top composites, the different bottom halves make the top's of the different-top composites appear more different than would be the case if the top halves were presented in isolation.

Alternatively, it is well known that inversion impairs other aspects of face processing in addition to holistic encoding, such as second-order configural encoding (Maurer et al., 2002; Mondloch et al., 2002). Hence, the increased errors for the inverted different-top decisions could equally reflect increased difficulty in detecting differences between the stimuli due to a disruption of second-order configural processing, that occurs in addition to the disruption of holistic encoding. Because these explanations cannot be distinguished from Experiment 1, we defer full discussion of this pattern for different-top composites until the section for Experiment 3, which sheds further light on this issue.

In summary, the main finding of Experiment 1 was that inversion disrupts the composite effect for facial expression. This concurs with the results of Hole et al. (1994) who used a similar paradigm to investigate the composite effect for facial identity. Given the similarity between Hole et al.'s (1994) findings and our own, it follows that the composite effects for identity and expression operate at a stage that is common to the perception of both facial identity and facial expression. A suitable level is what Bruce and Young (1986) characterise as structural encoding in their functional model of face processing.

EXPERIMENT 2

As we discussed, Hole (1994) found that the composite effect for facial identity is resistant to a photographic negative manipulation, whereas other work has shown that perception of second-order relations is not (Mondloch et al., 2002). Hence, we reasoned that further support for a common basis for identity and expression composite effects could be found by demonstrating that the composite effect for facial expression is similarly resistant to photographic negative. Thus, in contrast to Experiment 1, same-top decisions should not be facilitated when the composite pairs are shown in negative. Rather, the negative manipulation should disrupt second-order configural information producing an overall impairment in participants' ability to detect that the tops of the composites pairs are same/different.

METHOD

Participants

Fifteen members of the MRC CBU volunteer panel (3 men, 12 women; mean age = 23.2, $SD = 6.3$) took part in Experiment 2. None had participated in Experiment 1 and all had normal or corrected-to-normal vision.

Materials, design, and procedure

The experimental format was very similar to Experiment 1. The only difference was that the stimulus inversion manipulation was replaced with a photographic negative manipulation. This gave two repeated measures factors, top-half similarity (same-top and different-top composite pairs) and photographic format (positive and negative photographic format). The four different trial formats can be summarised as follows.

- Upright same-top composites: pairs of composite faces with same-top halves.
- Upright different-top composites: pairs of composite faces with different-top halves.
- Photographic negative same-top composites: as described in 1, except that greyscale values of the pixels in the composite images were inverted (i.e., photographic negative).
- Photographic negative different-top composites: as described in 2, except that greyscale values of the pixels in the composite images were inverted (i.e., photographic negative).

The photographic negative factor was blocked. Half of the participants completed the photographic positive block first, for the other half this ordering was reversed. In all other respects (i.e., stimulus location, task, presentation times, etc.), Experiment 2 was identical to Experiment 1.

RESULTS

Participants' mean correct reaction times and untransformed error proportions are summarized in Figure 3.

[FIGURE 3]

Reaction times

One participant was dropped from the analysis because they showed an unusual bias toward responding “same” and performed at chance overall (47% correct). Correct RTs greater than 2500 ms (2.53% of data) from the remaining 14 participants were excluded from the analysis. The remaining mean correct RTs were submitted to a two-factor ANOVA investigating top-half similarity (same-top and different-top composite pairs; repeated measure) and photographic format (original and photographic negative; repeated measure). The only significant effect was photographic format, $F(1, 13) = 7.22, p < .02$, reflecting longer RTs to make same-different decisions for images shown in photographic negative; all other $ps > .1$.

Error rates

Participants' error proportions were arcsin transformed and submitted to a two-factor repeated measures ANOVA investigating the same two factors listed above (i.e., top-half similarity and photographic format). There was a marginal nonsignificant effect of photographic format, $F(1, 13) = 3.81, p = .073$, reflecting a trend towards increased errors for same-different decisions when the images were shown in photographic negative; all other $ps > .1$.

COMBINED ANALYSIS OF EXPERIMENTS 1 AND 2

Reaction times

The patterns of RTs were different for two image manipulations used in Experiments 1 and 2—inverting the composites produced speeded “same-top” decisions in Experiment 1, whereas the photographic negative transform produced the opposite effect. To investigate whether this difference was statistically robust, correct RTs from Experiments 1 and 2 were submitted to a single ANOVA examining three factors: (1) Experiment (inversion in Experiment 1 and photographic negative in Experiment 2; between subjects); (2) stimulus format (original format and manipulated format, i.e., inversion or photographic negative; repeated measure), (3) and top-half similarity (same and different top half; repeated measure). The results showed a significant interaction between experiment and stimulus format, $F(1, 26) = 16.50, p < .001$, confirming that the effects of the two image manipulations (i.e., inversion and photographic negative) were qualitatively different. In other words, inversion and photographic negative had different effects on the composite effect for facial expression. Neither the main effect of experiment, $F(1, 26) = 2.00, p > .1$, nor any other interactions with the experiment factor reached statistical significance ($Fs < 1$).

Error rates

Arcsin transformed error proportions for Experiments 1 and 2 were submitted to a similar ANOVA with experiment (inversion in Experiment 1 and photographic negative in Experiment 2) as a between-subjects factor; stimulus format and top-half similarity were also included as repeated measures factors. The experiment factor did not interact with any of the repeated measure factors (all $Fs < 1.8, ps > .2$). However, the main effect of experiment was significant, $F(1, 26) = 8.50, p < .01$, reflecting an overall increase in error rates in Experiment 2 (photographic negative manipulation) relative to Experiment 1 (inversion manipulation).

DISCUSSION

Experiment 2 demonstrated further similarities between the composite effect for facial expression and facial identity, in the sense that photographic negative did not disrupt the effect by producing faster RTs in the negative condition. In fact, the only significant effect of photographic negative was an overall deleterious effect on participants' ability to perform the same-different matching task. In addition, there was a nonsignificant trend towards a greater cost for different-top rather than same-top decisions in the negative condition (expressed in both RT and error measures). Given that the effect of inversion was more evident for same-top than different-top decisions in Experiment 1, the trend towards a greater impact of photographic negative on different-top than same-top pairs concurs with the idea that photographic negative and inversion affect our composite expression matching task in different ways.

In the introduction we discussed the result of developmental studies that suggest that the composite effect for facial identity reflects holistic rather than second-order relations encoding. In addition, the observation that the identity composite effect is disrupted by inversion but not photographic negative has been

interpreted as support for the idea that photographic negative affects holistic encoding. Hence, by inference, the observation that the composite effect for expression shows the same effects with inversion and photographic negative manipulations is consistent with the idea that this effect also reflects holistic encoding.

Although Experiment 2 is consistent with Hole et al.'s (1999) observation that the composite effect for facial identity is similarly not affected by photographic negative transformation, it is not clear from Experiment 2 alone whether photographic negative had no impact on the composite effect or actually enhanced it. The absence of an additional baseline condition means that all we can conclude is that photographic negative does not produce the same pattern as inversion (i.e., it does not *reduce* the composite effect for facial expression). Whether the overall increase in RTs in the photographic negative condition reflects an enhanced composite effect or a disruption to another aspect of face perception (e.g., second-order configural encoding) is not clear. Consequently, we addressed this issue in Experiment 3 by including manipulations of both holistic encoding (i.e., composite and noncomposite images) and photographic format. This final experiment also gave us an opportunity to use a “same-different matching” task that was more similar to the original composite paradigm (Calder et al., 2000; Young et al., 1987) because it used both composite *and* noncomposite facial expressions shown in original and photographic negative format.

From Experiments 1 and 2 and our previous findings (Calder et al., 2000) we predicted that it should take longer to match the top halves of same-top expression composites relative to same-top expression noncomposites when the images are shown in their original (photographic positive) format. Moreover, if photographic negative has no effect on the composite effect for facial expressions, then the same pattern should be found when the images are shown in photographic negative. Alternatively, if the photographic negative manipulation actually enhances the composite effect for same-top composites, then this will be reflected in the form of a larger discrepancy between composite and noncomposite stimuli in the negative relative to original photographic conditions.

EXPERIMENT 3

METHOD

Participants

Sixteen members of the MRC Cognition and Brain Sciences Unit participant panel (13 women and 3 men; mean age = 22.8, *SD* = 3.7) were paid for participating in this experiment. All had normal or corrected-to-normal vision.

Materials

The experiment included both composites and noncomposite stimuli shown in both original and photographic negative formats. The composites were the same 27 stimuli used in Experiments 1 and 2; the noncomposites were prepared by misaligning the top and bottom halves of the same composite faces. For each composite there were two versions of the corresponding noncomposite—in one the top half was shifted to the right of the bottom half, in the other the top half was shifted to the left. To ensure equal numbers of composite and noncomposite stimuli the two types of noncomposite stimuli (top face half shifted to left or right of bottom face half) were counterbalanced across the trials.

Design and procedure

On each trial, two composite or noncomposite stimuli were presented simultaneously on the left and right of the visual display. As for Experiments 1 and 2, the position of the images was staggered (i.e., left image higher than right image, and vice versa) (Figure 1); this positioning was counterbalanced across the stimulus set. The participants' task was to indicate whether the top halves of the two faces were identical (same response) or different (different response) by pressing one of two labelled keys. The trials consisted of equal numbers of same-top and different-top composite and noncomposite face pairs randomly intermixed. Each participant completed two blocks of trials. In one block the composite and noncomposite stimuli were shown in their original format; in a second block they were shown in photographic negative. Order of block presentation was counterbalanced across participants.

In summary, three factors were investigated: Stimulus type (composite and noncomposite; repeated measure), top-half similarity (same-top halves and different-top halves; repeated measure), and photographic format (original and negative; repeated measure). This gave eight different trial formats—the

four formats listed below shown in original photographic format in one block, and the same four formats shown in photographic negative in a second block.

- Same-top composites: Pairs of composite faces with same-top halves.
- Same-top noncomposites: Pairs of noncomposite faces same-top halves.
- Different-top composites: Pairs of composite faces with different-top halves.
- Different-top noncomposites: Pairs of noncomposite faces with different-top halves.

On all trials the bottom halves of the composite or noncomposite pairs were different.

The composite images subtended a horizontal visual angle of approximately 4.9 degrees, while the noncomposite images subtended a visual angle of approximately 7.5 degrees. The top halves of both image types were separated by a visual angle of approximately 2 degrees. Note, the composite and noncomposite images had different visual angles and the top half of the noncomposite images had different positions depending on whether the top half was shifted to the left or right of its bottom half. To allow for this fact, half of the composites were presented in one of two positions: (1) The position corresponding to the top half of the left-of-bottom-half noncomposites, and (2) the position corresponding to the top half of the right-of-bottom-half noncomposites (Figure 1). This positioning was counterbalanced across the stimulus set. All other aspects of the experimental design (i.e., presentation times, etc.) were identical to Experiments 1 and 2.

RESULTS

Participants' mean correct RTs and untransformed error proportions are displayed in Figure 4.

[FIGURE 4]

Reaction times

Reaction times greater than 2500 ms (0.14% of data) were excluded from the subsequent analyses.

Participants' mean correct reaction times were submitted to a three-factor ANOVA examining photographic format (original and photographic negative; repeated measure), stimulus type (composite and noncomposite; repeated measure), and top-half similarity (same-top halves and different-top halves; repeated measure). The results showed a significant effect of stimulus type, $F(1, 15) = 8.08, p < .02$ (composite RTs > noncomposite RTs), qualified by a significant interaction between stimulus type and top-half similarity, $F(1, 15) = 6.20, p < .05$. Simple effects analyses showed a significant effect of stimulus type for same-top expression images (composite RTs > noncomposite RTs), $F(1, 15) = 8.64, p < .01$, but not for different-top expression images ($F < 1$). The three-factor ANOVA also showed a significant interaction between photographic format and top-half similarity, $F(1, 15) = 13.33, p < .005$, reflecting an overall increase in RTs to make correct "different-top" responses in the negative relative to original photographic format conditions, $F(1, 15) = 7.92, p < .02$, but no corresponding difference in RTs to make correct "same-top" decisions in the two formats, $F < 1$. No other effects reached statistical significance [main effect of photographic format, $F(1, 15) = 2.21, p > .1$, all other F s < 1].

In summary, the analysis of RTs showed the predicted pattern—a composite effect for correct same-top decisions (composite RTs > noncomposite RTs), but not correct different-top decisions. As also predicted, the composite effect was not significantly modulated by photographic format.

Error rates

Participants' error proportions were arcsin transformed and submitted to a three-factor repeated measures ANOVA investigating the same three factors outlined above (photographic format, stimulus type, and top-half similarity). There was no main effect of stimulus type ($F < 1$) but a significant interaction between stimulus type and top-half similarity, $F(1, 15) = 14.12, p < .005$. Simple effects analyses of this interaction showed *opposite effects* of stimulus type for the same-top pairs, $F(1, 11) = 6.72, p < .05$ (composite errors > noncomposite errors) and different-top pairs $F(1, 11) = 6.01, p < .05$ (composite errors < noncomposite errors). There were also significant main effects of top-half similarity, $F(1, 15) = 84.00, p < .0001$ (errors for different-top pairs > errors for same-top pairs), and photographic format, $F(1, 15) = 16.41, p < .0001$ (errors for negative format images > errors for original format images). However, these effects were qualified by a significant interaction between photographic format and top-half similarity, $F(1, 15) = 13.56, p < .005$, reflecting that while the negative format produced increased errors for the different-top images,

$F(1, 15) = 23.46, p < .0001$, a comparable significant increase was not observed for the same-top image condition, $F(1, 15) = 1.99, p > .1$. No other effects reached statistical significance, including the three-way interaction between photographic format, stimulus type, and top-half similarity ($F_s < 1$).

In summary, as predicted, a composite effect was evident for same-top pairs (composite errors > noncomposite errors). In addition, in line with Experiment 1, a disruption to holistic encoding (inversion in Experiment 1, noncomposites in Experiment 3) produced the opposite pattern (composite errors < noncomposite errors) for different-top pairs. Critically, the magnitude of the composite effect was not modulated significantly by converting the images to photographic negative. A detrimental effect of photographic negative was observed on participants' overall ability to make "different" decisions (negative errors > original image error); however, no corresponding increase in errors was found for the "same" decisions.

DISCUSSION

Our results showed faster correct RTs for same-top (but not different-top) noncomposites relative to composites in both original *and* photographic negative conditions. Furthermore, the photographic negative transform did not significantly modulate the size of the composite effect for facial expression (i.e., composite relative to noncomposite), and the only significant effect of photographic negative was to produce an overall increase in correct RTs and errors for different-top, but not same-top, decisions. In other words, photographic negative makes it harder to detect differences between two images, generating a tendency to respond that the images are the same. This is consistent with previous research showing that the size of the composite effect for facial identity is not affected by photographic negative, whereas the detection of changes in second-order spatial relations is (Kemp et al., 1990). The only discrepancy between the RT and error data was that noncomposites produced a reduction in errors for same-top images and an increase in errors for different-top images for both photographic formats; recall that a similar effect was found in Experiment 1 where inversion was used as to disrupt holistic encoding.

The results of Experiment 3, then, concur with the findings of Experiments 1 and 2. For the same-top trials a disruption to holistic encoding (as a result of inversion or the use of noncomposites) enables participants to concentrate on the top-half of the image alone and minimise the conflicting effect produced by encoding the two face halves as a whole (i.e., holistically). For the different-top conditions, a decision based on processing the top-halves of the faces in isolation would lead to an identical response ("different") as one based on coding both top and bottom halves combined. However, it is conceivable that processing the two halves in a holistic manner leads to a more marked tendency to respond "different", because differences between the tops halves of the composites are enhanced by a holistic encoding mechanism that forces the subject to take differences between the bottom halves of the image into account. Just why this effect should be reflected in errors and not RTs is not entirely clear, but consistent with Experiment 1. One possibility is that it may reflect a threshold effect that is only evident for the more difficult trials (i.e., the difficult different-top trials produce mostly errors, meaning that this would not be picked up by an RT analysis based on a correct responses only). In other words, the more difficult different-top trials may simply produce an incorrect response when shown in noncomposite (or inverted) form, rather than a graded overall increase in RTs for all noncomposite different-top trials. Finally, it is worth noting that consistent with Experiment 2, both RTs and error rates showed an overall increase for different-top but not same-top decisions in the negative photographic format condition.

GENERAL DISCUSSION

Our results can be summarised as follows.

- In line with the findings for facial identity, our results show that a composite effect for facial expression is found with a face matching task for which subjects are asked to indicate whether the top-halves of two composite (or noncomposite) faces are the same or different (Experiments 1, 2, and 3). This is consistent with the idea that the composite effect for facial expression reflects a disruption to perceptual encoding; a suitable level is what Bruce and Young (1986) identified as structural encoding.
- The composite effect for facial expression is disrupted by inverting the stimuli (Experiment 1; see also Calder et al., 2000), but not by converting the images to photographic negative (Experiments

- 2 and 3). This concurs with the observations for the composite effect for facial identity, and provides further support for the idea that both effects reflect a common perceptual mechanism.
- Finally, our results are consistent with the idea that the composite effect for facial expression reflects holistic encoding (which is disrupted primarily by inversion) rather than second-order configural encoding (which is disrupted by both inversion and photographic negative) (Carey & Diamond, 1994; Hole, 1994; Maurer et al., 2002).

Our findings concur with our previous study showing a composite effect for facial expression when the task was to identify the expression shown in the top (or bottom half) of a face as one of three emotions (e.g., anger, sadness, or fear) (Calder et al., 2000). This demonstrates that the composite effect for expression is robust occurring across different paradigms (identification and same-different matching tasks) and for different manipulations of holistic encoding (inversion and noncomposite conditions).

Our results show distinct similarities and differences with those of Hole and colleagues (Hole, 1994; Hole et al., 1999) who used a similar matching procedure to explore the composite effect for facial identity. A first obvious difference is that Hole (1994) found the composite effect for identity was only present for brief (80 ms) but not longer (1200 ms) presentation times. In contrast, we found a robust effect with stimulus presentation times of 2000 ms. As we discussed, this may be because the perceptual differences between two different identities (with similar neutral expressions, as used by Hole) are more salient than the differences between pictures of the same identity posing the same or different expressions, used in our own study.

In line with our own findings, Hole (1994) found that inverting the images generated shorter correct RTs for same-top facial identity composites relative to when the images presented in upright orientation (i.e., a reduced composite effect). However, Hole also found the same pattern for *different-top* composites, whereas we found that this pattern was restricted to same-top composites. In addition, although Hole found no significant effect of inversion on error rates for the same experiment, the trend in his data was comparable to our own (i.e., upright errors > inverted errors for same-top composite, and upright errors < inverted errors for different-top composites).

In a subsequent experiment Hole et al. (1999) investigated the effects of photographic negative *and* inversion on participants' ability to perform the same composite matching task for facial identity. Their results replicated the effects of inversion on the composite effect, i.e., reduced RTs to detect that the tops of two composite faces showed the same or different identities. In addition, consistent with our own findings, Hole et al. (1999) found that their composite effect was not modulated by photographic negative, i.e., photographic format (positive and negative) did not interact with orientation (upright and inverted). However, whereas we found that photographic negative produced an overall slowing in RTs, Hole et al. found that photographic negative *facilitated* participants' RTs to decide whether the tops of the composite faces showed the same or different identities.

Hole et al. (1999) attributed the puzzling beneficial effect of photographic negative to stimulus artefacts, such as increased salience of hair cues in the two different identities, a factor did not apply to our stimuli because the identities in the composite face pairs were always the same. This was addressed in a further experiment in which Hole et al. used pairs of composite faces posing different views (full face and three-quarters). An analysis of RT data for same-top composites showed an overall facilitation in RTs with inversion that was not modulated by photographic negative. In contrast, RTs to different-top composites showed no significant effect of orientation or photographic format. An analysis of the accuracy data for the same experiment revealed that same-top decisions were improved by inverting the stimuli, whereas photographic negative had no significant effect. For different-top composites, both inversion *and* photographic negative produced an increase in errors. Overall, then, Hole et al.'s findings are largely consistent with our own. The main difference being that we found that photographic negative produced an overall cost for RT and accuracy measures, but particularly for different-top composites.

Given the similarities between Hole and colleagues (Hole, 1994; Hole et al., 1999) results and our own, it seems likely the composite effects for facial identity and facial expression operate at the same level of perceptual encoding. In other words, a level that is common to identity and expression perception. In Bruce and Young's (1986) functional model of face recognition separate parallel routes coding different facial dimensions (i.e., identity, expression, lipspeech) emerge from a common front-end system that the authors referred to "structural encoding". Hence, a suitable locus for the composite effects for identity and

expression is the structural encoding mechanism. But what form of face processing might structural encoding involve?

In the introduction we presented data from a series of developmental studies which together show that different forms of configural encoding exist. Two forms that have been most often studied are second-order relations and holistic encoding, and there is evidence that the holistic encoding makes a disproportionately greater contribution to the composite effect than second-order spatial relations. Moreover, the fact that the composite effect is disrupted by inversion, but not a photographic negative manipulation, while the detection of changes in second-order spatial relations are disrupted by both (Carey & Diamond, 1994; Hole et al., 1999; Mondloch et al., 2002; Young et al., 1987), has led some to infer that holistic processing is largely resistant to photographic negative (Hole et al., 1999; Maurer et al., 2002). Our own results do not conflict with this interpretation of data. Therefore, we conclude that the composite effects for both facial identity and facial expression reflect holistic encoding of faces, and that holistic encoding comprises one of the mechanisms operating within the structural encoding system.

As discussed in the introduction, our earlier study showed that the configural cues underlying identity and expression processing showed qualitative differences (Calder et al., 2000). This finding also concurs with our work on the image-based analyses of facial information (Calder, Burton, Miller, Young, & Akamatsu, 2001), which showed that a single multidimensional framework, derived from a principal component analysis (PCA) of the greyscale pixel values of facial images, contained separate principal components (dimensions) for coding each of facial identity and facial expression. In other words, although a single multidimensional model was sufficient for coding these different facial cues, dissociations between identity and expression were implicit within the context of this single perceptual framework. Thus, in line with our previous work on PCA and configural processing, and the results of this present study, we propose that configural cues to identity and expression are encoded by the same perceptual system, but with subtle differences in configural cues relevant to identity and expression. It is also worth noting that our recent collaborative work with Cottrell and Branson (Cottrell, Branson, & Calder, 2002) has shown that a system based on PCA can provide an effective simulation of the composite effects for identity and expression; including the observation that subjects can selectively attend to the configural cues relating to a face's identity or its expression (Calder et al., 2000).

Finally, one additional aspect of our data that requires discussion is that all of our experiments showed a similar relationship between RT and accuracy data for same-top and different-top images. That being that different-top images produced shorter RTs and increased error rates relative to their same-top counterparts. We believe that this reflects the nature of the experimental task, in that for a same-different matching task of this sort a sensible strategy would be to examine the two top face halves for differences—if any are found, then respond “different”, if none are found by the point at which the images are removed from the screen, then respond “same”. We do not consider that this is a problem for our interpretation because any main effect of top-half similarity (same-top versus different-top) was incidental to the effects we have discussed (which were expressed as interactions). In other words, main effects were not present for both RTs and error rates in any single experiment, meaning that the analyses did not show evidence of speed-accuracy tradeoff.

In summary, our study concurs with the idea that composite effect for facial expression reflects a holistic encoding mechanism. In addition, we propose that the composite effects for expression and identity operate at a similar level of perceptual encoding. However, given that our other research has shown that the configural cues for identity and expression show certain qualitative differences, we would also want to add that subtle differences in the configural cues for identity and expression will be abstracted at this common level. Although a detailed analysis of these different configural cues was not the focus of this present study, it is important that this should be addressed in future work.

REFERENCES

1. 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:(11) , pp. 2049-2058.
3. Calder, A. J. , Burton, A. M. , Miller, P. , Young, A. W. and Akamatsu, S. ((2001)) A principal component analysis of facial expressions. *Vision Research* 41:(9) , pp. 1179-1208.
4. Calder, A. J. , Young, A. W. , Keane, J. and Dean, M. ((2000)) Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance* 26:(2) , pp. 527-551.
5. Carey, S. ((1981)) The development of face perception. *Perceiving and remembering faces* — In G. Davies, H. Ellis, & J. Shepard (Eds.); New York: Academic Press
6. Carey, S. ((1992)) Becoming a face expert. *Philosophical Transactions of the Royal Society, London B335* , pp. 95-103.
7. Carey, S. and Diamond, R. ((1994)) Are faces perceived as configurations more by adults than by children?. *Visual Cognition* 1:(2/3) , pp. 253-274.
8. Cottrell, G. W. , Branson, K. M. and Calder, A. J. ((2002)) Do expression and identity need separate representations?. — Paper presented at the 24th Annual Meeting of the Cognitive Science Society, Fairfax, VA, USA
9. Ekman, P. and Friesen, W. V. ((1976)) *Pictures of facial affect*. — Palo Alto, CA: Consulting Psychologists Press
10. Goldstein, A. G. and Chance, J. F. ((1964)) Recognition of children's faces. *Child Development* 35 , pp. 129-136.
11. Hole, G. J. ((1994)) Configurational factors in the perception of unfamiliar faces. *Perception* 23 , pp. 65-74.
12. Hole, G. J. , George, P. A. and Dunsmore, V. ((1999)) Evidence for holistic processing of faces viewed as photographic negative. *Perception* 28 , pp. 341-259.
13. Kemp, R. I. , McManus, C. and Pigott, T. ((1990)) Sensitivity to the displacement of facial features in negative and inverted images. *Perception* 19 , pp. 531-543.
14. Maurer, D. , Grand, R. le and Mondloch, C. J. ((2002)) The many faces of configural processing. *Trends in Cognitive Sciences* 6 , pp. 255-260.
15. Mondloch, C. J. , Grand, R. le and Maurer, D. ((2002)) Configural face processing develops more slowly than featural face processing. *Perception* 31 , pp. 553-566.
16. Rakover, S. S. ((2002)) Featural and configurational information in faces: A conceptual and empirical analysis. *British Journal of Psychology* 93 , pp. 1-30.
17. Tanaka, J. W. and Farah, M. J. ((1993)) Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology* 46A:(2) , pp. 225-245.
18. Tanaka, J. W. , Kay, J. B. , Grinnell, E. , Stansfield, B. and Szechter, L. ((1998)) Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition* 5:(4) , pp. 479-496.
19. White, M. ((2000)) Parts and wholes in expression recognition. *Cognition and Emotion* 4:(1) , pp. 39-60.
20. Young, A. W. , Hellawell, D. and Hay, D. C. ((1987)) Configurational information in face perception. *Perception* 16 , pp. 747-759.

FIGURES

Figure 1. Examples of the images used in Experiments 1 to 3 and their relative positions on the monitor. The left-hand display shows an example of two composite faces with the same top-half expressions (fear) and different bottom-half expressions (left bottom half = surprise, right bottom half = happy) (i.e., an example of a “same-top” trial). The correct response for this trial is “same”. The right-hand display shows the corresponding noncomposite trial from Experiment 3 in which the same images are shown but with the top and bottom face halves misaligned. In all experiments the image pairs were shown in staggered vertical positions with the right composite (or noncomposite) above the left for half of the trials and vice versa for the other half.



Figure 2. Data from Experiment 1. The left graph shows participants' mean correct reaction times (RTs) to identify the expression displayed in the top halves of the composite and noncomposite stimuli. The right graph shows participants' mean error proportions from the same experiment. Standard error bars for each data point are also shown. Same = same-top composites; diff = different-top composites.

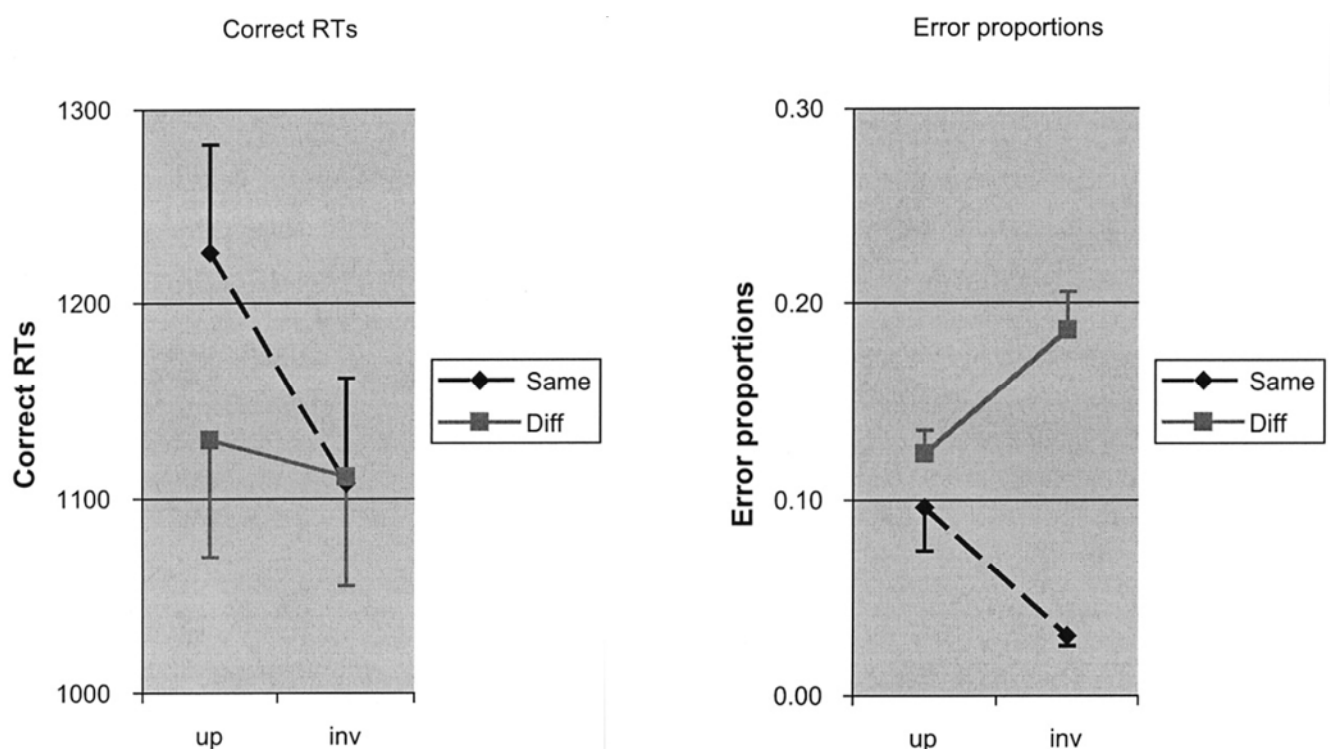


Figure 3. Data from Experiment 2. The left graph shows participants' mean correct reaction times (RTs) to identify the expression displayed in the top halves of the composite and noncomposite images presented in original and photographic negative formats. The right graph shows participants' mean error proportions from the same experiment. Standard error bars for each data point are also shown. Same = same-top composites; diff = different-top composites.

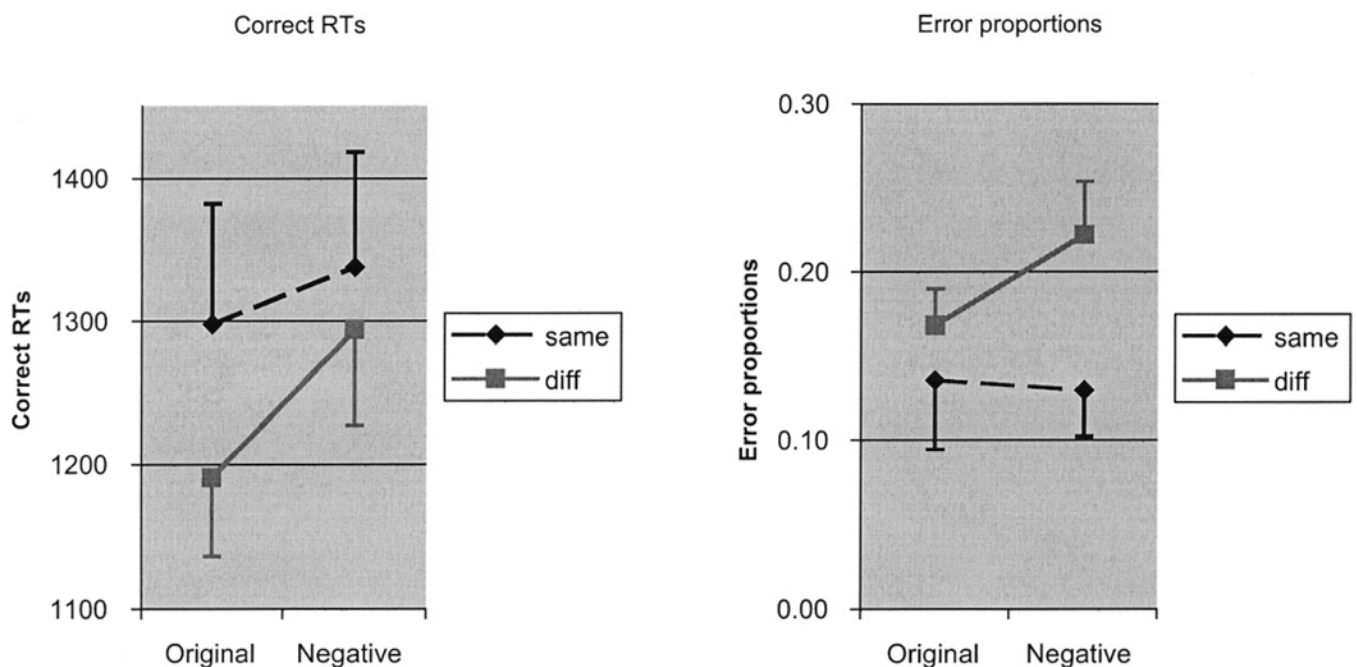


Figure 4. Data from Experiment 3. The top graphs show participants' mean correct reaction times (RTs) to identify the expression displayed in the top halves of the composite and noncomposite images presented in original and negative photographic formats. The bottom graphs show participants' mean error proportions from the same experiment. Standard error bars for each data point are also shown. Same = same-top composites; diff = different-top composites.

