

Asymmetric interference between sex and emotion in face perception

ANTHONY P. ATKINSON

University of Durham, Durham, England

JASON TIPPLES

University of Hull, Hull, England

D. MICHAEL BURT

University of Durham, Durham, England

and

ANDREW W. YOUNG

University of York, York, England

Previous research with speeded-response interference tasks modeled on the Garner paradigm has demonstrated that task-irrelevant variations in either emotional expression or facial speech do not interfere with identity judgments, but irrelevant variations in identity do interfere with expression and facial speech judgments. Sex, like identity, is a relatively invariant aspect of faces. Drawing on a recent model of face processing according to which invariant and changeable aspects of faces are represented in separate neurological systems, we predicted asymmetric interference between sex and emotion classification. The results of Experiment 1, in which the Garner paradigm was employed, confirmed this prediction: Emotion classifications were influenced by the sex of the faces, but sex classifications remained relatively unaffected by facial expression. A second experiment, in which the difficulty of the tasks was equated, corroborated these findings, indicating that differences in processing speed cannot account for the asymmetric relationship between facial emotion and sex processing. A third experiment revealed the same pattern of asymmetric interference through the use of a variant of the Simon paradigm. To the extent that Garner interference and Simon interference indicate interactions at perceptual and response-selection stages of processing, respectively, a challenge for face processing models is to show how the same asymmetric pattern of interference could occur at these different stages. The implications of these findings for the functional independence of the different components of face processing are discussed.

The Components of Face Processing and Their Functional Independence

A common conception of face perception is that its different component processes are to some extent functionally independent of each other (see, e.g., Bruce & Young, 1986; Ellis, 1989; Young, 1998). There is plenty of evidence for this view, from dissociations following brain injury (see, e.g., Campbell, Landis, & Regard, 1986; Humphreys, Donnelly, & Riddoch, 1993; Parry, Young,

Saul, & Moss, 1991; Young, Newcombe, de Haan, Small, & Hay, 1993) to the results of speeded judgment tasks in normal participants (see, e.g., Bruce, Ellis, Gibling, & Young, 1987; Campbell, Brooks, de Haan, & Roberts, 1996; Young, McWeeny, Hay, & Ellis, 1986). Evidence from neurophysiological studies of nonhuman primates and functional brain-imaging studies in humans indicates that this functional separation is underpinned by anatomically separable components of a distributed face-processing system (see, e.g., Hasselmo, Rolls, & Baylis, 1989; Puce, Allison, Bentin, Gore, & McCarthy, 1998; see Haxby, Hoffman, & Gobbini, 2000, for a review).

Although our appearances change gradually over the years, facial identity is invariant relative to the rapid changes that occur in our faces when we speak or emote. In their model of face perception, Bruce and Young (1986) proposed that, subsequent to an initial common stage of structural processing, distinct sets of processes are involved in recognizing different aspects of faces, including identity, emotion, and speech. More recently, Haxby et al. (2000) have proposed a model of face processing

A.P.A. was supported by a Leverhulme Trust Research Fellowship and a grant from the McDonnell Project in Philosophy and the Neurosciences (funded by the James S. McDonnell Foundation; see www.sfu.ca/neurophilosophy). J.T. and A.W.Y. were supported by MRC Co-operative Group Component Grant G9812945. We are grateful to Cathy Blacklock, Sefton Loftus, Kathryn Jones, and Laura Hallett for helping with data collection, and to the referees for helpful comments on an earlier version of this article. Correspondence concerning this article should be addressed to A. P. Atkinson, Department of Psychology, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, England (e-mail: a.p.atkinson@durham.ac.uk).

that shares some elements of Bruce and Young's model but fits the various components of face processing into two broadly defined streams. On the basis of neuroimaging and electrophysiological research, Haxby et al. proposed one route, leading from inferior occipital cortex to inferotemporal cortex, in which relatively invariant aspects of faces are represented, and another route, leading from inferior occipital cortex to superior temporal cortex, in which changeable aspects of faces resulting from movement of the facial features are represented. (In this article, we shall use the general term *expressions* as shorthand for the changes in the shapes of facial features created by muscle movements involved in talking and emoting.) These two processing streams are regarded as being largely functionally independent of each other, although Haxby et al.'s account of the model does allow for a modicum of interaction. Furthermore, the model posits that specific face perception functions are performed by the coordinated interaction of each of these two streams with other neural regions. For example, the perception of facial emotion involves the occipital to superior temporal stream operating in concert with regions involved in representing emotional information, such as the amygdala and the insula, whereas the perception of identity involves the occipital to inferotemporal stream operating in concert with more anterior temporal regions.

Face Processing Interference With the Garner Paradigm

The results of some studies in which a speeded-judgment interference task was employed have been interpreted as indicating that certain components of face processing can interact and so are not completely independent. Using the Garner paradigm (Garner, 1976), Schweinberger and colleagues found that facial identity judgments remain unaffected by task-irrelevant variations in either facial expression or facial speech, yet variations in facial identity did influence judgments about facial expression and facial speech (Schweinberger, Burton, & Kelly, 1999; Schweinberger & Soukup, 1998; see also Baudouin, Martin, Tiberghien, Verlut, & Franck, 2002).

In the Garner paradigm, participants are required to classify exemplars of a stimulus, such as a face, along a particular dimension while a second, task-irrelevant dimension of the stimulus is varied. The stimuli are presented in three different conditions, defined by the relationship between the task-relevant and task-irrelevant dimensions. In the control condition, the irrelevant dimension is held constant across variations in the relevant dimension. For example, all faces might express the same emotion when the task is to judge identity, and all faces might be of the same person when the task is to judge facial expression. In the orthogonal condition, both the irrelevant and relevant dimensions are varied (e.g., two people express each of two different emotions). In the correlated condition, changes in the irrelevant dimension covary with changes in the relevant dimension (e.g., Person A always expresses fear and Person B always expresses happiness). The logic

of the Garner paradigm is that a difference in response times (RTs) to the stimuli across these three conditions is evidence that the processing of the different stimulus dimensions is not independent; in other words, it is evidence that participants are unable to pay selective attention to one dimension without the other dimension's influencing their responses. The strongest evidence of independence is longer RTs in the orthogonal than in the control condition. Shorter RTs in the correlated condition than in the control condition, indicating facilitation of classification ("redundancy gain"), are consistent with but not strongly indicative of nonindependent or integral processing (for discussion, see Eimas, Tartter, Miller, & Keuthen, 1978; Green & Kuhl, 1991; Schweinberger & Soukup, 1998).

In Schweinberger and Soukup's (1998) first two experiments, participants classified computer-presented photographs of faces with respect to either identity (Person A or Person B) or expression (happy or sad) while ignoring the other, irrelevant dimension. RTs for identity judgments did not differ across the control, orthogonal, and correlated conditions. RTs for expression judgments, on the other hand, were significantly different across the three conditions; specifically, RTs for the correlated condition were reliably shorter than those for the control condition, which in turn were reliably shorter than RTs for the orthogonal condition. In subsequent experiments, Schweinberger and Soukup found a similar pattern of results when identity and facial speech judgments were compared. Identity judgments were not affected by variation in facial speech postures; however, when judging facial speech, participants were reliably faster in the correlated condition than in the control condition and reliably faster in the control condition than in the orthogonal condition. These *asymmetric dependencies* between the processing of facial identity, emotion, and speech indicate that observers are able to attend and respond to the identity of faces while ignoring emotional and speech expressions, but they are unable to ignore identity when attending and responding to either emotional expression or speech posture (Schweinberger & Soukup, 1998).

An alternative explanation for Schweinberger and Soukup's (1998) findings is that the interference effect is due simply to a difference in the processing speeds of the two types of information, as is indicated by different overall RTs for the two tasks. Such a speed-of-processing account predicts that faster tasks can interfere with slower tasks, but the opposite is not true (see the introduction to Experiment 2 for a more detailed discussion). This explanation is consistent with Schweinberger and Soukup's results insofar as participants were faster to classify identity than they were to classify emotion (a finding that is also consistent with electrophysiological evidence that identity processing occurs substantially earlier than expression processing: Münte et al., 1998). In order to test this alternative explanation, Schweinberger et al. (1999) manipulated the difficulty of identity- and emotion-judgment tasks by employing a computer morphing technique to create face stimuli that blended, to varying degrees,

either the identities of two faces within a given emotion or two emotions (happy and angry) within a given identity. Difficult stimuli were defined as 71:29 and 29:71 blends across the relevant dimension, and easy stimuli were defined as the endpoints of these continua (100:0 and 0:100). The irrelevant dimension was not morphed; it remained fixed at the endpoints. Employing these stimuli in Garner tasks, Schweinberger et al. replicated Schweinberger and Soukup's finding that variation in identity interfered with classifications of emotional expression, but variation in emotional expression did not interfere with classifications of identity. Importantly, they found that this asymmetric pattern of interference held even when identity was more difficult (and therefore took longer) to discriminate than emotion.

Also using the Garner paradigm, Baudouin et al. (2002) have demonstrated that schizophrenic individuals, as well as healthy controls, show this asymmetric interaction between the processing of facial expression and that of identity, despite the fact that schizophrenics tend to have impaired face-processing skills, especially in emotion recognition. Such impairments were manifest in the RTs of Baudouin et al.'s participants but not in their accuracy. Although the schizophrenic patients performed just as accurately as the healthy controls on the expression and identity classification tasks, the schizophrenics were considerably slower than the controls in every condition in each task. Moreover, this RT deficit was greater for emotion classification than for identity classification. Nevertheless, a typical pattern of Garner interference was obtained with both the schizophrenics and the healthy participants without any significant difference between the two groups. Variation in identity increased RTs for classification of facial expression, but variation in facial expression did not influence RTs for classification of identity.

Interpreting Asymmetric Interference on Facial Variants of the Garner Paradigm

One explanation that Schweinberger et al. (1999) offered for their findings of asymmetric Garner interference is that identity information can provide a reference for computing information about facial speech and emotion, but the opposite is not the case. Presumably, it is rather less likely that information about emotional and speech expressions can provide references for computing identity, for the very reason that such information is changeable. Cast in terms of Haxby et al.'s (2000) neurological model, we can think of Schweinberger et al.'s results as indicating some form of interaction between neurological systems that code invariant aspects of faces and those that encode changeable aspects. Note that this conception of Haxby et al.'s model implies that the connection between these two systems is one-way, or at least is considerably stronger in one direction than in the other. Although Haxby et al.'s depiction of their model shows a two-way arrow linking the distinct systems in which invariant and changeable aspects of faces are represented,

this is an assumption for which no evidence is presented. Moreover, Haxby et al. remark, "The degree of separation between the functional roles played by the different regions . . . is unclear" (p. 231). Their only suggestion of a possible interaction between these two systems is a one-way interaction, similar to our own proposal—namely, that the system that encodes invariant information about faces "may play a supportive role in the perception of expression, perhaps because different individuals can have characteristic expressions, such as a crooked smile or a wry grin, that we associate uniquely with them" (p. 231).

The sex of faces, like their identity, is a relatively invariant property. Thus, given our interpretations of Haxby et al.'s (2000) model and of the aforementioned findings of asymmetric Garner interference between facial identity and facial emotion judgments, we would expect task-irrelevant variations in emotional expression to interfere with sex judgments, but we would not expect irrelevant variations in sex to interfere with expression judgments. However, Le Gal and Bruce (2002) used the Garner paradigm to examine the relationship between facial expression and sex processing and found little evidence of interference between these two dimensions. In two experiments in this study, participants judged either the sex of faces or whether the faces expressed anger or surprise. In the first of these experiments, RTs to judge the sex of faces did not differ significantly among the correlated, control, and orthogonal conditions. RTs to judge emotional expression did not differ between the control and orthogonal conditions, although RTs were significantly shorter in the correlated condition than in these two conditions. This relationship was examined again in a second experiment, in which photographs of faces were cropped to remove the hairline and jawline in order to increase the difficulty of the sex classification task to approximately equal that of the emotion classification task, as indicated by equivalent overall RTs for the two tasks. For sex classifications, there were no significant differences among the three conditions. For expression classifications, RTs in the correlated condition were reliably shorter than RTs in the control condition, but there was no significant difference between the orthogonal and control conditions. In contrast to these speeded-judgment findings, however, a prior rating experiment showed that variation in emotional expression affected the perceived relative femininity or masculinity of faces; surprised expressions increased the judged femininity of faces and decreased their judged masculinity relative to angry expressions. These findings provide, in Le Gal and Bruce's view, "qualified support" (p. 242) for the independence of facial emotion and sex processing.

The Present Study

To examine further the utility of Haxby et al.'s (2000) model, we investigated the interaction between sex (a paradigmatic example of invariance) and (paradigmatically changeable) emotional expression using two different speeded-response interference paradigms. If an invariant

aspect of facial information—identity or sex—can provide references for our judgments about more changeable aspects of faces, including emotional and speech information, then we would expect to find asymmetric interference effects in experiments in which judgments about the sex of a face are pitted against judgments about its particular emotional expression. In Experiments 1 and 2, we employed a variant of the Garner paradigm in which participants classified the emotion or sex of faces whose irrelevant dimension (their sex in the case of emotion judgments, or their expressions in the case of sex judgments) either varied or remained constant. In Experiment 3, we sought additional evidence from the Simon paradigm (Simon, 1990; Simon & Acosta, 1982; Simon & Rudell, 1967). In our facial variant of the Simon paradigm, which we adapted from De Houwer, Hermans, and Eelen (1998), participants classified the emotional expression or sex of faces by saying a word that was either semantically congruent or semantically incongruent with the irrelevant dimension of the face. The Simon paradigm is widely thought to involve interference at an output stage of processing rather than at an earlier perceptual stage (see, e.g., Kornblum & Lee, 1995; Lu & Proctor, 1995; Umiltà & Nicoletti, 1990, 1992). Given that Garner interference effects are sometimes interpreted in terms of perceptual interaction, a finding of the same pattern of asymmetric interference with the facial variants of the Garner and Simon paradigms would raise questions about the location of the interference and about whether the two tasks in fact tap the same source of interference. We shall return to these issues in the General Discussion.

EXPERIMENT 1

Method

Participants. Thirty-two students from University College Winchester were recruited for Experiments 1A and 1B and were paid for their participation. Sixteen of these students (12 female) took part in Experiment 1A (age range, 19–26 years; mean age, 21.1 years), and the other 16 (12 female) took part in Experiment 1B (age range, 18–35 years; mean age, 20.9 years). All of the participants had normal or corrected-to-normal eyesight.

Stimuli and Apparatus. Sixteen face stimuli were used in this experiment, in the form of digitized black-and-white photographs. Eight of the pictures showed different people with fearful facial expressions, and the remaining pictures showed the same individuals with happy facial expressions. Each set of fearful and happy faces consisted of four pictures of men and four pictures of women.

Each face stimulus was a caricatured version of a face from the Ekman and Friesen (1976) pictures of facial affect, taken from the FEEST database of facial expression stimuli (Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). The expressions selected were computer caricatured by 50% using the techniques described by Calder, Young, Rowland, and Perrett (1997). Stimuli from the FEEST were chosen so that the expressions from which they were derived would be good exemplars of basic emotion categories, including those of fear and happiness. Not only are the original Ekman and Friesen pictures among the most used and thus well-normed sets of facial affect stimuli, but, in addition, their caricatured versions are identified more quickly than their noncaricatured counterparts with no loss in accuracy (Calder et al., 1997), and people perceive caricatured expressions as more emotionally intense than their

noncaricatured counterparts (Benson, Campbell, Harris, Frank, & Tovée, 1999; Calder et al., 2000). Another important feature of the selected stimuli was that the faces were cropped around the hairline, eliminating the possibility that participants might base their sex judgments on hairstyle. Our intention in using cropped hairlines was to make the sex-judgment task harder than usual, and by using caricatured expressions we hoped to make the expression-judgment task easier than usual. This was done to counter as much as possible the tendency for sex classification to be an easier task than emotion classification.

An Apple Macintosh PowerPC 8100/80 computer, running Superlab 1.74 experimental software, controlled the stimulus displays and recorded RTs. All the faces appeared in the center of a uniform gray background, which covered the whole screen of the 12-in. color monitor in grayscale mode. When presented on the screen, the faces had a mean height of 10 cm (range, 9–11 cm) and a mean width of 7.2 cm (range, 6–8 cm), subtending angles of 4.8° vertically and 3.4° horizontally at a viewing distance of 60 cm. Manual responses were recorded via the computer keyboard.

Design. The aim of Experiment 1A was to examine whether the sex of a face would influence judgments of its emotional expression, whereas the aim of Experiment 1B was to examine whether a face's emotional expression would influence judgments of its sex. Since the contrast between orthogonal and control conditions is critical for establishing integral versus independent processing in the Garner task, with the correlated condition being of less theoretical interest, we chose to simplify the experimental design by including only orthogonal and control conditions. Thus, in Experiment 1A we compared RTs for classifying expression (happy vs. fearful) when the faces were all of the same sex with RTs for classifying expression when the faces depicted males and females in equal proportions. Similarly, in Experiment 1B we compared RTs for classifying faces as male or female when the faces all had the same expression with RTs for classifying sex when the faces had an equal mixture of happy and fearful expressions.

The set of 16 faces provided 8 stimuli for classifying the expressions of male faces (4 happy and 4 fearful) and 8 stimuli for classifying the expressions of female faces (4 happy and 4 fearful). Similarly, these 16 faces provided 8 stimuli for classifying the sex of happy faces (4 male and 4 female) and 8 stimuli for classifying the sex of fearful faces (4 male and 4 female). For the tasks of classifying the expressions of faces that were either male or female (mixed sex) and classifying the sex of faces that were either happy or fearful (mixed emotion), 16 stimuli were available. Thus, in order to ensure an equal number of stimuli in each block of trials, the 16 stimuli were divided into two matched blocks for the orthogonal conditions (i.e., mixed sex in Experiment 1A and mixed emotion in Experiment 1B).

For Experiment 1A, each orthogonal (mixed-sex) block consisted of one male (E.M. or J.J.) and one female (C. or M.O.), each with both happy and fearful expressions, and the control (single-sex) blocks consisted of two males (P.E. and W.F.) or two females (M.F. and S.W.), each with both happy and fearful expressions. Similarly, for Experiment 1B each orthogonal (mixed-emotion) block consisted of one male (E.M. or J.J.) and one female (C. or M.O.), each with both happy and fearful expressions, and each control (single-emotion) block consisted of two males (P.E. and W.F.) and two females (M.F. and S.W.) showing either all happy expressions or all fearful expressions. For both experiments, each of the four different stimuli within each block was presented eight times, yielding a total of 32 trials per block. All trials within each block were presented in a pseudorandom order.

Eight different versions of Experiment 1A and eight different versions of Experiment 1B were created, and the participants were randomly assigned to one of these 16 versions. These different versions of the experiments were created in order to control for (1) the order of presentation of the four blocks within an experiment (four differ-

ent orders determined by a Latin square design) and (2) the stimulus–response mapping—that is, the way in which the two response keys were matched with the responses (“fearful” and “happy” in Experiment 1A, and “male” and “female” in Experiment 1B).

Procedure. The participants were seated in a quiet room in front of the computer, approximately 60 cm from the monitor screen. Written instructions were presented on the screen and summarized orally by the experimenter. Each trial began with a fixation cross (+), which appeared in the center of the screen for 1,500 msec. The fixation cross was followed immediately by a stimulus face, which also appeared in the center of the screen. The face remained on the screen until the participant’s keypress triggered its offset, at which point the fixation cross appeared again, followed by the next stimulus. The response keys were “z” and “/,” which were located 17 cm apart on the row of the keyboard second from the bottom. The participants were asked to respond as quickly and accurately as possible to the emotional expression displayed by the face (for Experiment 1A) or to the face’s sex (for Experiment 1B) and to ignore other aspects of the face. The experiments began with a practice block consisting of each of the 16 faces presented once, in pseudorandom order.

Results

Incorrect responses accounted for 2.8% of the data overall. Outliers, defined as RTs greater than 3,000 msec, accounted for only two datapoints, one in each task; these were excluded from the RT and error analyses.

Response times. The mean correct RTs were initially analyzed in a mixed-design ANOVA with task (sex classification vs. emotion classification) as the between-subjects variable and condition (orthogonal vs. control) as the repeated measures variable. There were main effects of task [$F(1,30) = 15.01, p < .005; \eta^2 = .33$] and condition [$F(1,30) = 9.48, p < .01; \eta^2 = .24$]. The main effect of task indicated that, averaged across condition, the participants were reliably slower to judge the emotions ($M = 668$ msec, $SEM = 23$) than they were to judge the sex

($M = 544$ msec, $SEM = 23$) of the faces. The main effect of condition indicated that, averaged across tasks, the participants were reliably slower in the orthogonal condition ($M = 622$ msec, $SEM = 18$) than in the control condition ($M = 590$ msec, $SEM = 15$). Of greater theoretical interest was the task \times condition interaction [$F(1,30) = 6.59, p < .05; \eta^2 = .18$]. This interaction is depicted in Figure 1, which indicates that there was a clear effect of condition for emotional expression classifications but not for sex classifications. This was confirmed by simple main effect analyses of the effect of condition for each task separately. When the participants were required to judge the emotion of the faces, they were slower to respond in orthogonal blocks ($M = 698$ msec, $SEM = 26$) than in control blocks [$M = 638$ msec, $SEM = 22; F(1,15) = 22.19, p < .005; \eta^2 = .59$]. However, when the participants were required to judge the sex of the faces, there was no difference between responses to orthogonal blocks ($M = 546$ msec, $SEM = 26$) and those to control blocks [$M = 541$ msec, $SEM = 22; F(1,15) = .10, p > .5; \eta^2 = .007$].

In order to compare RTs for every combination of each relevant and each irrelevant dimension, separate ANOVAs were carried out for each task, with the repeated measures variables of condition, sex, and expression. For expression classifications, there were significant main effects of condition [$F(1,15) = 20.7, p < .001; \eta^2 = .58$] and emotion [$F(1,15) = 10.17, p < .01; \eta^2 = .404$], but no significant interactions. The new finding revealed by this ANOVA is that the participants were slower to judge fearful facial expressions as fearful ($M = 704$ msec, $SEM = 37$) than they were to judge happy facial expressions as happy ($M = 640$ msec, $SEM = 24$), averaged across condition. For sex classifications, no main effects of condition, sex, or expression reached significance. No interac-

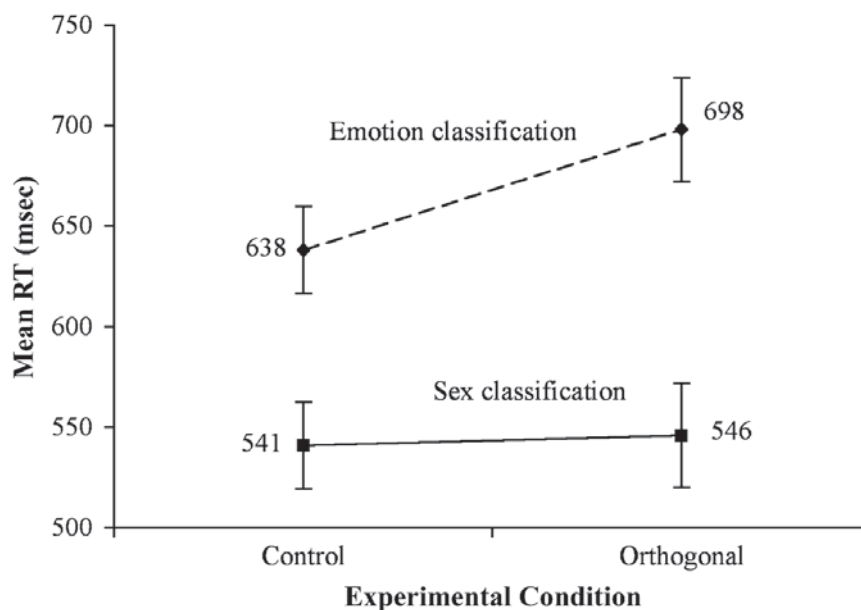


Figure 1. Mean response times (RTs) for classifying the emotion (Experiment 1A) and sex (Experiment 1B) of faces in the control and orthogonal conditions of the Garner paradigm.

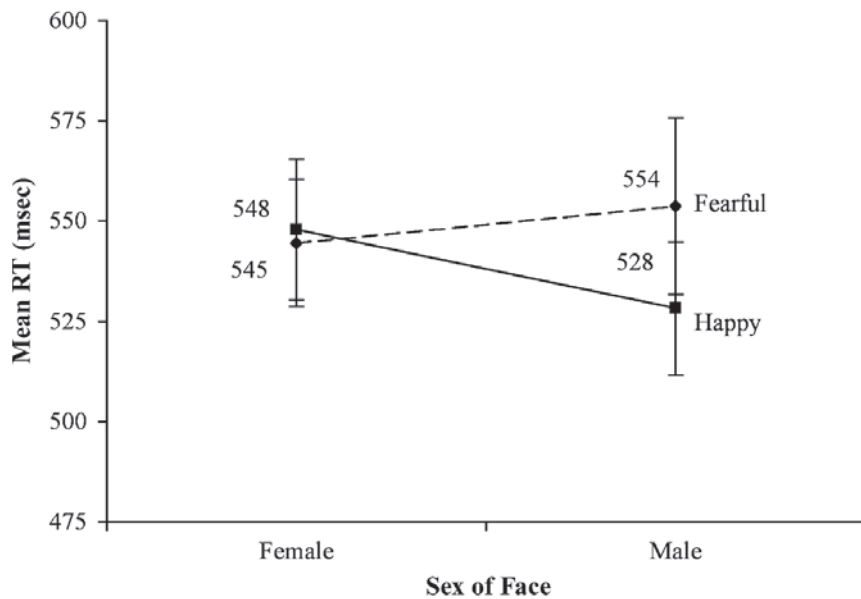


Figure 2. Mean response times (RTs) for classifying the sex of faces (Experiment 1B) as a function of the emotion on the faces.

tions reached significance either, but the emotion \times sex interaction approached significance [$F(1,15) = 4.4, p = .053; \eta^2 = .227$]. Inspection of Figure 2 reveals that the participants were somewhat slower to judge a fearful male face as male than they were to judge a happy male face as male, but there was essentially no difference in their RTs to judging fearful versus happy female faces as female.

Errors. A subsidiary set of analyses was used to examine error rates. Two separate one-way repeated measures ANOVAs were conducted to examine the percentage of errors for each block type in each task. For emotion classification, error rates were higher for the orthogonal blocks ($M = 5.67\%$, $SEM = 1.33$) than for the control blocks ($M = 1.95\%$, $SEM = 0.56$) [$F(1,15) = 12.12, p < .01; \eta^2 = .45$]. For sex classification, error rates were not significantly different between the orthogonal ($M = 2.05\%$, $SEM = 0.51$) and the control ($M = 1.56\%$, $SEM = 0.47$) blocks [$F(1,15) = 0.63, p > .1; \eta^2 = .04$].

Discussion

The main result of this experiment was that RTs were longer in the orthogonal than in the control condition of the emotion classification task, whereas there was no significant difference in RTs between the orthogonal and control conditions for sex classification. The error rates followed the same pattern, with more errors in the orthogonal condition than in the control condition for emotion classification but not for sex classification, thus confirming that the RT results were not due to a speed-accuracy trade-off. Despite the lack of an overall interference effect for sex classification, the more finely grained RT analysis of the intratask relationships yielded some evidence to suggest that a face's expression can influence sex classification; RTs to male faces portraying fearful expressions tended to be longer than RTs to male faces portraying happy expressions.

On the standard interpretation of Garner interference, the overall asymmetric interference effect obtained in Experiment 1 can be construed as evidence that although people are able to attend selectively to the sex of faces while ignoring variations in facial expression, they are less able to attend selectively to facial expressions while ignoring variations in the sex of the faces. (However, the nearly significant emotion \times sex interaction for sex classification suggests that for the male faces in this experiment, the participants could not entirely ignore variations in emotional expression.)

EXPERIMENT 2

Our second experiment had two aims. One was to replicate our finding of asymmetric Garner interference between sex and emotion classification, given that Le Gal and Bruce (2002) did not find such an effect. The other aim of this experiment was to examine whether or not this asymmetric interference would still be evident when the sex and emotion classification tasks were more equal in difficulty, which would argue against an account of the asymmetric relationship in terms of relative processing speeds.

According to a speed-of-processing account, the asymmetric pattern of interference is at least partially a result of the mismatch in task difficulty (Le Gal & Bruce, 2002; Melara & Mounts, 1993; Schweinberger et al., 1999), as is indicated by longer RTs and higher error scores for emotion classifications in comparison with sex classifications. The basic argument would run as follows: Information concerning the sex of faces (being the more discriminable dimension) is computed before the emotional expression (the less discriminable dimension) has been determined. Consequently, when the task is to classify sex in the orthogonal condition of the Garner paradigm, a

response can usually be initiated before information about the face's emotional expression has been fully processed. But when the task is to classify emotion in the orthogonal condition, the sex of the face is determined before a response regarding its emotion can be initiated, and thus information regarding the sex of the face is available as a reference for processing information regarding its emotional expression.

Attempts to equate or reverse the initial difficulty of the paired tasks in facial variants of the Garner paradigm do not provide strong support for such an argument, however. Le Gal and Bruce (2002) did not find any significant interference effects between sex and emotion classifications in an experiment in which both classifications were assumed to be approximately equal in difficulty, as is evidenced by approximately equal mean RTs in the two tasks (in fact, sex classifications were marginally though not significantly slower than emotion classifications). This is admittedly not strong evidence against the mismatch-in-difficulty argument for sex versus emotion classification, since Le Gal and Bruce found no interference effect in an experiment in which emotion classifications took longer (i.e., were more difficult) than sex classifications. Yet a mismatch in task difficulty is unlikely to be able to account entirely for our own finding of an asymmetric relationship between facial emotion and sex classification. As was noted earlier, Schweinberger et al. (1999) found that equating and even reversing task difficulty did not eliminate the same asymmetric pattern of interference between identity and expression classifications. This suggests that we should also expect equating or reversing task difficulty to preserve our interference effects, given our assumptions that invariant and changeable aspects of faces are represented in distinct processing routes and that the sex of faces, like their identity, is a relatively invariant property.

Despite this reason for doubting a relative speed-of-processing account of our asymmetric interference effect, it behoved us to test directly the validity of such an account. Consequently, we attempted to vary the difficulty of (and thus the RTs for) the sex and emotion classification tasks in Experiment 2 by morphing the faces across the task-relevant dimension (cf. Schweinberger et al., 1999).

Method

Participants. Thirty-two students from the University of Durham were recruited; 16 (8 female; age range, 19–21 years; mean age, 20.1 years) took part in Experiment 2A, whereas the other 16 (8 female; age range, 19–22 years; mean age, 20.9 years) took part in Experiment 2B. All of the participants had normal or corrected-to-normal eyesight.

Stimuli and Apparatus. The bases for the stimulus set in Experiment 2 were faces from the FEEST set (Young et al., 2002) comprising four males (E.M., J.J., P.E., and W.F.) and four females (C., M.F., M.O., and N.R.) with both fearful and happy caricatured (+50%) expressions. (Thus, the faces were the same as those used in Experiment 1, with the exception of one of the female faces.) Two sets of morphed continua were created from these original faces. For Experiment 2A (emotion classification), each of the eight identities was morphed across emotion (e.g., J.J.'s face morphed from happy to

fearful) with identity and thus sex held constant. For Experiment 2B (sex classification), the faces were morphed across sex and thus identity, with expression held constant, so that one particular female face and one particular male face were morphed together with their fearful expressions and again with their happy expressions. (C. was paired with W.F., M.F. with E.M., M.O. with J.J., and N.R. with P.E.) The choice of the particular female and male faces to be morphed together was based on the criterion that they should have as many Facial Action Coding System (Ekman & Friesen, 1978) action units in common as possible for each emotion (action unit data obtained from Young et al., 2002). Six equidistant morph levels were selected for the morphed continua of emotion and sex: 75:25, 65:35, 55:45, 45:55, 35:65, and 25:75 (the ratios expressing proportions of initial image to final image). This resulted in a total of 48 different stimuli in each set, each of which was presented four times during the course of the experiment, resulting in a total of 192 trials for Experiment 2A and 192 trials for Experiment 2B.

The stimulus presentation and recording of RTs was controlled by Superlab 1.74 experimental software running on an Apple Macintosh G3 computer with a 19-in. color monitor in grayscale mode. The dimensions of the stimuli on the screen were the same as those in Experiment 1. Manual responses were recorded via the computer keyboard.

Design. For reasons similar to those that apply to Experiment 1, two orthogonal and two control conditions were created for each of Experiments 2A and 2B. For Experiment 2A, each orthogonal (mixed-sex) block consisted of one female and one male (C. with W.F. or M.O. with J.J.), each morphed across emotion, and the control (single-sex) blocks consisted of two females (M.F. and N.R.) or two males (E.M. and P.E.), each morphed across emotion. For Experiment 2B, each orthogonal (mixed-emotion) block consisted of one female–male pair (C. with W.F. or M.O. with J.J.) morphed across identity (and thus sex) for their fearful expressions and again for their happy expressions, whereas the control (single-emotion) blocks consisted of two female–male pairs (M.F. with E.M. and N.R. with P.E.) morphed across either their happy or their fearful expressions. For both Experiments 2A and 2B, within each block, each of the 12 different stimuli (2 morphed continua \times 6 morph levels) was presented four times, yielding a total of 48 trials per block. All trials within each block were presented in a pseudorandom order.

As in Experiment 1, eight different versions of Experiment 2A and eight different versions of Experiment 2B were created, permitting counterbalancing of block order and stimulus–response mapping. The participants were randomly assigned to 1 of these 16 different versions.

Procedure. The procedural details were the same as those for Experiment 1, except that this time the practice block consisted of each of the eight morphed continua presented twice, once at the 75:25 morph level and once at the 25:75 morph level, in pseudorandom order. As in Experiment 1, the task was to respond as quickly and accurately as possible to the emotional expression displayed by each face (for Experiment 2A) or to the face's sex (for Experiment 2B) while ignoring other aspects of the face.

Results

Classification performance. Figure 3 depicts the percentages of “fearful” and “male” responses for emotion and sex classification, respectively, as a function of morph level and experimental condition. With respect to the control and orthogonal conditions of the emotion classification task, it can be seen that except for the most ambiguous morph levels (i.e., those with either 55:45 or 45:55 proportions of the two emotions), the stimuli were very consistently classified as displaying either fear or

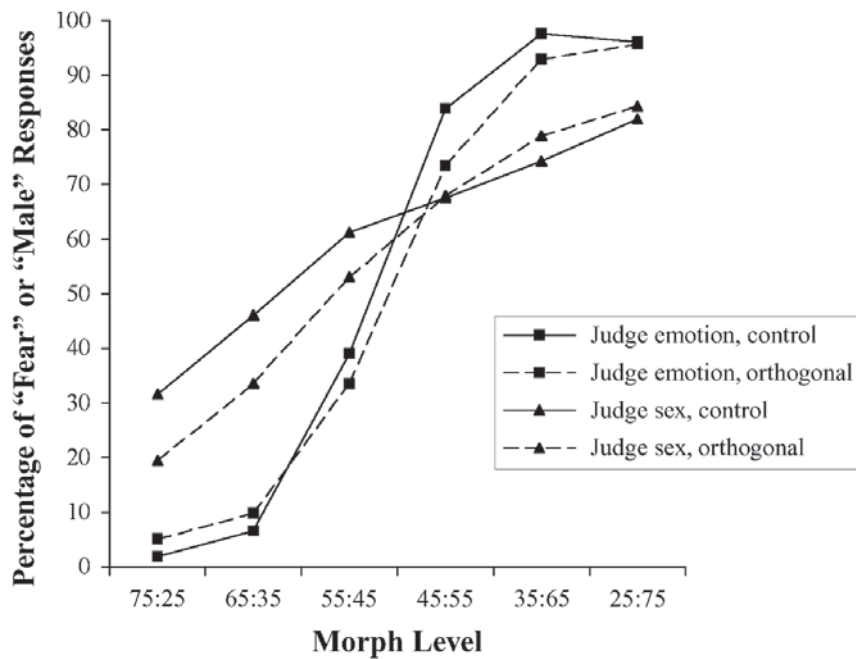


Figure 3. Percentages of “fearful” or “male” responses, depending on task and condition in Experiment 2, as a function of morph level (happy:fearful, or female:male).

happiness. With respect to the control and orthogonal conditions of the sex classification task, on the other hand, the stimuli were not as consistently classified as male or female, even for the morph levels toward the endpoints of the continua. These data were initially subjected to an ANOVA across tasks, with repeated measures on the variables of condition and morph level. This mixed-design ANOVA revealed a significant main effect of condition [$F(1,30) = 4.4, p < .05; \eta^2 = .13$] and a highly significant main effect of morph level [$F(1.7,50.9) = 228.97, p < .001; \eta^2 = .88$, Greenhouse–Geisser corrected]. The main effect of condition reflected the fact that the mean ratio of “female” to “male” or “happy” to “fearful” responses was reliably greater in the control conditions ($M = 57.4:42.6, SEM = 1.8$) than in the orthogonal conditions ($M = 54.0:46.0, SEM = 1.3$). Pairwise comparisons (Bonferonni corrected, $\alpha = .05$) for the variable of morph level revealed that, averaged over task and condition, all morph levels differed significantly from each other. The main effect of task approached significance [$F(1,30) = 3.89, p = .058; \eta^2 = .12$], reflecting a tendency for the mean ratio of “female” to “male” responses in the sex classification task ($M = 58.4:41.6, SEM = 1.9$) to be greater than the mean ratio of “happy” to “fearful” responses in the emotion classification task ($M = 53.0:47.0, SEM = 1.9$). These main effects were modified by a significant task \times morph level interaction [$F(5,150) = 26.53, p < .001; \eta^2 = .47$], which was in turn modified by a significant task \times condition \times morph level interaction [$F(5,150) = 7.01, p < .001; \eta^2 = .19$].

To follow up on this significant three-way interaction, separate ANOVAs were conducted for each task, with repeated measures on the variables of condition and morph level. There were highly significant main effects of morph

level for both emotion classification [$F(2.17,32.56) = 481.17, p < .001; \eta^2 = .97$] and sex classification [$F(1.34,20.02) = 34.53, p < .001; \eta^2 = .7$; Greenhouse–Geisser corrected in both cases]. Pairwise comparisons (Bonferonni corrected, $\alpha = .05$) for the variable of morph level revealed that, for emotion classification, all morph levels differed significantly from each other except for levels 35:65 and 25:75, whereas for sex classification, morph levels 55:45 and 45:55 did not differ significantly from each other. There was no main effect of condition for either sex or emotion classification, although in the case of the latter there was a nonsignificant trend toward a higher mean ratio of “happiness” responses to “fearful” responses in the control ($M = 54.2:45.8, SEM = 1.4$) than in the orthogonal ($M = 51.8:48.2, SEM = 1.3$) conditions [$F(1,15) = 4.13, p = .06; \eta^2 = .22$]. There were significant condition \times morph interactions for both emotion classification [$F(2.92,43.81) = 3.55, p < .05; \eta^2 = .19$, Greenhouse–Geisser corrected] and sex classification [$F(5,75) = 5.19, p < .001; \eta^2 = .26$]. As Figure 3 indicates, these interactions reflected differences between the control and orthogonal conditions for some but not all morph levels, which was confirmed by simple main-effects analyses. For emotion classification, the consistency of response was higher in the control than in the orthogonal conditions for morph levels 75:25, 45:55, and 35:65, whereas for sex classification the consistency of response was higher in the orthogonal than in the control conditions for morph levels 75:25 and 65:35 [all $ps < .05$; morph level 55:45 approached significance for sex classification ($p = .057$)].

Classification response times. In order to analyze the RTs to stimuli of different levels of difficulty on the task-relevant dimension, we collapsed the data from the six

individual morph levels into three levels of difficulty: 1 (75:25 and 25:75), 2 (65:35 and 35:65), and 3 (55:45 and 45:55). Since our primary goal was to select certain levels of morphing to test whether asymmetric interference would still be evident when task difficulty was equated, and because it is not very meaningful to analyze RTs at near-chance levels of performance, we entered into the RT analyses the data for difficulty levels 1 and 2 only. We initially performed an ANOVA across tasks, with repeated measures on the variables of condition (control vs. orthogonal) and difficulty level (1 vs. 2). There was no main effect of task [$F(1,30) = 0.06, p > .5; \eta^2 = .002$], indicating that RTs for the classification of sex ($M = 900$ msec, $SEM = 68$) were similar to those for the classification of emotion ($M = 877$ msec, $SEM = 68$). The main effects of condition and difficulty did not reach significance either (all F s < 0.2). However, there was a significant task \times difficulty interaction [$F(1,30) = 7.47, p < .05; \eta^2 = .2$], a significant task \times condition interaction [$F(1,30) = 15.67, p < .001; \eta^2 = .34$], and a significant condition \times difficulty interaction [$F(1,30) = 4.24, p < .05; \eta^2 = .12$]. These two-way interactions were modified by a significant task \times difficulty \times condition interaction [$F(1,30) = 4.78, p < .05; \eta^2 = .14$].

To analyze these interactions further, we performed separate ANOVAs for each task, with repeated measures on the variables of condition (control vs. orthogonal), difficulty level (1 [easier] vs. 2 [more difficult]), and irrelevant dimension of the face (male vs. female in the case of emotion classification, and fear vs. happiness in the case of sex classification). For emotion classification, there was a significant main effect of condition [$F(1,15) = 15.22, p < .005; \eta^2 = .5$], reflecting the expected result—namely, that RTs in the orthogonal condition ($M = 914$ msec, $SEM = 46$) were reliably longer than RTs in the control condition ($M = 835$ msec, $SEM = 42$). There was also a significant main effect of difficulty [$F(1,15) = 6.97, p < .05; \eta^2 = .32$], reflecting the fact that the participants were slower to respond to morphed faces of Difficulty Level 2 ($M = 906$ msec, $SEM = 49$) than to morphed faces of Difficulty Level 1 ($M = 842$ msec, $SEM = 40$). There was no significant main effect of the sex of the face, and none of the interactions was significant (all F s < 1).

For sex classification, there was a significant main effect of condition [$F(1,15) = 5.21, p < .05; \eta^2 = .26$], reflecting the fact that, overall, RTs in the orthogonal condition ($M = 856$ msec, $SEM = 74$) were reliably shorter than RTs in the control condition ($M = 933$ msec, $SEM = 98$). There was also a significant main effect of difficulty [$F(1,15) = 7.86, p < .05; \eta^2 = .34$], which was due to the fact that, overall, the participants were faster to respond to morphed faces of Difficulty Level 2 ($M = 867$ msec, $SEM = 80$) than to morphed faces of Difficulty Level 1 ($M = 922$ msec, $SEM = 90$). These main effects were modified by a significant condition \times difficulty interaction [$F(1,15) = 8.95, p < .01; \eta^2 = .37$]. Simple main-effects analyses for this significant interaction revealed that RTs for faces of Dif-

iculty Level 1 were significantly shorter in the orthogonal condition ($M = 846$ msec, $SEM = 75$) than in the control condition [$M = 1,000$ msec, $SEM = 108; F(1,15) = 11.56, p < .005; \eta^2 = .44$], but that for faces of Difficulty Level 2 there was no significant difference in RTs between the control ($M = 896$ msec, $SEM = 89$) and the orthogonal ($M = 859$ msec, $SEM = 79$) conditions [$F(1,15) = 1.14, p > .1; \eta^2 = .07$].

Discussion

Morphing the faces across the task-relevant dimension clearly affected the difficulty of the corresponding task. The more the faces were blended on the relevant dimension, the more difficult the task became, as is indicated by both longer RTs and the decrease in consistency of the participants' responses. Crucially, these manipulations of task difficulty had the effect of equating the overall RTs across the emotion and sex classification tasks. Nevertheless, we still found an asymmetric pattern of Garner interference. The direction of this asymmetric interference was the same as that found in Experiment 1—that is, variation in the sex of the faces interfered with emotion classification performance more than variation in expression interfered with sex classification performance. Although sex classifications remained relatively unaffected by changes in facial expression, there was some evidence of an interaction, insofar as RTs were reliably shorter in the orthogonal condition than in the control condition for faces of Difficulty Level 1 (75:25 and 25:75 morphs). In sum, the results of this experiment suggest that differences in processing speed cannot account for the asymmetric relationship between facial emotion and sex processing.

EXPERIMENT 3

Experiments 1 and 2 demonstrated the value of distinguishing the processing of invariant aspects of faces from that of changeable aspects in interpreting asymmetric interference on facial variants of the Garner paradigm. The purpose of Experiment 3 was to determine whether the implementation of a different interference paradigm—a facial variant of De Houwer et al.'s (1998) affective Simon paradigm—would converge on the same pattern of findings.

De Houwer et al. (1998) investigated the effect of an irrelevant dimension on speeded forced-choice judgments about faces by using a variant of the affective Simon paradigm (De Houwer, Crombez, Baeyens, & Hermans, 2001; De Houwer & Eelen, 1998). Like the Garner paradigm, the Simon paradigm requires participants to respond to stimuli along a particular dimension while a second dimension of the stimulus is varied. However, unlike in the Garner paradigm, the required response is related to the irrelevant dimension. In De Houwer et al.'s (1998) study, participants classified the sex or identity of faces with a verbal response that was either congruent or incongruent with the valence of the facial expressions. The stimuli were photographs of two male and two female faces with both positive (happy) and negative (fearful, angry, disgusted,

and sad) expressions. In the sex classification task, half of the participants were asked to say “positive” if the face was male and “negative” if the face was female, whereas the remaining participants were asked to give the converse responses (i.e., “negative” to male faces and “positive” to female faces). In the identification task, the participants were asked to say “positive” to two of the four stimulus faces (one male and one female) and “negative” to the other two faces. (The female and male faces assigned the “positive” and “negative” responses were determined in a prior learning phase of the experiment.) In both tasks, the participants were told to ignore the expressions on the faces.

Vocal RTs were reliably shorter on congruent than on incongruent trials in the identity classification task in De Houwer et al.’s (1998) study, but not in their sex classification task. That is, for identity but not sex classification, the participants were faster when the valence of the facial expression matched the valence of the correct response (“positive” to faces with positive emotional expressions and “negative” to faces with negative emotional expressions) than when they did not match. Although there was no significant congruency effect in the sex classification task, RTs were shorter for the congruent than for the incongruent trials in both instances of this task. De Houwer et al. (1998) speculated that further studies with more statistical power might show this trend to be significant.

Our modified version of De Houwer et al.’s (1998) task employed exactly the same faces as those used in Experiment 1. In contrast to Experiments 1 and 2, the required response in Experiment 3 was vocal and semantically related to the irrelevant dimension. Participants classified emotional expression either by saying “male” for happy faces and “female” for fearful faces or by saying the converse (i.e., “female” for happy faces and “male” for fearful faces), and they classified the sex of the faces either by saying “fearful” for male faces and “happy” for female faces or by saying the converse. Thus, our Simon task is analogous to a Stroop task (MacLeod, 1991; Stroop, 1935) insofar as the relevant and the irrelevant dimensions in the Simon task can be semantically congruent (e.g., requiring responses of “male” to happy male faces and “female” to fearful female faces) or semantically incongruent (e.g., requiring responses of “male” to happy female faces and “female” to fearful male faces). The two paradigms differ to the extent that, in the Stroop but not in the Simon paradigm, the relevant stimulus dimension overlaps with the response set and the irrelevant feature dimension (see Kornblum & Lee, 1995).

Method

Participants. Twenty-eight psychology students (15 female) from University College Winchester took part in the experiment. All were 19 to 45 years of age ($M = 26.2$ years) and had normal or corrected-to-normal eyesight.

Stimuli and Apparatus. In Experiment 3, we employed the same stimuli as in Experiment 1—that is, 16 digitized black-and-white photographs of four male and four female faces from the FEEST database of facial expression stimuli (Young et al., 2002),

with both happy and fearful caricatured expressions. These stimuli were presented in the same dimensions and on the same computer and monitor as they were in Experiment 1, but this time the stimulus presentation and recording of RTs was controlled by PsyScope 1.2.4 experimental software (Cohen, MacWhinney, Flatt, & Provost, 1993). Vocal responses were recorded via a headset microphone connected to the computer via a response box and the keyboard.

Design. Each participant undertook two tasks: sex classification and emotion classification. Half of the participants classified the sex of the faces first and then classified the emotion of the faces; the other half of the participants completed the tasks in the reverse order. For sex classification, half of the participants were required to respond “happy” to male faces and “fearful” to female faces, whereas the remaining participants were asked to respond “fearful” to male faces and “happy” to female faces. For emotion classification, half of the participants were required to respond “male” to happy faces and “female” to fearful faces, whereas the remaining participants were asked to respond “male” to fearful faces and “female” to happy faces. Four versions of each experiment were created—one for each combination of task order and response mapping—and the participants were randomly assigned to them. Within each task, there were five blocks of 16 trials. Each of the 16 faces was presented once in each block, in a pseudorandom order.

Procedure. The participants were seated in a quiet room in front of the computer, approximately 60 cm from the monitor screen. Written instructions were presented on the screen and summarized orally by the experimenter. Each trial began with the warning GET READY! presented in the center of the screen for 1,000 msec. This was followed by a 1,000-msec blank screen interval, at the end of which a stimulus face appeared in the center of the screen. The face stimulus remained on the screen until the participant’s vocal response triggered its offset. Following a 200-msec interval, a central fixation cross appeared in the center of the screen until its offset was triggered by the experimenter’s pressing one of three keys on the keyboard, which coded the vocal response and any erroneous activation of the voice key. An interval of 1,000 msec preceded the beginning of the next trial. The participants were asked to respond as quickly and accurately as possible to either the sex or the emotional expression of the faces, depending on the task, and to ignore other aspects of the faces.

Results

Response times. Incorrect responses and trials in which the vocal response failed to activate the voice key accounted for 9.2% of the total responses and were not included in the RT analyses. Mean correct RTs were initially analyzed with a two-way ANOVA, with task (emotion classification vs. sex classification) and congruency (congruent vs. incongruent) as repeated measures variables. There were main effects of task [$F(1,27) = 39.98$, $p < .001$; $\eta^2 = .6$] and congruency [$F(1,27) = 28.27$, $p < .001$; $\eta^2 = .51$]. The main effect of task showed that responses were slower for emotion classification ($M = 900$ msec, $SEM = 23$) than for sex classification ($M = 773$ msec, $SEM = 20$). The main effect of congruency showed that responses were slower on incongruent ($M = 852$ msec, $SEM = 19$) than on congruent ($M = 821$ msec, $SEM = 19$) trials. These main effects were modified by a significant task \times congruency interaction [$F(1,27) = 4.69$, $p < .05$; $\eta^2 = .15$], which is depicted in Figure 4. As can be seen from this figure, the participants were slower to respond in the incongruent conditions than in the congruent conditions when classifying emotion, but not when

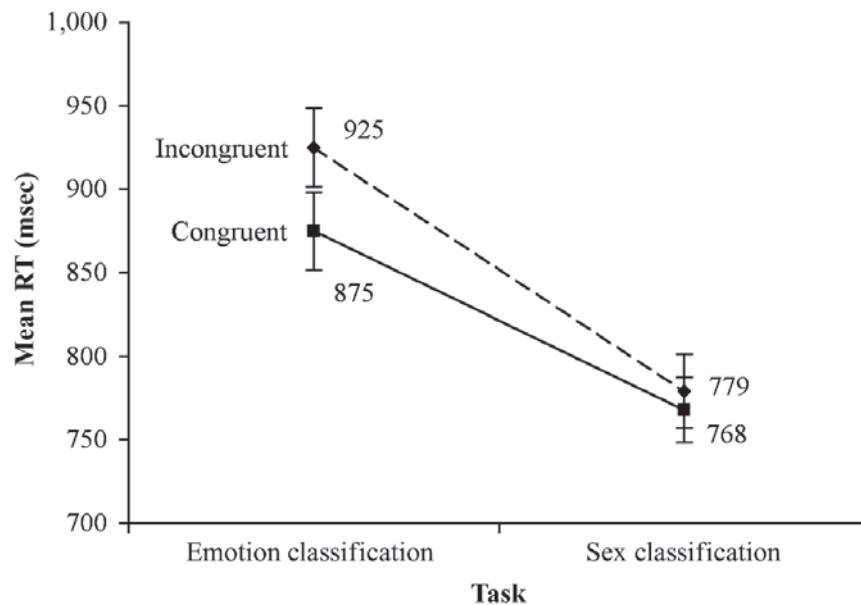


Figure 4. Mean response times (RTs) for classifying the emotion and sex of faces in the incongruent and congruent conditions (Simon paradigm) of Experiment 3.

they were classifying sex; furthermore, they were faster in classifying emotion than in classifying sex. This was confirmed by simple main-effect analyses of congruency for each task separately. When judging emotional expressions, the participants were significantly slower when the correct responses were incongruent with the irrelevant dimension (i.e., the sex) of the face ($M = 925$ msec, $SEM = 24$) than when they were congruent [$M = 875$ msec, $SEM = 23$; $F(1,27) = 27.4$, $p < .001$; $\eta^2 = .5$]. In the sex classification task, there was no significant difference between RTs to incongruent ($M = 779$ msec, $SEM = 22$) and to congruent [$M = 768$ msec, $SEM = 19$; $F(1,27) = 1.02$, $p > .1$; $\eta^2 = .036$] stimuli.

It was not possible to compare RTs for every combination of each relevant and each irrelevant dimension for each task in this experiment, as was done for Experiment 1. This is because in Experiment 3 the participants were not presented with all possible combinations of each relevant and each irrelevant dimension in each condition. Consider, for example, those participants who classified emotion by saying “male” for happy faces and “female” for fearful faces and classified sex by saying “happy” for male faces and “fearful” for female faces. These participants saw only happy male and fearful female faces in the congruent conditions of both the emotion classification and sex classification tasks, and they saw only fearful male and happy female faces in the incongruent conditions of both tasks.

Errors. We examined error rates through a subsidiary set of analyses. A two-way repeated measures ANOVA with variables of task (emotion classification vs. sex classification) and congruency (congruent vs. incongruent) was conducted to examine the percentage of errors for each condition in each task. There were main effects of task [$F(1,27) = 9.68$, $p < .01$; $\eta^2 = .26$] and congruency

[$F(1,27) = 26.29$, $p < .001$; $\eta^2 = .49$]. The main effect of task indicated that error rates were reliably higher for emotion classification ($M = 11.4\%$, $SEM = 1.5$) than for sex classification ($M = 6.0\%$, $SEM = 1.1$). The main effect of congruency indicated that error rates were reliably higher in the incongruent condition ($M = 10.8\%$, $SEM = 1.2$) than in the congruent condition ($M = 6.6\%$, $SEM = 1.0$). These main effects were modified by a significant task \times congruency interaction [$F(1,27) = 12.71$, $p < .005$; $\eta^2 = .32$], which is depicted in Figure 5. As this figure indicates, the participants made approximately twice as many errors in the incongruent ($M = 15.1\%$, $SEM = 2.1$) than in the congruent ($M = 7.7\%$, $SEM = 1.2$) condition when classifying emotion, but when classifying sex they showed essentially no difference in their error rates between these two conditions.

Discussion

The participants were significantly slower to classify emotion when the required response (“male” or “female”) was incongruent with the sex of the face in comparison with when the response and the sex of the face were congruent. No such significant congruency effect was evident for sex classification, however. The error rates followed the same pattern, with more errors in the incongruent than in the congruent condition for emotion classification but not for sex classification, confirming that the RT results were not due to a speed–accuracy trade-off. Thus, we have found a facial variant of a semantic Simon effect (De Houwer, 1998) but not of an affective Simon effect—at least not in a sex classification task, which corroborates De Houwer et al.’s (1998) finding. This asymmetry in the interactions between the processing of the emotion and the sex of faces mirrors the asymmetric dependency found in Experiments 1 and 2.

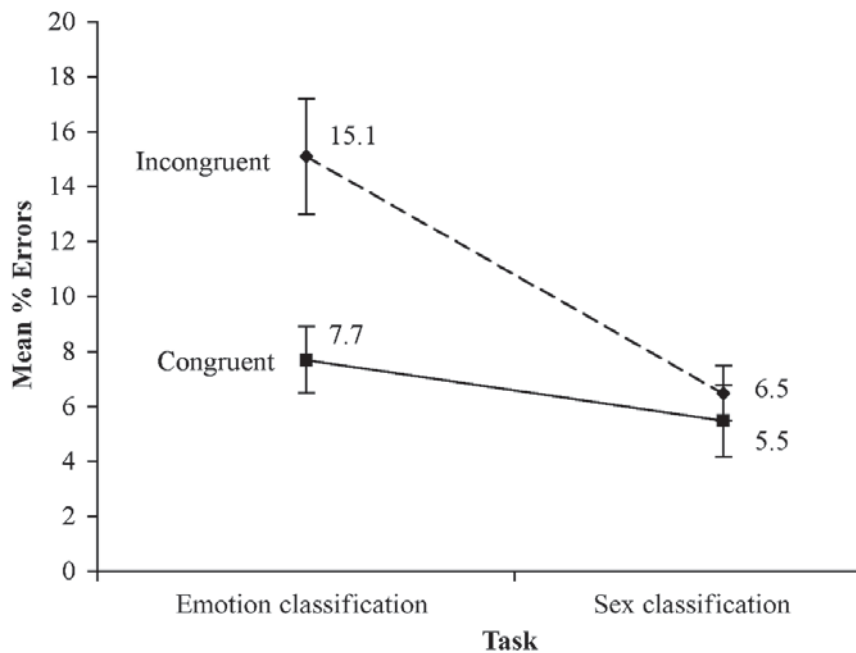


Figure 5. Mean percentage errors for classifying the emotion and sex of faces in the incongruent and congruent conditions (Simon paradigm) of Experiment 3.

It is worth noting that the sex classification task in Experiment 3 was completed more slowly than the emotion classification task in Experiment 1 (773 vs. 668 msec). This too counts against a purely speed-based account of the asymmetric interference, because the shorter RTs in the Experiment 1 emotion classification task than in the Experiment 3 sex classification task indicate that information about emotion must have been available during the sex classification task in Experiment 3, yet there was no interference in the latter task.

GENERAL DISCUSSION

This study provides convergent evidence, from two different interference paradigms, of an asymmetric relationship between the processes underlying performance on facial emotion and sex classification tasks. Emotion classification was significantly influenced by irrelevant variations in the sex of the faces (Experiments 1 and 2) and by the semantic relatedness of the vocal response to the sex of the faces (Experiment 3). Sex classification, on the other hand, was less influenced by both irrelevant variations in facial expression and the semantic relatedness of the vocal response to the facial expression. Nevertheless, Experiments 1 and 2 provided evidence that information concerning emotional expressions did impinge to a limited extent on sex classification performance. Specifically, the participants in Experiment 1 were slower to classify male faces as male when those faces portrayed fearful expressions in comparison with when they portrayed happy expressions, and in Experiment 2, sex classifications were slower in the control than in the orthogonal conditions for faces morphed across sex at the easiest of the difficulty levels.

The lack of Garner interference between sex and emotion classification in Le Gal and Bruce's (2002) study is not consistent with our results. There are several possible reasons for the different findings in the two studies, including the fact that Le Gal and Bruce employed surprised and angry expressions whereas we used expressions of fear and happiness. Although we have no specific reason to suspect that different emotions might produce different patterns of Garner interference, an avenue for future research will be to examine the robustness of the asymmetric pattern of interference by conducting studies in which different combinations of emotional expressions are employed.

Haxby et al.'s (2000) face-processing model provides a useful framework for interpreting the findings of asymmetric dependencies between identity processing and facial and speech expression processing, as reported by Schweinberger and Soukup (1998), Schweinberger et al. (1999), and Baudouin et al. (2002). In this model, invariant and changeable aspects of faces are represented through distinct processing routes. According to the interpretation of Haxby et al.'s model that we presented in the introduction, asymmetric Garner interference indicates some form of interaction between these distinct neurological systems. The direction of the asymmetric dependencies is consistent with the further plausible assumption that identity information (which is paradigmatically invariant) is much more likely to provide a reference for computing information about facial speech and emotion (which is paradigmatically changeable) than vice versa. The sex of faces, like their identity, is a relatively invariant property. Moreover, whereas the extent to which sex processing and identity processing are independent has been a matter of debate (see Young, 1998, for discussion) and a degree of

independence has been convincingly established (Bruce et al., 1987), a variety of recent evidence suggests that the sex and identity of faces are processed through a single route (see, e.g., Calder, Burton, Miller, Young, & Akamatsu, 2001; Dubois et al., 1999; Ganel & Goshen-Gottstein, 2002; Goshen-Gottstein & Ganel, 2000), as Haxby et al.'s (2000) conception of involvement of a neurological pathway in the analysis of invariant properties of faces would suggest. Thus, our finding of asymmetric Garner interference between sex and emotion classification is consistent with this particular reading of Haxby et al.'s model.

Our explanation of the asymmetry of the interference on facial variants of the Garner paradigm in terms of Haxby et al.'s (2000) model has so far relied on the intuition that changeable properties of faces, by their very nature, provide a less stable basis for further computations and are more context sensitive than relatively invariant properties of faces. There is a growing body of evidence consistent with this intuition. Our judgments of others' emotional states are influenced by our expectations regarding those people and the social contexts in which the emotions are expressed (see, e.g., Ekman & Friesen, 1971; Hess & Kirouac, 2000; Kirouac & Hess, 1999). For example, the expectation (at least in Western cultures) is that females will tend to be more facially expressive, express fear and sadness more readily, and smile more frequently than males, whereas males will tend to express anger more readily than females (Brody & Hall, 2000; Kring & Gordon, 1998; LaFrance, Hecht, & Paluck, 2003). Our suggestion is that the human brain's expression-processing systems are sensitive to such expectations insofar as they can draw upon information regarding the sex and identity of the person displaying the emotion.

According to Schweinberger and Soukup (1998) and Schweinberger et al. (1999), interactions at a perceptual level of processing are the presumed cause of asymmetric Garner interference in face processing. Similarly, Le Gal and Bruce (2002; see especially p. 242) assumed that it is the *absence* of interactions at a perceptual level of processing that explains their finding of no Garner interference between expression classification and sex classification. We have indicated how our conception of Haxby et al.'s (2000) model of face processing is consistent with such an interpretation, insofar as it suggests that such an interaction occurs between neurologically separate systems that represent the changeable properties of faces on one hand, and the relatively invariant ones on the other. Shortly, we will elucidate how this model is nevertheless closer in spirit and detail to models of face processing that posit parallel and independent processes than it is to paradigmatic "parallel-contingent" models.

The possibility of interaction at an even later stage of processing is suggested by the results of our third experiment, in which we used a facial variant of the Simon paradigm. Given the similarity of the Simon and Stroop paradigms, theoretical accounts of Stroop interference clearly have the potential to explain findings of interference on facial variants of the Simon paradigm. A common aspect

of theoretical accounts of Stroop interference is an appeal to response competition: The extent to which processing of the task-irrelevant dimension and that of the task-relevant dimensions compete for control of the participant's response depends on the degree of relatedness or overlap between the required response and the stimulus dimensions (see, e.g., Kornblum & Lee, 1995; MacLeod, 1991). Thus, the interference effect observed in Experiment 3 suggests that the locus of this Simon interference is at a response-selection stage of processing.

Given that the Simon paradigm employed in Experiment 3 yielded the same asymmetric pattern of interference as that observed with the Garner paradigm in Experiments 1 and 2, the challenge is to find a model that can account parsimoniously for this finding of convergence. Although an explanation in terms of response conflict might be appropriate for findings of Simon interference, it is difficult to see how Garner interference effects could be due to response conflicts, given that in that paradigm the irrelevant dimension is not mapped onto the same responses as is the relevant dimension. We note, though, that it is possible that an interaction between two or more stimulus dimensions can occur at a response-selection stage of processing for reasons other than response conflict. For example, the interested reader might want to consider Nosofsky and Palmeri's (1997a, 1997b) "exemplar-based random walk model" of stimulus categorization, which accounts for Garner interference effects in terms of post-perceptual operations at a memory-retrieval or response-selection stage of processing, and how this model might be extended to account for the asymmetric pattern of interference on facial variants of the Garner paradigm.

We have been promoting the utility of Haxby et al.'s (2000) model of face processing, our interpretation of which is consistent with the findings of asymmetric Garner interference between the processing of a face's sex and identity on the one hand, and the processing of facial emotion and speech expressions on the other. This model locates the source of that interaction at a perceptual stage of processing. To the extent that Simon interference indicates interaction at a later, response-selection stage of processing, Haxby et al.'s model would need to be extended in order to account for our finding of such an effect. Be that as it may, our finding of the same pattern of asymmetric interference in both our Garner and our Simon experiments indicates the pervasiveness of the principle that invariant dimensions of a stimulus are more useful referents for computing information about changeable aspects of that stimulus than vice versa. Moreover, Haxby et al.'s model preserves a high degree of functional separation between anatomically distinct processing routes, with only a relatively small degree of dependence or contingency, and then mostly in one direction and likely not mandatory. This model is a parallel-contingent model, but one that emphasizes the "parallel" much more than the "contingent." It is thus closer in spirit and detail to models that posit parallel and independent processes, such as that of Bruce and Young (1986), than to the paradigms

of parallel-contingent models (see discussion and references in Schweinberger et al., 1999, p. 1113). As such, Haxby et al.'s model is still consistent with a range of findings indicative of a high degree of functional separation (if not complete independence) between certain components of face processing, which we mentioned in the introduction—namely, between those components that encode changeable properties of faces versus those that encode invariant ones. Consider, for example, neuropsychological dissociations between identity and expression processing. Haxby et al.'s model is consistent with preserved expression processing despite impaired identity processing (due to damage to inferior but not to superior temporal cortex) and with preserved identity processing despite impaired expression processing (due to damage to superior but not to inferior temporal cortex). This is because there is relatively minimal functional reliance between the two anatomically separate components in the core system of Haxby et al.'s model, in comparison with the extent to which each of these components relies on the contributions of different components in the extended system. One intriguing prediction of this interpretation, given our aforementioned assumption about the direction of interaction between these components, is that those prosopagnosics who have preserved expression recognition abilities might be expected not to be able to use facial identity information (and perhaps not even information about the sex of faces) in their judgments of facial expression; such patients might not show the asymmetric pattern of interference in the facial variant of the Garner task, for example.

Conclusion

We have provided convergent evidence from two different interference paradigms that facial emotion classification can be significantly influenced by the sex of faces, but sex classification remains relatively impervious to facial emotion. These results are consistent with the hypothesis that information about invariant aspects of faces is much more likely to influence the processing of changeable aspects of faces than information about their changeable aspects is to influence the processing of their invariant aspects. The findings of asymmetric interference on facial variants of the Garner paradigm, including those reported here, suggest interactions at a perceptual level of processing, whereas our finding of asymmetric interference in the facial variant of the Simon paradigm suggests interaction at a memory-retrieval or response-selection stage of processing. A challenge for face processing models is therefore to show how the same asymmetric pattern of interference could occur at different stages of processing or, alternatively, how interactions at a single stage of processing could account for the same asymmetric pattern of interference across two paradigms that place apparently rather different processing demands on the subject. Although Haxby et al.'s (2000) model does not obviously explain the Simon interference effect, a strong point of this model is that it can account for asym-

metric Garner interference while preserving a high degree of functional separation between the processing routes for invariant and for changeable aspects of faces. Future research could capitalize on the ability of electrophysiological recordings to determine not just the relative timing of different perceptual judgments about faces (Müntz et al., 1998), but also their interactions (Rahman, Sommer, & Schweinberger, 2002), to address the issue of the locus of interference in the different paradigms.

REFERENCES

- BAUDOIN, J. Y., MARTIN, F., TIBERGHIE, G., VERLUT, I., & FRANCK, N. (2002). Selective attention to facial emotion and identity in schizophrenia. *Neuropsychologia*, *40*, 503-511.
- BENSON, P. J., CAMPBELL, R., HARRIS, T., FRANK, M. G., & TOVÉE, M. J. (1999). Enhancing images of facial expressions. *Perception & Psychophysics*, *61*, 259-274.
- BRODY, L. R., & HALL, J. A. (2000). Gender, emotion, and expression. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 338-349). New York: Guilford.
- BRUCE, V., ELLIS, H., GIBLING, F., & YOUNG, A. (1987). Parallel processing of the sex and familiarity of faces. *Canadian Journal of Psychology*, *41*, 510-520.
- BRUCE, V., & YOUNG, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305-327.
- CALDER, A. J., BURTON, A. M., MILLER, P., YOUNG, A. W., & AKAMATSU, S. (2001). A principal component analysis of facial expressions. *Vision Research*, *41*, 1179-1208.
- CALDER, A. J., ROWLAND, D., YOUNG, A. W., NIMMO-SMITH, I., KEANE, J., & PERRETT, D. I. (2000). Caricaturing facial expressions. *Cognition*, *76*, 105-146.
- CALDER, A. J., YOUNG, A. W., ROWLAND, D., & PERRETT, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society of London: Series B*, *264*, 919-925.
- CAMPBELL, R., BROOKS, B., DE HAAN, E., & ROBERTS, T. (1996). Dissociating face processing skills: Decision about lip-read speech, expression, and identity. *Quarterly Journal of Experimental Psychology*, *49A*, 295-314.
- CAMPBELL, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lip-reading: A neurological dissociation. *Brain*, *109*, 509-521.
- COHEN, J., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, *25*, 257-271.
- DE HOUWER, J. (1998). The semantic Simon effect. *Quarterly Journal of Experimental Psychology*, *51A*, 683-688.
- DE HOUWER, J., CROMBEZ, G., BAEYENS, F., & HERMANS, D. (2001). On the generality of the affective Simon effect. *Cognition & Emotion*, *15*, 189-206.
- DE HOUWER, J., & EELEN, P. (1998). An affective variant of the Simon paradigm. *Cognition & Emotion*, *12*, 45-61.
- DE HOUWER, J., HERMANS, D., & EELEN, P. (1998). Affective Simon effects using facial expressions as affective stimuli. *Zeitschrift für Experimentelle & Angewandte Psychologie*, *45*, 88-98.
- DUBOIS, S., ROSSION, B., SCHILTZ, C., BODART, J. M., MICHEL, C., BRUYER, R., & CROMMELINCK, M. (1999). Effect of familiarity on the processing of human faces. *NeuroImage*, *9*, 278-289.
- EIMAS, P. D., TARTTER, V. C., MILLER, J. L., & KEUTHEN, N. J. (1978). Asymmetric dependencies in processing phonetic features. *Perception & Psychophysics*, *23*, 12-20.
- EKMAN, P., & FRIESEN, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality & Social Psychology*, *17*, 124-129.
- EKMAN, P., & FRIESEN, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- EKMAN, P., & FRIESEN, W. V. (1978). *The facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

- ELLIS, A. W. (1989). Neuro-cognitive processing of faces and voices. In A. W. Young & H. D. Ellis (Eds.), *Handbook of research on face processing* (pp. 207-215). Amsterdam: North-Holland.
- GANEL, T., & GOSHEN-GOTTSTEIN, Y. (2002). Perceptual integrality of sex and identity of faces: Further evidence for the single-route hypothesis. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 854-867.
- GARNER, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, **8**, 98-123.
- GOSHEN-GOTTSTEIN, Y., & GANEL, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: Evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1198-1214.
- GREEN, K. P., & KUHL, P. K. (1991). Integral processing of visual place and auditory voicing information during phonetic perception. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 278-288.
- HASSELMO, M. E., ROLLS, E. T., & BAYLIS, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, **32**, 203-218.
- HAXBY, J. V., HOFFMAN, E. A., & GOBBINI, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, **4**, 223-233.
- HESS, U., & KIROUAC, G. (2000). Emotional expression in groups. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 368-381). New York: Guilford.
- HUMPHREYS, G. W., DONNELLY, N., & RIDDOCH, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia*, **31**, 173-181.
- KIROUAC, G., & HESS, U. (1999). Group membership and the decoding of nonverbal behavior. In P. Philippot, R. S. Feldman, & E. J. Coats (Eds.), *The social context of nonverbal behavior* (pp. 182-210). Cambridge: Cambridge University Press.
- KORNBLUM, S., & LEE, J. W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 855-875.
- KRING, A. M., & GORDON, A. H. (1998). Sex differences in emotion: Expression, experience, and physiology. *Journal of Personality & Social Psychology*, **74**, 686-703.
- LAFRANCE, M., HECHT, M. A., & PALUCK, E. L. (2003). The contingent smile: A meta-analysis of sex differences in smiling. *Psychological Bulletin*, **129**, 305-334.
- LE GAL, P. M., & BRUCE, V. (2002). Evaluating the independence of sex and expression in judgments of faces. *Perception & Psychophysics*, **64**, 230-243.
- LU, C.-H., & PROCTOR, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, **2**, 174-207.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MELARA, R. D., & MOUNTS, J. R. W. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, **21**, 627-645.
- MÜNTE, T. F., BRACK, M., GROOTHEER, O., WIERINGA, B. M., MATZKE, M., & JOHANNES, S. (1998). Brain potentials reveal the timing of face identity and expression judgments. *Neuroscience Research*, **30**, 25-34.
- NOSOFSKY, R. M., & PALMERI, T. J. (1997a). Comparing exemplar-retrieval and decision-bound models of speeded perceptual classification. *Perception & Psychophysics*, **59**, 1027-1048.
- NOSOFSKY, R. M., & PALMERI, T. J. (1997b). An exemplar-based random walk model of speeded classification. *Psychological Review*, **104**, 266-300.
- PARRY, F. M., YOUNG, A. W., SAUL, J. S., & MOSS, A. (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical & Experimental Neuropsychology*, **13**, 545-558.
- PUCE, A., ALLISON, T., BENTIN, S., GORE, J. C., & MCCARTHY, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, **18**, 2188-2199.
- RAHMAN, R. A., SOMMER, W., & SCHWEINBERGER, S. R. (2002). Brain-potential evidence for the time course of access to biographical facts and names of familiar persons. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 366-373.
- SCHWEINBERGER, S. R., BURTON, A. M., & KELLY, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. *Perception & Psychophysics*, **61**, 1102-1115.
- SCHWEINBERGER, S. R., & SOUKUP, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology: Human Perception & Performance*, **24**, 1748-1765.
- SIMON, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. Gilmour Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31-86). Amsterdam: North-Holland.
- SIMON, J. R., & ACOSTA, E., JR. (1982). Effect of irrelevant information on the processing of relevant information: Facilitation and/or interference? The influence of experimental design. *Perception & Psychophysics*, **31**, 383-388.
- SIMON, J. R., & RUDELL, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, **51**, 300-304.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- UMILTÀ, C., & NICOLETTI, R. (1990). Spatial stimulus-response compatibility. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 89-116). Amsterdam: North-Holland.
- UMILTÀ, C., & NICOLETTI, R. (1992). An integrated model of the Simon effect. In J. Alegria, D. Holender, J. Junca de Morais, & M. Radeau (Eds.), *Analytic approaches to human cognition* (pp. 331-350). Amsterdam: North-Holland.
- YOUNG, A. W. (1998). *Face and mind*. Oxford: Oxford University Press.
- 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., PERRETT, D. I., CALDER, A. J., SPRENGELMEYER, R., & EKMAN, P. (2002). *Facial expressions of emotion: Stimuli and tests (FEEST)*. Bury St. Edmunds: Thames Valley Test Company.

(Manuscript received June 30, 2003;
revision accepted for publication February 11, 2005.)