

Evaluating the independence of sex and expression in judgments of faces

PATRICIA M. LE GAL and VICKI BRUCE
University of Stirling, Stirling, Scotland

Face recognition models suggest independent processing for functionally different types of information, such as identity, expression, sex, and facial speech. Interference between sex and expression information was tested using both a rating study and Garner's selective attention paradigm using speeded sex and expression decisions. When participants were asked to assess the masculinity of male and female angry and surprised faces, they found surprised faces to be more feminine than angry ones (Experiment 1). However, in a speeded-classification situation in the laboratory in which the sex decision was either "easy" relative to the expression decision (Experiment 2) or of more equivalent difficulty (Experiment 3), it was possible for participants to attend selectively to either dimension without interference from the other. Qualified support is offered for independent processing routes.

A number of cues (e.g., gait, clothing, behavior, the presence or absence of breasts) may be used to determine someone's sexual identity. We also use a variety of cues to determine someone's internal emotional state. Both types of information are also clearly signaled by the face and can be interpreted accurately by perceivers from the face alone. However, the face also provides information about other personal information (e.g., identity, age, race) and sends out other communicative signals (e.g., facial speech and gaze). How are all these different messages from the face deciphered?

Currently, there is substantial evidence to suggest that different types of information may be extracted from faces independently (for a review, see Young, 1998). There is strong evidence, for example, reviewed later in this article, that the emotional expression of a face is analyzed independently of information about identity. The independence of major routes, such as expression analysis and identity, can be contrasted with the dependence seen within a particular route, between sequential stages of processing (e.g., assessing the familiarity of a face is an early stage and naming a relatively late stage within the identification route in Bruce & Young's, 1986, model). In this study, we explored the hypothesis that expression and sex¹ are analyzed via independent routes, but we did not address the question of any stages that might be involved within specific routes. Expression and sex might be expected to be independent, since, in some models, sex classification forms an early stage in the identification route (e.g., H. D. Ellis, 1986), whereas in other analyses, sex classification is proposed to

be independent of identification but also logically independent from emotional expression. Experimental and neuropsychological evidence relating to the proposed processing independence of sex and expression routes is presented later in this introduction.

In this study, we employed Garner's (1974, 1976) selective attention paradigm to investigate whether or not processing resources are shared when participants make sex and expression judgments to faces. Garner's paradigm, introduced in detail later, has recently been used to test the theory of independent face processing routes by Schweinberger and colleagues (Schweinberger, Burton, & Kelly, 1999; Schweinberger & Soukup, 1998). The independence of sex and expression processing is of particular interest, since analyses in this instance may be applied to the same information source. Some of the features in the region of eyebrows and eyes, which are useful for distinguishing male faces from female faces, vary considerably when faces express surprise and anger. A change in expression from anger to surprise might be expected to "feminize" a face, for example, and lead to uncertainty in the classification of a male face as male.

In the remainder of the introduction, we briefly outline what is understood about the information used to categorize faces by sex and expression. We then review models of face processing in which different kinds of categorization are achieved by independent routes. Finally, we introduce Garner's selective attention methodology, which we apply to the question of whether sex and expression decisions are made independently.

How Do We Tell Male Faces From Female Faces?

Potential cues to discriminate male faces from female faces include superficial texture (e.g., stubble), relational information (e.g., nose size relative to the chin), and three-dimensional information (e.g., brow and chin protuberance), in addition to simple features (e.g., brow width or

This research was conducted when the first author was a PhD student funded by the United Kingdom Medical Research Council. The authors thank the referees for comments that have helped to improve the paper. Correspondence should be addressed to V. Bruce, Department of Psychology, University of Stirling, Stirling, Scotland FK9 4LA (e-mail: vicki.bruce@stir.ac.uk).

brow-to-eye distance). Laboratory studies of human perception of faces and statistical analysis of actual variation in face images suggest that it is the combination of a number of different discriminating cues that allows us to classify faces by sex effortlessly and accurately, even when hairstyle is concealed (e.g., Brown & Perrett, 1993; Bruce et al., 1993; Burton, Bruce, & Dench, 1993; O'Toole, Abdi, Deffenbacher, & Valentin, 1993).

Distance between the eyebrows and eyelids is a measure that was found to be one of the better discriminators between male and female faces (Burton et al., 1993). This distance is smaller in males than in females. Campbell, Wallace, and Benson (1996) found little overlap between scores for brow-to-lid distances for male and female faces looking straight ahead. Campbell et al. (1996) went on to show that speeded sex decisions to faces with gaze averted were slower than when the faces were looking straight ahead and that male faces were judged to be less masculine with eyes averted. Averting gaze increases the brow-to-lid distance, and Campbell et al. (1996) suggest that this is a particularly salient feature for males. The increased distance in faces looking down slows responses. Similar findings were also reported in a later study (Campbell, Benson, Wallace, Doesbergh, & Coleman, 1999) in which the men and women whose faces were used as stimuli were asked to raise and lower their brows. These last findings are of particular relevance to the experiments reported here. If brow-to-lid distance also varies as the result of a change in expression (which it does when the face changes from being angry to being surprised), then sex decisions might also be facilitated or delayed depending on what expression is shown.

How Do We Tell Angry Expressions From Surprised Expressions?

For expression decisions, posed and spontaneous expressions, photographs of actors or line drawings, and even schematic faces have been used in attempts to identify the underlying structural information mediating recognition. The most comprehensive measurement system to describe the way in which facial muscles and features change during the expression of emotions is called the *facial action coding system* (FACS) and was developed by Ekman and Friesen (1976, 1978; cited in Rosenberg, 1997). FACS allows the description of visible changes in the face using 44 action units (AUs), together with the underlying muscle groups involved. FACS-coded events can be classified into emotion categories that match facial events with emotional events. In a study of AUs that were generated when actors portrayed felt emotions, Gosselin, Kirouac, and Doré (1997) found that the following were generated for anger and surprise: For anger, the "brow lowerer" (lowers eyebrows and pulls them together), the "upper lid raiser" (raises the upper lid exposing more of the upper portion of the eyeball), the "lip tightener" (tightens the lips making them appear more narrow), the "jaw drop" (parts lips so that the space between the teeth can be seen), the "lower lip depressor" (pulls the lower lip down and flattens the chin

boss), and the "lip stretcher" (stretches lips horizontally) were frequently generated. For surprise, the "upper lid raiser," the "jaw drop," the "inner brow raiser" (raises only the inner part of the eyebrow), the "outer brow raiser" (raises only the outer part of the eyebrow), the "mouth stretch" (stretches mouth open), and the "lips part" (parts lips to a limited extent) were common. A great deal of displacement therefore occurs around the eyes (and the mouth) during the expression of both anger (brows down) and surprise (brows up).

So, both expression and sex discrimination rely in part on variations in brow-to-lid distance. When eyebrows are raised, this indicates that the face is surprised, but this should also indicate that the face is female. When eyebrows are lowered, this indicates the face is angry, but this should also indicate that the face is male.

Independent Processing Routes

Contemporary models of face recognition (e.g., Bruce & Young, 1986) suggest independent information processing routes for the derivation of functionally different information from faces. It is not necessary, for example, to recognize someone's identity to know what emotion they are expressing or to read their lips. Different types of analyses appear to be carried out independently on the same stimulus. In some cases, this functional independence builds on the logical independence of information sources for the different tasks: For example, although the transient shape of the mouth may be important to distinguish expressions, the identification of a person may rely on features such as hairstyle that are irrelevant for expression perception. Functional independence of information processing routes is particularly interesting when independent analyses are shown to be applied to the same information source. For example, the functional independence demonstrated between expression analysis and lipreading (Campbell, Landis, & Regard, 1986) is intriguing, given that both expressions and facial speech rely on momentary configurations of the mouth. In this paper, we report new experiments that examined whether sex and expression information are processed independently. This is of interest because of the extent to which the features in the region of eyebrows and eyes that are useful for distinguishing male faces from female faces overlap perceptually with those that are used to distinguish anger from surprise.

Evidence for Independence

On what grounds might we propose that expression and sex classification should be independent? Certainly, there is strong evidence for independence of expression analysis and identification, and, on some accounts, sex classification forms a stage within the identification route (H. D. Ellis, 1986). Experimental evidence for the independence of identity and expression comes from a study by Young, McWeeny, Hay, and A. W. Ellis (1986). Familiarity facilitated decisions to pairs of faces when the participant was asked to decide whether or not the faces were of the same person, but familiarity had no effect on the time

taken to match pairs of faces for expression. Neuropsychological evidence comes from double dissociations between identification and expression analysis in brain-injured patients (e.g., see Young, Newcombe, De Haan, Small, & Hay, 1993) and is enhanced by dissociations observed in neurophysiological and neuroimaging studies (e.g., Rolls, 1992; Sergent, Ohta, MacDonald, & Zuck, 1994).

However, there is also evidence that sex classification is in turn independent of identification. For example, A. W. Ellis, Young, and Flude (1990) investigated the effect of familiarity, sex, and expression decisions made during a Phase 1 task (first exposure to stimulus faces) on later decisions to familiarity, sex, and expression in a Phase 2 task (second exposure). They found that any of the three Phase 1 decisions speeded later familiarity decisions, that neither familiarity nor sex Phase 1 decisions speeded later sex decisions, and that neither familiarity nor expression Phase 1 decisions speeded later expression decisions. The results of this study are interpreted as supporting the idea that priming is confined to the system that mediates identity and that sex decisions do not form part of this route. Further evidence for the independence of identity and sex decisions comes from a study by Bruce, H. D. Ellis, Gibling, and Young (1987; see also Bruce, 1986). Whether a face was easy or difficult to classify as male had no effect on how quickly participants judged a face as familiar or not. If sex classification formed an early stage of an identification route, then such an effect should have been observed.

There is also some evidence that sex information is independent from other facial dimensions. In an audiovisual study to investigate the independence of sex information and the analysis of facial speech, Green, Kuhl, Meltzoff, and Stevens (1991) found that mismatched male and female voices to female and male videotaped faces did not interfere with speech comprehension. From this experimental evidence then, it is likely that sex and expression are independent from identity analysis and that sex information is independent from facial speech analysis, but it is less certain that sex and expression are independent from each other. The appropriate experimental comparisons have not yet been made.

Evidence for the neuropsychological dissociation of sex and expression processing is nonexistent. Neither are there unequivocal reports of the dissociation of sex and identity. A possible reason for this lack of knowledge may be that prosopagnosic patients have not always been tested for the recognition of face gender, since the recognition of familiar people (and what sex they are, through knowledge about that person) from voice and gait may remain unimpaired. This does not always apply, and there have been reports of severely prosopagnosic patients who also cannot tell the gender of a face (e.g., Cole & Perez-Cruet, 1964). The tests, however, have not always been rigorous, and Cole and Perez-Cruet mention that their patient was impaired at gender discrimination on the basis of his inability to recognize that characters in a soap opera on television with the sound turned off were women: Their con-

cern was more with his inability to recognize familiar people. The patients in Tranel, Damasio, and Damasio's (1988) study were also impaired at identity but retained intact facial expression, age, and gender recognition. A similar dissociation between identity and expression and gender was also reported by Humphreys, Donnelly, and Riddoch (1993), whose patient (G.K.) was relatively unimpaired at identifying famous faces but was poor at the expression and gender of unfamiliar faces. There have also been numerous studies supporting the double dissociation of expression and identity (e.g., Calder et al., 1996; Parry, Young, Saul, & Moss, 1991). The double dissociation of expression and facial speech analysis has also been reported (Campbell et al., 1986). From this neuropsychological evidence then, it is likely that sex and expression are independent from identity analysis and that expression analysis is independent from facial speech analysis, but, again, it is less certain that sex and expression are independent from each other.

In summary, there is strong evidence for the independence of certain processing routes, particularly expression and identity, but the evidence addressing the specific independence of sex and expression processing is sparse. There is evidence to suggest that both are independent from identity analysis, but none to offer a direct comparison of each with the other.

Recent Evidence Suggests More Complex Relationships

Although much of the observed functional independence is consistent with logical separations in the structural information needed to perform different tasks, recent studies suggest that the relationship between different types of categorization is more complex than simple independence. For example, Schweinberger and Soukup (1998, Experiments 1 and 2; see also Schweinberger et al., 1999) suggest asymmetric independence between identity and expression. They found that, although it is possible in a speeded-classification task (after Garner, 1974, 1976) to attend to identity without any interference from expression, it may not be possible to attend selectively to expression. These findings are at variance with an earlier study by Etcoff (1984), who also used Garner's (1974) selective attention methodology. Participants were asked to sort cards of faces into separate identity (Person A, Person B) or expression (happy or sad) categories. Etcoff found that there was no time cost when information irrelevant to the task (e.g., identity, when classifying expressions as happy or sad) was kept constant relative to when it was varied orthogonally. Schweinberger and Soukup suggest that procedural differences, including the increased sensitivity of their method of measuring reaction time, may account for this difference.

Further evidence of complexity comes from studies on other functional facial information. Schweinberger and Soukup (1998, Experiments 3–7), again using Garner's speeded-classification format, found an asymmetric relationship between identity and facial speech similar to that

found for identity and expression. That is, it was possible to ignore the different speech actions depicted when identifying a face, but it was not possible to ignore identity when classifying speech. This latter finding supports a study by Walker, Bruce, and O'Malley (1995), using "McGurk" blend illusions (McGurk & MacDonald, 1976), where visible facial speech influences what a speaker is heard to say. Walker et al. found that when the identities of the faces and voices that were combined were incongruent, participants familiar with the speakers reported fewer blends than did participants who were unfamiliar with them, again showing an influence of identity on facial speech processing.

Garner's (1974) selective attention paradigm offers an advantage over unidirectional experimental methods (e.g., Campbell et al., 1999; Campbell et al., 1996), in that the procedure is bidirectional: Both processing routes under investigation may be examined in the same experiment. Current measurement techniques also allow the detection of small, systematic differences in reaction time (RT) between conditions. For these reasons, the focus of the present study was an exploration of sex and expression judgments using Garner's speeded-classification method. A detailed consideration of Garner's methods follows in the next section.

Garner's Selective Attention Paradigm

Garner (1974, 1976) was interested in determining which stimulus properties could be said to occupy their own processing dimension and which (even though they might be defined by the experimenter as occupying another dimension) require shared processing resources. If selective attention is possible in the presence of irrelevant variation, then the two "dimensions" under investigation can be declared independent or "separable." If selective attention is not possible, then the dimensions show "integrality." His work was focused on seemingly simple dimensions, such as shape and color. The present study was concerned with testing the idea of functional independence between the stimulus properties of faces that determine whether the underlying emotion of the bearer is anger or surprise and those that determine their maleness or femaleness. In this way, the Garner paradigm was extended to the seemingly much more complex dimensions of expression and sex.

In a typical Garner experiment, participants are required to make speeded two-choice classificatory judgments to four types of stimuli representing the crossing of two different dimensions. When the dimensions are expression and sex, participants are instructed to ignore sex when deciding whether the face is angry or surprised or are instructed to ignore expression when deciding whether the face is male or female. The stimuli are presented in three different experimental conditions termed the *control*, *orthogonal*, and *correlated* conditions. In a between-subjects design, one group of participants makes sex classifications for all three conditions and one group makes expression classifications for all three conditions.

In the control condition, the stimuli vary along only the relevant dimension and the irrelevant dimension is held con-

stant. A blocked presentation of stimuli is crucial to Garner's method. For example, in the control condition, when the relevant dimension is sex, the participant classifies the face as either male or female when all the faces in the first block are one of the expressions only (e.g., angry). In the second block, all the faces seen are the other version of the expressions (e.g., surprised). In the orthogonal condition, stimuli vary orthogonally along both the relevant dimension and the irrelevant dimension. When the relevant dimension is sex, the faces in both blocks can be either angry or surprised. In the correlated condition, there is covariation of the relevant and irrelevant dimensions. When the relevant dimension is sex, in this condition, all the faces in the first block that are male will also show just one of the expressions (e.g., surprised), as will the female faces seen (e.g., angry), with this contingency reversed in the second block. The use of blocked presentation ensures all the faces are seen an equal number of times. Even though the participant is asked to ignore irrelevant variation in all instances, the influence of the irrelevant dimension when kept constant (control), varied unexpectedly (orthogonal), or made contingent on each version of the relevant dimension (correlated) can be estimated.

Garner (1976) outlines four types of dimensional interaction and their effects on different perceptual processes (RT facilitation, selective attention). He gives experimental evidence to back up the four possible outcomes and indicates that a further two could be predicted. The four for which there is evidence are the following: (1) *integral* (facilitation with correlated dimensions and interference with orthogonal dimensions), (2) *configural* (no facilitation with correlated dimensions and interference with orthogonal dimensions), (3) *separable* (no facilitation with correlated dimensions and no interference with orthogonal dimensions), and (4) *asymmetric separable* (facilitation with correlated dimensions and asymmetric interference with orthogonal dimensions).

Two further possible outcomes were mentioned by Garner for which, at that time, he could cite no evidence. These were facilitation with correlated dimensions and no interference with orthogonal, named as *optional separable* (Garner notes this could be called the "ideal interaction," since the participant can use redundancy in the correlated condition to facilitate discrimination and ignore it when it is irrelevant to the classification task), and a lack of facilitation with correlated dimensions and asymmetric selective attention with orthogonal ones, named as *asymmetric configural*.

Garner claims that one of the characteristics of integral dimensions is that the stimulus is seen as an integrated whole. This creates a cost to the perceiver in the increased uncertainty of the orthogonal condition. Although Garner's explanation of the processes underlying the perceptual experience of the viewer has been questioned by later researchers (e.g., attribute level processing driven by context rather than holistic processing; Melara & Marks, 1990), the view that slowed RTs in the orthogonal condition demonstrate a failure of selective attention is generally accepted (e.g., Schweinberger & Soukup, 1998).

It is not the case, however, that a redundancy gain (shown by speeded reaction times in the correlated condition relative to the control condition) is unambiguous support for integrality. Garner claims that, by some process of stimulus redefinition, it is possible for the participant to benefit from the additional or "redundant" information available here. Although a redundancy gain is consistent with integrality, this is not the only reason faster RTs in the correlated condition might be expected. Participants responding to separable dimensions could also use the redundant information here. For example, if dimensions are separable, participants may be free to choose which of the dimensions they process first, choosing either the easier dimension to discriminate or the one they prefer, which could lead to a gain. The conditions for such a "selective serial processing" strategy are outlined by Biederman and Checkosky (1970) and Felfoldy and Garner (1971). The presence therefore, of a redundancy gain is best considered additional evidence for integrality, with the critical comparison to be made between control and orthogonal conditions (see also Ashby & Townsend, 1986).

The main aim of the present study was to investigate the relationship between the sex and expression dimensions using Garner's selective attention classification methodology. Independence of processing routes should allow selective attention to each dimension, which would therefore give rise to the pattern described as *separable* by Garner. However, Schweinberger and colleagues (Schweinberger et al., 1999; Schweinberger & Soukup, 1998) have applied the Garner method to the dimensions of expression and identity to yield evidence of *asymmetric* rather than complete separability, and so it is possible that a similar relationship might be observed for the dimensions of expression and sex. Furthermore, since variation in expression and sex both affect the same features of the face (e.g., in brow-to-lid distance), there could be grounds for supposing that these dimensions would be *integral* (or *configural*) in Garner's terms. Before describing the Garner experiments, we report an experiment in which we examined directly how expression moderates masculinity using a (unidirectional) rating task. If expression influences rated masculinity as predicted, then the two dimensions might be expected to show integrality in the Garner experiments that follow.

EXPERIMENT 1

Experiment 1 was conducted as a preliminary investigation, prior to the Garner experiments (Experiments 2 and 3). If sex and expression information are independent, then rated masculinity will not be sensitive to expression (surprise or anger). If variations in expression affect the perceived masculinity of faces, then a difference in ratings due to expression is expected. Male faces especially should appear less masculine when surprised, since the increased brow-to-lid distance is a cue associated with female faces. Female faces should look less feminine when angry, since the reduced brow-to-lid distance is characteristic of male faces.

Method

Participants. The participants were recruited while viewing a gallery exhibition. Twenty-four members of the public (8 men, 16 women) agreed to take part.

Stimuli and Apparatus. Angry and surprised expressions from 5 male and 5 female actors made up the stimulus set (Figure 1). Stimulus faces were checked for consistency of recognition in a previous assessment procedure (Le Gal, 1999). The mean expression recognition scores and standard deviations (*SDs*) for each subset of faces are as follows: female angry, 93% (*SD* = 7.5); male angry, 94% (*SD* = 8.0); female surprised, 87% (*SD* = 9.8); male surprised, 83% (*SD* = 8.7). Stimuli were printed out, a single image per page, and made into small booklets for presentation to raters.

Preliminary analyses were carried out to check that the brow-to-lid distances of male and female faces of this stimulus set were different from each other. Measurements were made in the following manner using the method proposed by Campbell et al. (1996) (Figure 2). The vertical distance between the eyelid and the underside of the brow immediately above the pupil divided by the diameter of the iris is called *ratio minA*. The vertical distance between the corner of the eye at either the inner canthus or the outer canthus (whichever is greater) and the underside of the brow divided by the diameter of the iris is called *ratio maxA*. It was found that neutral male and female faces of this set are distinguished clearly from each other with regard to this feature: mean minA female = 1.12, (*SD* = 0.18), mean minA male = 0.70 (*SD* = 0.16) [$t(4) = 4.02, p < .02$, two-tailed]; mean maxA female = 1.72 (*SD* = 0.17), mean maxA male = 1.23 (*SD* = 0.14) [$t(4) = 4.95, p < .01$, two-tailed]. There was no overlap at all between the scores from the sets of female and male faces on either measure.

To see whether sex differences for this feature were sustained for expressive faces, two separate analysis of variance (ANOVAs) were carried out on minA and maxA ratios, with sex (male, female) and expression (angry, surprised) as factors. For the minA ratio, there was a main effect of expression [$F(1,8) = 109.1, p < .001$], with surprise ratios higher. There was no main effect of sex ($p > .1$), although the trend was for female brow-to-lid distances to be larger. For the maxA ratio, there was also a main effect of expression [$F(1,8) = 214.8, p < .01$], with surprise ratios higher. There was again no main effect of sex ($p > .1$), although the trend was again for female distances to be larger. There was no significant interaction in either of these analyses. Unlike the neutral faces, there was some overlap in the measures derived from male and female faces when these were expressing.

These measurements show that differences between male and female faces should be clearly visible if the viewer looks at the brow-to-lid distance when these faces are expressively neutral. This distinction between male and female faces is not so sharp if this feature is considered when the faces are distorted by angry and surprised expressions. Therefore, it was predicted that ratings of masculinity and femininity should be influenced by variations in expression.

Design. A 2×2 within-subjects design was used, with sex of stimulus face (male, female) and expression (angry, surprised) at two levels. Ratings were the dependent measure.

Procedure. The participants were invited to rate the faces (20 in total), according to how feminine or masculine they appeared. Ratings ranged from 1 (*very feminine*) to 15 (*very masculine*). The participants were asked to try to ignore the expressions shown on the faces. Each participant viewed the faces sequentially in two blocks. Half the participants saw females first, and half saw males first. The faces were ordered randomly within blocks, and two different random orders were used. Equal numbers of participants viewed each order. The task took approximately 5 min.

Results and Discussion

Rating scores are shown in Figure 3. Male faces were thought to be an average of 2.4 scale intervals less mascu-



Figure 1. Stimulus set for Experiment 1. Far left column (from top): female angry Faces 2, 7, 9, 16, and 17. Center left: female surprised; Center right: male angry Faces 21, 22, 24, 32, and 41. Far right: male surprised. Face numbers identify which faces from Le Gal (1999) were selected for use.

line and female faces were thought to be an average of 1.9 intervals more feminine when the faces were surprised relative to when they were angry. A 2×2 ANOVA revealed main effects of sex [$F(1,23) = 46.8, p < .0001$] and expression [$F(1,23) = 92.5, p < .0001$] but revealed no interaction ($p > .1$).

Experiment 1 provided some evidence to suggest that angry and surprised expressions may modify judgments

of masculinity and femininity and that this response is consistent with the way in which brow-to-lid distances vary when these measurements are taken from expressively neutral faces.² Having found that physical variation of brow-to-lid distance does impact on gender ratings, we examined whether there is interaction between these two types of coding using Garner's interference task in Experiment 2.

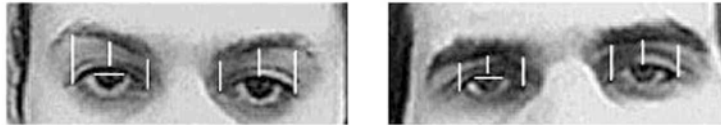


Figure 2. Measurements of the eyelid that distinguish female (left) from male (right) expressively neutral faces (after Campbell et al., 1996).

EXPERIMENT 2

A test of the modification of sex judgments by expression was carried out by placing the stimuli in an RT experiment using the selective attention paradigm described by Garner (1974, 1976). If, as Campbell et al. (1996) suggest, a large eyelid area is perceptually important for the classification of sex, an experiment that compares sex classifications with expression classifications for stimuli in which this feature is contrasted should reveal interference if it is impossible to attend selectively to just the relevant dimension.

It was predicted, on the basis of the results of Experiment 1, that sex decisions would be influenced by variation in expression. It was not known whether expression classification would be influenced by variation in sex. Although there is no necessary equivalence between how sex and identity information are processed, expression classification has been found susceptible to interference from identity information in this speeded-classification format

(Schweinberger et al., 1999; Schweinberger & Soukup, 1998).

One group of participants made sex classificatory decisions, and another group made expression classificatory decisions. A mixed analysis of the entire data set was used to determine whether the effects of the Garner conditions differed for the two tasks (groups). Irrespective of whether this interaction was observed, however, more detailed analyses were planned within each task to investigate the detailed effects of the different dimension combinations within each Garner condition.

Method

Participants. Twenty-seven women volunteered as participants, and each received either course credit or £3 for their participation.³ The data from 24 participants, with a mean age of 21.5 years ($SD = 5.8$) were included in the data set, since 3 participants failed to achieve at least 50% correct responses within the time window in each cell for the expression classification task.

Stimuli and Apparatus. The face stimuli used were as for Experiment 1. They were presented using SuperLab software on a light background on a Macintosh Quadra 800. Stimulus size measured 6×9.9 cm, which subtended an angle of 5.7° horizontally and 9.4° vertically at a viewing distance of 0.6 m. Each participant's head was kept level and in one place with the use of a chinrest.

Design. The experiment had a 2×3 mixed design, with the between-subjects variable of task at two levels (sex classification, expression classification) and condition at three levels (correlated, control, orthogonal). Dependent measures were RTs and errors.

Procedure. The procedure for the sex classification task was as follows: The 12 participants in this group were told that the task involved making a decision as quickly and as accurately as possible as to whether the face presented on the computer screen was male or female. Prior to the main task, the participants saw each stimulus face (printed out and mounted on pieces of card) for 3–5 sec, so that they would be familiar with the range of stimuli included. Half saw the male faces first, and half saw the female faces first. Each subset was ordered randomly. Once seated in front of the computer, the participants were asked to ignore the expressions on the faces while making their speeded responses and to use their index fingers to press one key for male and another for female. The keys were 15 cm apart on the keyboard, and hand of response was counterbalanced.

Each order of the three conditions was seen by 2 participants. Each of the conditions was made up of two consecutive blocks of 100 trials. Blocks contained 20 practice trials ordered randomly, followed by 80 test trials ordered randomly. In the correlated condition, Block 1 contained male angry and female surprised faces (the sex decision is assumed to be easy in this block), and Block 2 contained male surprised and female angry faces (the sex decision is assumed to be hard in this block). In the control condition, Block 1 contained male and female angry faces, and Block 2 contained male and female surprised faces. In the orthogonal condition, Block 1 contained male angry and surprised and female angry and surprised faces, and

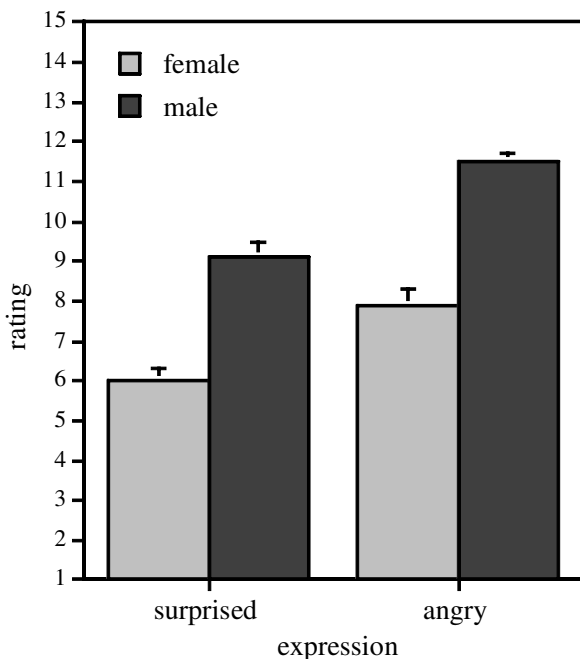


Figure 3. The effect of surprised and angry expressions on ratings of face masculinity in Experiment 1. 1 = very feminine; 15 = very masculine.

Block 2 contains the same. In this way, each stimulus was seen for a total of 8 test trials in each condition.

The procedure for the expression classification task was as above for sex classification, with task-relevant details changed as appropriate. For example, the participants were asked to ignore sex while making their expression decision. For expression classifications, the makeup of the correlated and orthogonal conditions remained the same, and the control condition contained angry and surprised male faces in Block 1 and contained angry and surprised female faces in Block 2.

For both tasks, each trial started with the presentation of a fixation cross, which was replaced after 1,550 msec by a stimulus face. The face remained on the screen until a key was pressed when it was immediately replaced by the next fixation cross. Criteria for inclusion in the data set were 50% correct and within the time window (150–2,500 msec) in each cell. The participants with errors and outlying responses greater than 10% overall were excluded. The experiment took 30–35 min to complete.

Results

Errors and outliers. Errors for the participants who met the scoring criteria were few. The participants in the sex classification group made average incorrect responses of 1.6% (correlated, 1.7%; control, 1.7%; orthogonal, 1.1%). The participants in the expression classification group made average incorrect responses of 2% (correlated, 1.3%; control, 2.7%; orthogonal, 2.2%). Error rates were not analyzed further. Outliers made up 0.03% of responses to sex and 0.2% of responses to expression.

RTs: Comparisons between classification tasks across conditions. Mean correct RTs for both groups in all experimental conditions can be seen in Figure 4. Inspection of the figure shows that the participants made sex classi-

fications more quickly than expression classifications. Also, the participants in the expression classification group showed faster reactions to stimuli in the correlated condition than in the control and orthogonal conditions. A 2×3 mixed ANOVA, with task as the between-subjects variable and condition as the within-subjects variable, showed a significant main effect of group [$F(1,22) = 12.6, p = .002$], a significant main effect of condition [$F(2,44) = 8.4, p = .001$], and a significant interaction between task and condition [$F(2,44) = 3.8, p = .03$]. Mean RTs for sex and expression responses were 522 msec ($SD = 87$ msec) and 679 msec ($SD = 142$ msec), respectively. A simple main effects analysis [$F(2,88) = 11.8, p < .001$] confirmed that the interaction reflected an effect of condition when making classification decisions for the expression task but not for the sex classification task.

RTs: Comparisons within each classification task.

RTs for every combination of each relevant and irrelevant dimension may be seen in Figure 5. In Figure 5A, the relevant variable is sex; in Figure 5B, the relevant variable is expression. Although differences between each of the combinations of relevant and irrelevant variables may be seen within each condition, comparisons between the control and orthogonal conditions are critical indicators of Garner interference. In neither task do these conditions appear to be very different from each other. Separate $3 \times 2 \times 2$ ANOVAs were carried out for each task, with the within-subjects variables of condition, sex, and expression.

For sex classifications, the overall mean RTs were 510, 528, and 528 msec for correlated, control, and orthogonal conditions, respectively. No main effects of condition, sex, or expression reached significance. A significant three-way interaction [$F(2,22) = 4.5, p = .02$] is interpreted as reflecting faster RTs to female surprised expressions than to female angry expressions and faster RTs to male angry expressions than to male surprised expressions in the correlated condition. Although the slight redundancy gain in the correlated condition observed in this task did not reach significance in the analysis across subjects, an analysis across items did show a significant main effect of condition [$F(2,16) = 8.05, p = .004$]. Although a slight redundancy gain within this dimension was observed, sex classification appears to have been unaffected by irrelevant expression variation in the control and orthogonal conditions.

For expression classifications, the mean RTs per condition were 619, 701, and 716 msec for correlated, control, and orthogonal conditions, respectively. A significant main effect of condition [$F(2,22) = 7.3, p = .004$] was further confirmed by an analysis across item [$F(2,16) = 95.82, p < .0001$]. However, Newman-Keuls comparison ($\alpha = .05$) of condition means showed that, although RT in the correlated condition was significantly different from RT in the control and orthogonal conditions, these latter were no different from each other. Expression classification therefore also showed a redundancy gain in the correlated condition; however, stimulus specific effects like those for sex classification were not observed for expression. For the

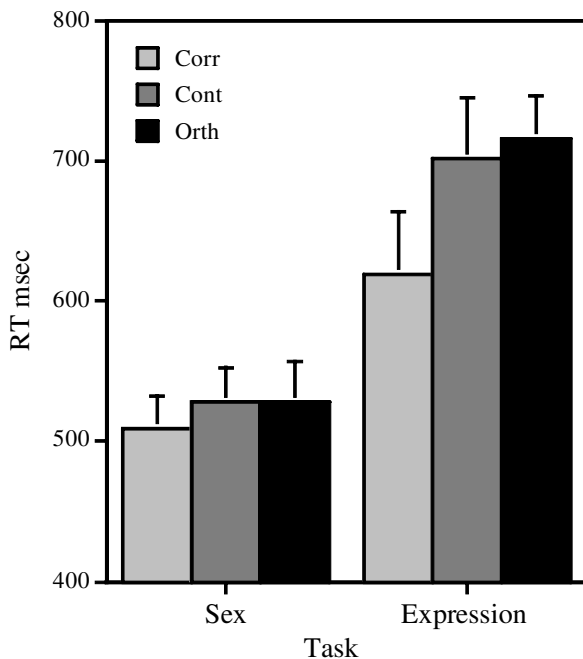


Figure 4. Mean reaction times in Experiment 2 for sex and expression classification groups in each condition. Vertical bars represent positive standard errors of the mean.

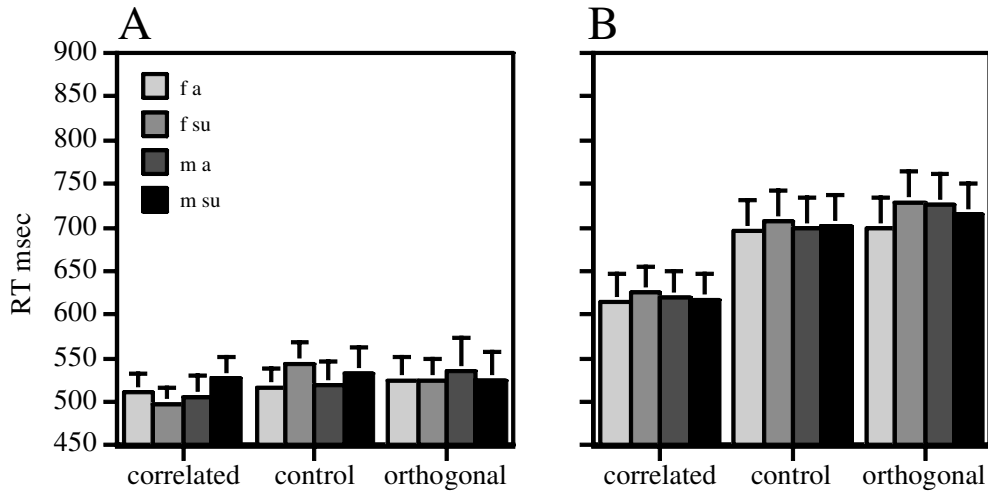


Figure 5. Mean reaction times for every combination of each relevant and irrelevant dimension for Experiment 2. (A) The relevant dimension is sex; (B) The relevant dimension is expression. Vertical bars represent the positive standard errors of the means. f = female face; m = male face; a = angry; su = surprised.

expression dimension also, there was no significant difference between control and orthogonal conditions.

Discussion

Sex classifications were faster than expression classifications, suggesting that sex is more easily discriminable than expression for these stimuli. The comparisons critical for establishing integrality—those between control and orthogonal conditions—were not significantly different either for expression classifications or for sex classifications, which would suggest that it is possible to attend to each dimension selectively. This result therefore appears to support the suggestion of independent processing routes for expression and sex. There is also some evidence that the participants were able to use correlated irrelevant information to enhance their performance, and this was most marked for the expression classification task. Furthermore, the analysis of intratask relationships did yield slight evidence to suggest that variation in expression may affect the ease with which contingency information is accessible to participants making sex classifications in the correlated condition (enhanced RTs to surprised female and angry male faces). Garner (1976, p.101) discusses how “higher order” interactions may take place with complex stimuli to influence conditions within dimensions, but, unless interference is demonstrated between control and orthogonal conditions, the overall conclusion of dimensional separability can be sustained.

However, the obvious difference in task difficulty indicated by longer RTs to expression than to sex in the baseline condition makes meaningful comparison of these two dimensions difficult. Melara and Mounts (1993), in a series of experiments to investigate the roles of practice, response mode, and baseline discriminability in modifying the asymmetric interference between color–word and

color dimensions, found that discriminability dominated interactions. Matching baseline discriminabilities between words and colors removed almost all interference, and certainly asymmetric interference, between these dimensions in a Garner experiment (Experiments 1c and 1w). When discriminability is mismatched, Melara and Mounts (1993, p. 627) propose that information from the more discriminable dimension will intrude on the less discriminable one, causing asymmetric interference and resulting in a “mandatory failure of selective attention.” Mismatching task difficulty may therefore in some cases lead to the presence of interference, with dimensions shown eventually to be separable. In order to eliminate the possibility of the more discriminable dimension intruding on the processing of the less discriminable dimension, RTs in the control conditions should not be significantly different from each other.

This issue could be a concern when faces are used as stimuli. However, Schweinberger et al. (1999) used morphed photographic stimuli to vary discriminability systematically across expression and identity dimensions and concluded that differences in perceptual salience were not responsible for the asymmetric interference that they observed between these facial dimensions.

While these initial results therefore support the independence of sex and expression processing, the redundancy gain in the expression classification task, together with the (nonsignificant) increase in orthogonal RTs over the control condition, is (weakly) suggestive of asymmetric interference. This implies that sex information may be processed independently of expression but that expression classification might be modified by the sex of the sender’s face. Also, complex interactions between different stimulus groupings revealed in the correlated condition for sex classification may become more important interdimen-

sionally when the discriminability of each dimension is matched. In Experiment 3, we attempted to match baseline RTs in order to make a more appropriate comparison.

EXPERIMENT 3

Because the discrimination of sex from facial features is readily affected by hairline (O'Toole et al., 1993) and jawline (Brown & Perrett, 1993), cropping the stimuli used in Experiments 1 and 2 more severely was predicted to slow sex decisions without affecting too greatly the time taken to respond to expression. Again, on the basis of the results of Experiment 1, it was predicted that sex classification would be influenced by variation in expression. It remained an open question as to whether expression classification would be influenced by variation in sex, although the redundancy gain and slight attenuation of RTs in the orthogonal condition of Experiment 2 suggested that it might be.

Method

Participants. Forty-eight participants (32 women, 16 men; see note 3) volunteered, each participant received either course credit or £3 for participation. The data from 24 participants, with a mean age of 23.1 years ($SD = 5.8$), were included, since 22 participants failed to achieve at least 50% correct responses within the time window in each cell for the sex classification task (much more difficult with the cropped stimuli). Only 2 participants failed these criteria for the expression classification task.

Stimuli and Apparatus. The face stimuli were the same as in Experiments 1 and 2, except that they were cropped to remove hairline and jawline information. Examples of the modified stimuli used in this experiment are shown in Figure 6. Stimulus size was slightly smaller, at 6×8.5 cm. The images subtended an angle of 5.7° horizontally and 8° vertically at a viewing distance of 0.6 m. Other viewing conditions remained the same as in Experiment 2.

Design. The experiment had a 2×3 mixed design, with a between-subjects variable of task at two levels (sex classification, expression classification) and condition at three levels (correlated, control, orthogonal). Dependent measures were RTs and errors.

Procedure. The procedure was as described for Experiment 2, with new stimulus versions printed out and viewed on cards prior to the speeded-classification tasks.

Results

Errors and outliers. Errors for the participants who met the scoring criteria were few. The participants in the sex classification group made average incorrect responses of 2.9% (correlated, 2.7%; control, 3.1%; orthogonal, 2.9%). The participants in the expression classification group

made average incorrect responses of 1.9% (correlated, 1.3%; control, 2.4%; orthogonal, 2.1%). Error rates for these participants were not analyzed further. Outliers made up 0.5% of responses to sex decisions and 0.3% of responses to expression.

Errors made by the participants excluded from the data set for sex classifications made up 11.6% of responses, which were also approximately evenly distributed across conditions (correlated, 11.8%; control, 11.9%; orthogonal, 11.2%). Of these errors, 26% were made to female angry faces, 15% to female surprised faces, 27% to male angry faces, and 32% to male surprised faces. The vast majority of the disqualifying responses by the 2 participants who did not achieve inclusion criteria for the expression classification were outliers. Overall, 2.3% errors were made by these participants (correlated, 5.9%; control, 0.3%; orthogonal, 0.6%). Of these errors, 50% were made to female surprised faces, 23% to male surprised faces, 23% to male angry faces, and 5% to female angry faces. Virtually all the errors made occurred with long (excluded) RTs.

There was then no sign of an effect of condition on the combined error rates for all 48 participants (mean errors = 2.0%, 2.1%, and 2.2% for correlated, control, and orthogonal conditions, respectively, within the expression task, and 8.6%, 8.8%, and 8.3% for correlated, control, and orthogonal conditions, respectively, for the sex task).

RTs: Comparisons between classification tasks across conditions. Mean correct RTs for both groups in all experimental conditions can be seen in Figure 7. Mean RT for sex classifications was 759 msec ($SD = 117$), whereas mean RT for expression classifications was 708 msec ($SD = 120$). Performance was equivalent across the three conditions for sex classification. For the expression task, there appears to have been some advantage for the correlated condition, but the mean RT in the orthogonal condition was also faster than that in the control condition. However, a 2×3 mixed ANOVA, with task as the between-subjects variable and condition as the within-subjects variable, showed no significant main effects of task or condition and no significant interaction between task and condition. The lack of a main effect of task is taken to indicate that difficulty was approximately equivalent in each task when RT alone is considered. Although no main effects or any interaction have been revealed, the results presented in Figure 7 suggest that there may have been a condition effect for expression classification,



Figure 6. Examples of the cropped faces used in Experiment 3. From left to right: female angry, female surprised, male angry, and male surprised.

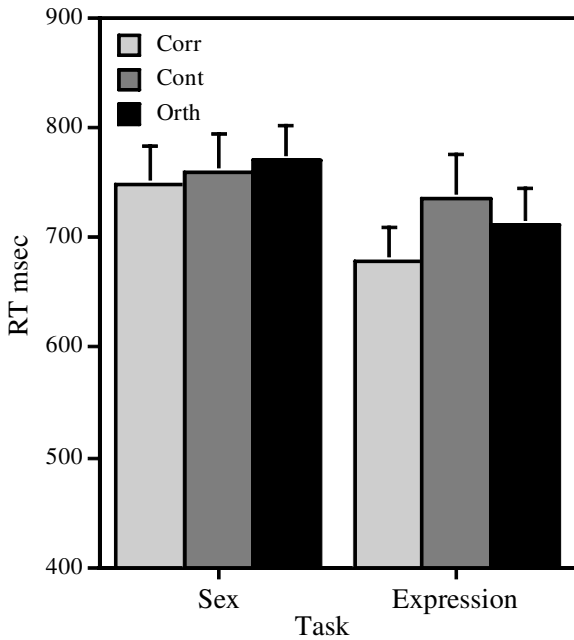


Figure 7. Mean reaction times in Experiment 3 for sex and expression classification groups in each condition. Vertical bars represent positive standard errors of the mean.

though not for sex. Dimensions were again analyzed separately to check for interference within each task.

RTs: Comparisons within each classification task. RTs for every combination of each relevant and irrelevant dimension may be seen in Figure 8. In Figure 8A, the relevant variable is sex; in Figure 8B, the relevant variable is expression. As for Experiment 2, differences between each of the combinations of relevant and irrelevant variables can be seen within each condition. In particular, the irrelevant variable may have been affecting RT in the corre-

lated condition for expression classification. However, comparisons between the control and orthogonal conditions were critical indicators of Garner interference. Separate $3 \times 2 \times 2$ ANOVAs were carried out, with within-subject variables of condition, sex, and expression.

For sex classifications, there were no significant effects of condition, sex, or expression, and there were no significant interactions. An analysis across items confirmed this.

For expression classifications, the mean RTs were 678, 736, and 712 msec for correlated, control, and orthogonal conditions, respectively. There was a significant main effect of condition [$F(2,22) = 4.2, p = .03$], and there was a significant three-way interaction between condition, expression, and sex [$F(2,22) = 3.96, p = .03$]. A Newman-Keuls comparison ($\alpha = .05$) of condition means showed that RT in the correlated condition was significantly different from RT in the control condition only. Further analysis of the three-way interaction showed that, for the correlated condition, female surprised faces are responded to more quickly than were female angry faces; for male faces, this was reversed, with male angry faces achieving faster RTs than male surprised faces [$F(2,33) = 5.4, p = .009$]. An items analysis confirmed this pattern of responses, with a significant main effect of condition [$F(2, 16) = 21.3, p < .0001$], a significant two-way interaction between condition and expression [$F(2, 16) = 3.6, p = .05$], and a significant three-way interaction between condition, expression, and sex [$F(2, 16) = 7.2, p = .006$]. A Newman-Keuls comparison of condition means again showed no significant difference between control and orthogonal conditions.

Discussion

The results of Experiment 3 suggest that the attempt to match task difficulty across sex and expression classifications was a partial success. For the sex classification task, although RTs were slowed by an appropriate amount for comparability with the expression task, accuracy was com-

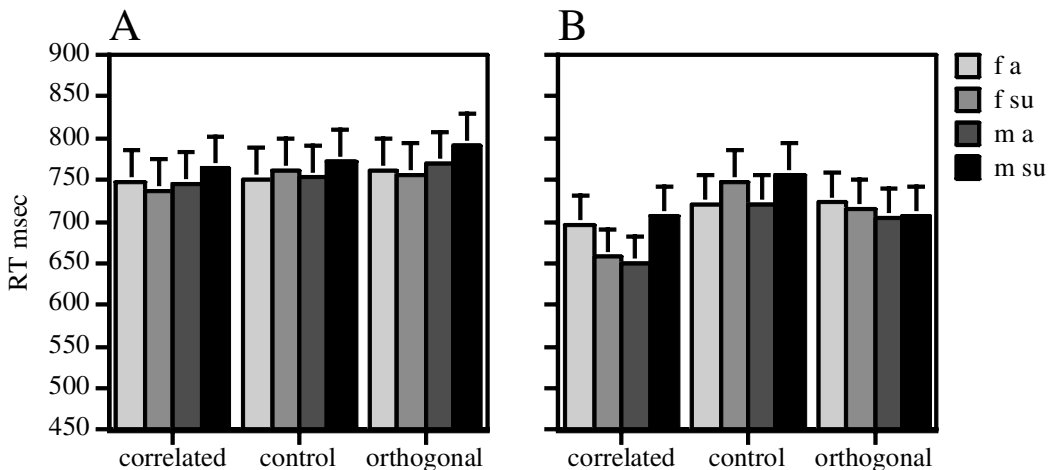


Figure 8. Mean reaction times for every combination of each relevant and irrelevant dimension for Experiment 3. (A) The relevant dimension is sex; (B) The relevant dimension is expression. Vertical bars represent the positive standard errors of the means. f = female face; m = male face; a = angry; su = surprised.

promised, and this meant that a high proportion of the participants failed to meet the criteria for inclusion in this speeded task. An analysis of RT alone from the included participants lends support to the tentative conclusion from Experiment 2 that participants are able to attend selectively to each dimension independently of the other. This indicates that, provided the structural information is sufficient to make the correct sex judgment, variations in expression are processed independently.

For expression classifications (cf. sex classifications in Experiment 2), the accessibility of contingent information was influenced by the irrelevant dimension, with male faces responded to more quickly as angry and with female faces responded to more quickly as surprised. This indicates that when stimulus information contained within the value on the irrelevant dimension corresponds to stimulus information within the value on the relevant dimension used to make a classificatory decision, responses may be speeded, and when it conflicts, responses may be slowed. This effect appears to be restricted to the correlated condition. Within-dimensional analysis confirmed the absence of any increase in difficulty from the control condition to the orthogonal condition. The influence of irrelevant variation in the correlated condition might be interpreted as a higher order effect arising for complex stimuli (see Garner, 1976, p. 101) but is not inconsistent with the overall demonstration of separability that arises from the comparison between control and orthogonal conditions.

For sex classifications, errors made by the large number of participants excluded from the RT data analysis also reflect an influence of irrelevant variation—this time, of expression. Nearly twice as many errors were made to female angry faces than to female surprised faces, and more were made to male surprised faces than to male angry faces. In spite of being shown which faces belonged in each set, the participants classified faces as either male or female early in the task on the basis of the information available. This pattern of errors is in line with the ratings made in Experiment 1 and indicates that, for still photographs, an expressive face (at least an angry face and a surprised face) can alter our perception of someone's gender. These results are also consistent with an observation by Burton et al. (1993) that faces misclassified by discriminant function analysis were also those that raters saw as atypical for their sex. Importantly, however, the errors were distributed evenly across the Garner conditions indicating that this aspect of stimulus assessment was not affected by increasing difficulty across conditions. This is suggestive of automatic processing specific to the values within each dimension, rather than a failure of selective attention between dimensions.

In summary, when RTs of the participants who achieved a high level of accuracy at sex classification were compared with those of the participants who maintained a similar level of accuracy at expression classification, no failure of selective attention for either dimension was observed. However, "congruency" effects, which emerged in the correlated condition when the participants classified expres-

sion, and the types of errors made by the participants who did not achieve criterion accuracy at classifying the faces for sex suggest that the participants were indeed influenced by changes in irrelevant variation when comparisons were made within dimensions.

GENERAL DISCUSSION

Variation in expression was able to modify the rated masculinity of male and female faces when the participants were allowed unlimited viewing time in Experiment 1. However, during speeded classification of the same faces in the laboratory, selective attention to both sex and expression dimensions was possible under conditions in which sex classifications were both easy (Experiment 2) and difficult (Experiment 3). In none of the groups carrying out sex or expression classification was a significant increase in RT from control to orthogonal conditions (the outcome considered critical to establishing dimensional integrality) detected. It therefore appears that expression and sex are separable dimensions, in Garner's terms, consistent with independent processing routes for expression and sex in functional models of face perception. Yet the rating experiment (Experiment 1) revealed a pattern more consistent with integrality. How are these results to be reconciled?

One possible explanation is that the rating of the masculinity/femininity of a face may be different from the time-stressed classification of male/female. For example, masculinity/femininity judgments to faces of known sex may be influenced by stereotypical attributions (anger is more masculine), whereas speeded classification could be conducted so quickly that judgments are immune from such influences. Moreover, the participants in Experiment 1 knew which faces were male and which were female, since these were presented in blocks. A face that is rated as masculine in appearance will not necessarily be classified as male, and some dissociation between sex classification and gender ratings has also been noted in previous work (Burton et al., 1993).

However, the observation that errors in Experiment 3 (from the participants excluded from the experiment) mirrored the ratings data of Experiment 1 suggests that the two tasks may not be so different. In this case, errors were made presumably because surprised faces in these circumstances looked more female and angry faces looked more male. This is probably because a major cue to sex, the brow-to-lid distance, is congruent with these decisions. If many more errors had been made in the orthogonal condition than in the control condition, rather than being equally spread across conditions as they were, then such a pattern would indicate a dimensional sharing of resources in Garner's (1976) terms. However, the fact that the errors made were distributed equally across the different Garner conditions suggests that the errors were due to higher order properties of the stimuli rather than due to a failure of selective attention.

Indeed, even for the participants included in the RT analyses, separate analyses of each of the classificatory tasks of

Experiments 2 and 3 showed that variations in expression consistent with features that play a role in the discrimination of sex from faces may be used to enhance responding to both sex (Experiment 2) and expression (Experiment 3) when these dimensions are correlated. Clearly, there is separability of sex and expression processing in Garner's terms (Experiments 2 and 3) but some interaction between the two dimensions as well (Experiment 1, ratings; Experiments 2 and 3, correlated conditions; Experiment 3, errors).

We suggest that there is a distinction to be drawn between information processing operations that may be independent (or separable, in Garner's terms) and interactions at the level of stimulus pattern, which may affect the way an item is assigned to a category. For example, an angry female face will be more distant from the prototypical appearance of "female" than will a surprised (or neutral) female face. This will lead to an angry female face being rated as less feminine in appearance and (depending on how much other information is present in the pattern) could lead to the item being misclassified as male in a speeded decision when pattern information is limited. However, once the item has been categorized in one way or another, whether or not this classification is the same as the one the experimenter came up with, it is still possible to attend selectively either to the sex or to the expression of that face. Similar arguments can be applied to the other kinds of item in our experiments, and can explain why there was some variation in ease of responding to different items within the correlated conditions of the experiments.

This analysis would seem to argue strongly for the use of a Garner paradigm in order to investigate the independence (or interference) of information processing operations, since other kinds of method may simply reveal interactions at the level of stimulus pattern rather than cognitive process. Laboratory experiments suggesting that functionally different types of information interact do not always include equivalent tasks for both dimensions under consideration. For example, the inclusion of gaze direction classifications for baseline (control) and filtering (orthogonal) tasks, as well as a baseline measure for sex classifications, would extend Campbell et al.'s (1996) findings with regard to sex-of-face and gaze direction dimensions.

However, there are also a number of difficulties associated with using Garner's selective attention paradigm with faces as stimuli. These range from ensuring matched discriminability of dimensions to interpreting outcomes when these are complex and to mapping any results achieved on to models of face perception. Indeed, research to date records conflicting findings even using the Garner logic to examine the same question (Etcoff, 1984; cf. Schweinberger et al., 1999, Schweinberger & Soukup, 1998). Faces as stimuli are more complex than are those explored by Garner (e.g., red and green squares and triangles, left and right positioned parentheses), and the logic that has been used here and by Etcoff, Schweinberger, and others has relied on making the equation between a processing route (e.g., to derive sex, expression, or identity) and a stimulus dimension. Given these difficulties, we suggest that, al-

though powerful, the Garner method should not be used alone but that results from this paradigm should be set alongside those obtained using other methods.

In this study, we obtained different answers to our question about the independence of sex and expression processing depending on the task and measure used. We have argued that there is a distinction to be drawn between higher order interactions at the level of the stimulus pattern and information processing operations that operate on that pattern. In terms of the information processing acts that allow faces to be categorized in terms of sex and expression, the data reported here best fit that pattern described as separable. There was a slight indication that the classification of sex information in Experiment 2 and the classification of expression information in Experiment 3 may have been optional separable (this may be linked to overall discriminability), but this qualification does not alter the substantive conclusion of separability between the dimensions. The major implication of this finding is that of independent processing systems for sex and expression information. That is, sex and expression processing are carried out by separate functional systems.

In conclusion, although the perceived masculinity of a face may be modified by variations in expression in a rating task, in a speeded-classification task, in circumstances in which task difficulty and accuracy of responding are matched, both sex and expression may be attended to selectively. The evidence presented here using the Garner paradigm thus provides qualified support for functional models of face processing (e.g., Bruce & Young, 1986), suggesting independent processing routes for sex and expression information.

REFERENCES

- ASHBY, F. G., & TOWNSEND, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93*, 154-179.
- BIEDERMAN, I., & CHECKOSKY, S. F. (1970). Processing redundant information. *Journal of Experimental Psychology*, *83*, 486-490.
- BROWN, E., & PERRETT, D. I. (1993). What gives a face its gender? *Perception*, *22*, 829-840.
- BRUCE, V. (1986). Influences of familiarity on the processing of faces. *Perception*, *15*, 387-397.
- BRUCE, V., BURTON, A. M., HANNA, E., HEALEY, P., MASON, O., COOMBES, A., FRIGHT, R., & LINNEY, A. (1993). Sex discrimination: How do we tell the difference between male and female faces? *Perception*, *22*, 131-152.
- BRUCE, V., ELLIS, H. D., GIBLING, F., & YOUNG, A. W. (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.
- BURTON, A. M., BRUCE, V., & DENCH, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception*, *22*, 153-176.
- CALDER, A. J., YOUNG, A. W., ROWLAND, D., PERRETT, D. I., HODGES, J. R., & ETCOFF, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology*, *13*, 699-745.
- CAMPBELL, R., BENSON, P. J., WALLACE, S. B., DOESBERGH, S., & COLEMAN, M. (1999). More about brows: How poses that change brow position affect perceptions of gender. *Perception*, *28*, 489-504.
- CAMPBELL, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lipreading: A neurological dissociation. *Brain*, *109*, 509-521.

- CAMPBELL, R., WALLACE, S., & BENSON, P. J. (1996). Real men don't look down: Direction of gaze affects sex decisions on faces. *Visual Cognition*, **3**, 393-412.
- COLE, M., & PEREZ-CRUET, J. (1964). Prosopagnosia. *Neuropsychologia*, **2**, 237-246.
- ELLIS, A. W., YOUNG, A. W., & FLUDE, B. M. (1990). Repetition priming and face processing: Priming occurs within the system that responds to the identity of a face. *Quarterly Journal of Experimental Psychology*, **42A**, 495-512.
- ELLIS, H. D. (1986). Processes underlying face recognition. In R. Bruyer (Ed.), *The neuropsychology of face perception and facial expression* (pp. 1-27). Hillsdale, NJ: Erlbaum.
- ERCOFF, N. L. (1984). Selective attention to facial identity and facial emotion. *Neuropsychologia*, **22**, 281-295.
- FELFOLDY, G. L., & GARNER, W. R. (1971). The effects on speeded classification of implicit and explicit instructions regarding redundant dimensions. *Perception & Psychophysics*, **9**, 289-292.
- GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- GARNER, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, **8**, 98-123.
- GOSSELIN, P., KIROUAC, G., & DORÉ, F. Y. (1997). Components and recognition of facial expression in the communication of emotion by actors. In P. Ekman & E. L. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)* (pp. 243-267). Oxford: Oxford University Press.
- GREEN, K. P., KUHL, P. K., MELTZOFF, A. N., & STEVENS, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, **50**, 524-536.
- 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.
- LE GAL, P. M. (1999). *Cognitive aspects of emotional expression processing*. Unpublished doctoral thesis, University of Stirling.
- MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746-748.
- MELARA, R. D., & MARKS, L. E. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 398-414.
- 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.
- O'TOOLE, A. J., ABDI, H., DEFFENBACHER, K. A., & VALENTIN, D. (1993). Low-dimensional representation of faces in higher dimensions of the face space. *Journal of the Optical Society of America A*, **10**, 405-410.
- PARRY, F. M., YOUNG, A. W., SAUL, J. S. M., & MOSS, A. (1991). Dissociable face processing impairments after brain injury. *Journal of Clinical & Experimental Neuropsychology*, **13**, 545-558.
- ROLLS, E. T. (1992). Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas. *Philosophical Transactions of the Royal Society of London: Series B*, **335**, 11-21.
- ROSENBERG, E. L. (1997). The study of spontaneous facial expressions. In P. Ekman & E. L. Rosenberg (Eds.), *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS)* (pp. 3-17). Oxford: Oxford University Press.
- 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.
- SERGENT, J., OHTA, S., MACDONALD, B., & ZUCK, E. (1994). Segregated processing of facial identity and emotion in the human brain: A PET study. *Visual Cognition*, **1**, 349-369.
- TRANSEL, D., DAMASIO, A. R., & DAMASIO, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, **38**, 690-696.
- WALKER, S., BRUCE, V., & O'MALLEY, C. (1995). Facial identity and facial speech processing: Familiar faces and voices in the McGurk effect. *Perception & Psychophysics*, **57**, 1124-1133.
- 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.

NOTES

1. The terms *sex* and *gender* are used interchangeably throughout the introduction and discussion, using whichever term was used by the authors of each paper we discuss. In all our published research, we have used the term *sex* to refer to biological sexual identity, and we retain this usage here for Experiments 2 and 3 in which classification of faces as either (biologically) male or female was required. However, the term *gender* is the preferred usage for Experiment 1 where an assessment along the masculinity/femininity dimension of faces of known sex was required.

2. The finding here that both male and female surprised faces were considered more feminine than were angry faces is not completely analogous to the rating study carried out by Campbell et al. (1996), in which there was an interaction between "maleness" and "femaleness" for faces looking down relative to faces looking at the viewer, such that only male faces looking down were thought more "female."

3. Pilot experiments that equated the number of male and female participants and compared their performance showed no differences, and, therefore, no attempt was made to examine data for sex differences in Experiments 2 and 3.

(Manuscript received March 29, 2000;
revision accepted for publication April 26, 2001.)