NONVERBAL EMOTION PRIMING: EVIDENCE FROM THE 'FACIAL AFFECT DECISION TASK'

Marc D. Pell

ABSTRACT: Affective associations between a speaker's voice (emotional *prosody*) and a facial expression were investigated using a new on-line procedure, the Facial Affect Decision Task (FADT). Faces depicting one of four 'basic' emotions were paired with utterances conveying an emotionally-related or unrelated prosody, followed by a yes/no judgement of the face as a 'true' exemplar of emotion. Results established that prosodic characteristics facilitate the accuracy and speed of decisions about an emotionally *congruent* target face, supplying empirical support for the idea that information about discrete emotions is shared across major nonverbal channels. The FADT represents a promising tool for future on-line studies of nonverbal processing in both healthy and disordered individuals.

KEY WORDS: emotion; facial expression; nonverbal communication; on-line processing; prosody.

Introduction

Researchers interested in emotive communication have paid particularly close attention to how facial cues, and to a lesser extent vocal-prosodic cues (emotional *prosody*), contribute to the regulation of interpersonal behavior. These varied efforts have yielded increasingly sophisticated descriptions of the physical and psychoperceptual attributes of overt facial and prosodic displays of emotion, of autonomic events associated with these signals, and of the cognitive and neural architecture involved in their encoding and decoding (see discussions by Baum & Pell, 1999; Öhman, 1987; Scherer, 1993). Relatively few studies have reported detailed

---

Marc D. Pell is affiliated with the School of Communication Sciences and Disorders, McGill University, Montréal, Canada.

The author gratefully acknowledges the service of Elmira Chan, Sarah Addleman, and Marta Fundamenski for running the experiment, and for helpful comments received from M. Harris, K. Scherer and anonymous reviewers of an earlier draft of the paper. This research was funded by the Natural Sciences and Engineering Research Council of Canada.

Address correspondence to Marc D. Pell, Faculty of Medicine, School of Communication Sciences and Disorders, McGill University, 1266 avenue des Pins ouest, Montréal (Qc) CANADA H3G 1A8 ; e-mail: marc.pell@mcgill.ca.

Journal of Nonverbal Behavior, 29(1), Spring 2005 © 2005 Springer Science+Business Media, Inc.
DOI: 10.1007/s10919-004-0889-8

information on the processing of emotional prosody or facial expressions during on-line mediation of these events, and data exemplifying how *simultaneous* activations of prosodic and facial displays of emotion are registered and integrated by decoders in real time are scarce. New on-line data are essential to provide clues about how nonverbal stimuli presented in separate but conjoined channels are evaluated and (presumably) combined to form a coherent impression of a communication partner's affective status, and perhaps to derive other types of inferences during natural discourse. These concerns motivated the current study which adopted a new on-line approach to examine issues in the processing and recognition of nonverbal emotion displays, and in particular, to obtain empirical support for the assumption that emotional prosody and emotional faces share an overlapping cognitive structure during input processing (e.g., Borod, 2000; Bowers, Bower, & Heilman, 1993).

Affective Bias in Nonverbal Processing

The past decade has seen burgeoning evidence of affective bias or *priming* in cognitive processing, where encoding of a ''target'' event—typically, a written word, picture, or facial stimulus—is systematically influenced by shared affective evaluations of a spatially or temporally contiguous stimulus ''prime'' (see Wentura, 2000 for a recent review and critique). This literature emphasizes that sensory events are *automatically* appraised by decoders to register central details about their affective significance; in particular, it is now widely accepted that the positive or negative *valence* of prime-target events plays an instrumental role in how most stimuli are processed (Bargh, 1988; Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Fiske & Pavelchak, 1986; Hermans, De Houwer, & Eelen, 1996). The *arousal* component of sensory input is also known to impact in an important manner on the cognitive processing system (Bargh, Chaiken, Raymond, & Pratto, 1992; Kitayama & Howard, 1994; Williams, Matthews, & MacLeod, 1996), although this dimension has been studied relatively little in studies of affective bias. When applied to nonverbal displays of emotion, evaluations of arousal are thought to index the relative intensity or physiological involvement of the encoder in the emotional expression, cues which influence the decoder's psychophysiological response to the event and attributions about its immediate importance (Kitayama & Howard, 1994; Scherer, 1989). It is widely accepted that procedures for determining both valence and arousal supply central details about the affective nature of most stimuli (Robinson, 1998; Russell, 1980; Scherer, 1986).

However, when complex social stimuli such as nonverbal displays of emotion are processed, stored knowledge referring to how differentiated emotion states such as *happiness* or *anger* are expressed through these channels may also be activated to provide greater specification of these events (Bower, 1981; Ekman, Sorenson, & Friesen, 1969; Izard, 1977; Niedenthal & Halberstadt, 1995). Indeed, there is evidence that facial expressions are perceived categorically with respect to certain basic emotions (Bimler & Kirkland, 2001; Etcoff & Magee, 1992) and that vocal and facial expressions exhibit strong 'modal' tendencies by emotion across a range of human cultures (Ekman et al., 1969; Scherer, Banse, & Wallbott, 2001; Scherer & Wallbott, 1994). These findings reinforce the idea that on-line evaluation of emotional prosody and emotional face stimuli may include processes which activate specialized emotion knowledge tied to these particular events, associations that are more differentiated than for many other types of "affective" input. Moreover, these procedures may be shared to some extent or linked in some manner for stimuli encountered in distinct nonverbal channels (Bowers et al., 1993).

A number of researchers have documented the biasing effects of emotional prosody (Kitayama & Howard, 1994; Schirmer, Kotz, & Friederici, 2002; Wurm, Vokoch, Strasser, Calin-Jageman, & Ross, 2001) or an emotional face stimulus (Murphy & Zajonc, 1993; Niedenthal, 1990) on processing of an associated target event, especially for target *words*. Investigators who have focused specifically on the connections between emotional prosody and emotional faces (i.e., prosody-face, face-prosody) in studies of affective bias are relatively few, but noteworthy (de Gelder, Bocker, Tuomainen, Hensen, & Vroomen, 1999; de Gelder, Pourtois, Vroomen, & Bachoud-Levi, 2000; de Gelder & Vroomen, 2000; Massaro & Egan, 1996).

The idea that processing mechanisms devoted to facial and vocal displays of emotion are strongly linked was addressed recently by de Gelder and Vroomen (2000), who examined whether sentences intoned in a happy or sad voice would prime decisions about congruent or incongruent faces (Experiments 1 & 2), or whether happy or fearful faces would prime decisions about congruent or incongruent sentence prosody (Experiment 3). The authors' measures considered how quickly and accurately subjects could identify the happy/sad (or happy/fearful) emotion of the actor based on nonverbal cues present in each trial. The authors reported that when emotional prosody served as a prime to facial targets, evaluation latencies were significantly slower for inconsistent than consistent pairings suggesting "mandatory influence of voice tone on judgements of facial expression'' (p. 301). Moreover, integration of affective detail from vocal and facial inputs appeared to occur very early during perceptual processing, based on results of a related ERP study conducted by the investigators (de Gelder et al., 1999; de Gelder & Vroomen, 2000). Although implications of their third experiment involving face to voice priming were far less definitive, these data and related work (Massaro & Egan, 1996) point to substantial, potentially bidirectional ties between affect processing centers underlying prosody and face decoding which may bias processing of either event. These preliminary conclusions, particularly those demonstrating prosody-to-face priming by emotion for which the data were far more robust, merit re-examination to extend our understanding of how emotional prosody influences on-line judgements of related or unrelated emotional faces in a new, preferably more implicit evaluative context.

The Facial Affect Decision Task

In the psycholinguistic literature, and in specific studies of affective *word* priming, the lexical decision task (LDT, Onifer & Swinney, 1981; Swinney, 1979) has acted as an important investigative tool. This approach, which evokes a word/non-word decision about a visual letter string presented in conjunction with a prime stimulus of some description, is believed to confer several advantages to those interested in cognitive attributes of word processing (see Wentura, 2000 for a discussion). For example, it has been emphasized that the response requirements of the LDT—a yes/no judgement of the lexical status of the word—are unrelated to *conscious* evaluation of constituent features of either the prime or the target, allowing researchers relatively direct insight into the activation of associated knowledge in memory. Given the hypothesis that discrete emotion concepts or knowledge prototypes are important in the recognition of facial emotion (Bower, 1981; Ekman, 1992; Niedenthal & Halberstadt, 1995; Young et al., 1997), one way to advance specific ideas about *nonverbal* decoding and supporting cognitive mechanisms is to require subjects to determine the 'candidacy' of a facial target stimulus as representing a ''true'' or ''false'' (i.e., known or unknown) expression of emotion in a highly analogous manner to requiring a written word lexical decision.

Such a paradigm was described briefly in a neuropsychological case study by Rapczak, Comer, and Rubens (1993) who recorded a yes/no acceptability judgement about recognizable exemplars of basic human emotions or pictures of facial ''grimaces,'' designating 'true' and 'false' expressions of emotion, respectively. Presumably, executing a 'facial affect decision' of this nature mandates an analysis of emotional features of the face that includes reference to known prototypes in the mental store (among other meaning activations); however, following assumptions of the LDT, subjects are not required to consciously access these attributes to generate a response, nor to retrieve verbal labels corresponding to specific emotional features. By requiring subjects to engage in this form of semantic analysis of face targets coupled with an on-line 'prime' utterance bearing congruent or incongruent emotional prosody, a constructive new context for revealing points of emotional bias between nonverbal processing channels is thus created.

The present study adopted this rationale and methodological approach in the construction of the *Facial Affect Decision Task* (FADT), introduced here as a future tool for elucidating the nature, strength, and timecourse of activations underlying the evaluation of nonverbal displays. Given the evidence that emotional prosody influences the processing of static faces in a cross-modal *evaluation* paradigm involving happy or sad stimuli (de Gelder & Vroomen, 2000), the nature of prosody-face interactions was explored further in the new context of the FADT, and for a larger set of emotional categories (*happiness, pleasant surprise, sadness, and anger*). We hypothesized that facial affect decisions would be facilitated in both accuracy and speed when the emotional value of a prosodic prime and of a facial target stimulus were related rather than unrelated, providing empirical support for the notion that cognitive substrates of emotional prosody and emotional face decoding are not fully independent. We predicted that congruity effects in our performance measures would revolve significantly around the emotion *identities* of prime-target events (Bower, 1981; Hansen & Shantz, 1995; Niedenthal & Halberstadt, 1995). At the same time, displays representing each of our four emotions were presumably associated with unique two-factor combinations of underlying valence (positive/negative) and arousal (high/low) characteristics of the stimuli (e.g., Russell, 1980). Thus, our design allowed initial commentary on the interplay of valence, arousal, and emotion-specific influences on nonverbal processing. For exploratory purposes, the *timecourse* of nonverbal emotion priming was also considered briefly by setting the stimulus onset asynchrony (SOA) of the face target to either 300 or 1000 ms following onset of the prosodic stimulus (e.g., Fazio et al., 1986; Neely, 1991), although the precise effects of this temporal manipulation in the context of the FADT could not be anticipated with certainty.

Method

Subjects

Forty-eight students from McGill University were recruited to the study through campus ads and compensated a small amount for their participation. Subjects comprised an equal number of female and male participants who were right-handed, native speakers of Canadian English with an average of 16 years of formal education. All subjects displayed good hearing and normal or corrected-to-normal vision, as determined by self-report.

Materials

(a) Stimulus elicitation and selection. Auditory and visual stimuli used in the FADT were chosen from an inventory of nonverbal displays elicited from a group of eight 'encoders' (see Pell, 2002 for extended details of this dataset). Four male and four female adults were recorded posing neutral expressions and a series of 'basic' emotions through their face (static photographs) and through their ''tone of voice'' (short utterances). Emotional target displays captured in each channel were: happiness, pleasant surprise, sadness, anger, disgust, and neutral. In the facial channel, encoders were also prompted to produce ''grimaces'' that involved movements of the jaw, lips, and brow that would not be construed by decoders as real emotional expressions (Rapcsak et al., 1993). Multiple exemplars of each target were generated by each encoder, accompanied by examiner cues and feedback (Borod et al., 1998; Ekman & Friesen, 1976). Tokens were captured directly onto digital recording media for transfer and editing on a PC.

Emotional features of selected displays were established perceptually by 32 'decoders' who used a checklist to rate facial and vocal stimuli in separate tasks. Independent perceptual data were gathered on: (a) the group consensus about the emotional identity of each display based on an eight-choice, open response paradigm (choices: neutral, happiness, pleasant surprise, sadness, anger, disgust, ''not a real emotional expression'' (to validate 'false' facial expressions), and ''other'' (subjects could specify)); (b) the perceived ''intensity'' of correctly identified, intended target meanings along a 5-point scale of increasing magnitude; (c) the *valence* of each token along an 11-point Likert scale ranging from ''very negative'' (−5) to ''very positive'' (+5). Valence ratings were collected without reference to potential emotional associations of the stimuli for a separate group of 14 decoders. Collectively, these background data contributed to the highly controlled selection of experimental tokens for manipulation of priming conditions within the FADT: a minimum criterion of 80% consensus about the emotional identity of facial targets and 65% for prosodic primes[1] governed selection of stimuli used in the actual experiment, where chance assignment of intended meanings was always 12.5%. Only displays conveying *happiness, pleasant surprise, sadness,* and *anger* served as targets in the current study, allowing the valence (positive, negative) and arousal (low, high) of experimental stimuli presented in the experiment to be further controlled, where: *happiness* = positive valence-low arousal; *pleasant surprise* = positive valence-high arousal; *sadness* = negative valence-low arousal; *anger* = negative valence-high arousal (e.g., Russell, 1980).

(b) Prosodic stimuli. Ten vocal displays representing each of the four emotions (*happiness, pleasant surprise, sadness, anger*) and ten exemplars of ''neutral'' affect served as prosodic primes (50 exemplars in total). Vocal stimuli took the form of nonsense utterances (e.g., *Someone migged the pazing*) pronounced according to English phonotactic constraints. Stimuli of this nature serve as an effective carrier of emotion in speech which isolate meaningful input to prosodic cues (Pell & Baum, 1997; Scherer et al., 2001). Experimental primes ranged in length from 1.19 to 2.74 seconds owing to natural differences in tempo, when intoned in the five emotional modes (further acoustic specification of this dataset is underway). ''Filler'' primes in the form of a 300 Hz pure tone stimulus 1.75 s in length (steady-state) also preceded a subset of facial targets to decrease expectancies about prosodic stimuli in the experiment. All auditory stimuli were prepared using CSL hardware and software (Kay Elemetrics, 4300B).

(c) Facial stimuli. Ten static faces representing each of the four emotions served as 'true' emotional target expressions for facial affect decisions (''YES'' response targets, $n = 40$). An equal number of faces portraying a grimace or 'false' emotional expression were also selected (''NO'' response targets, $n = 40$). Facial stimuli were modelled roughly after Ekman and Friesen's (1976) series, depicting a color photograph of an encoder's face, hair, and shoulders, saved as digital bitmaps measuring 17.1 × 17.1 cm for computer presentation. 'True' and 'false' portrayals were elicited from the same encoders controlling for any influence of encoder identity and related visual properties of the stimuli on facial affect judgements. Samples of 'true' and 'false' targets posed by one male and one female encoder are supplied in Figure 1. Table 1 summarizes major perceptual features associated with vocal and facial stimuli entered into the experiment.

Figure 1. Examples of 'true' and 'false' facial emotion targets presented in the experiment for one female and one male encoder.

TABLE 1

Perceptual Characteristics of Prosodic and Facial Stimuli Entered into the Experiment (*n* = 10/category)

Stimulus emotion identity	Emotion recognition[a] (%)	Emotion intensity[b] (0–5)	Stimulus valence[c] (−5 to 5)
Prosodic Primes			
Happiness	78.6	4.0	+2.8
Pleasant Surprise	68.7	3.8	+2.1
Sadness	90.2	3.6	−2.2
Anger	75.6	3.5	−2.2
Neutral	69.5	–	+0.1
Facial Targets			
Happiness	99.4	3.9	+2.9
Pleasant Surprise	98.2	4.0	+2.5
Sadness	88.0	3.7	−2.6
Anger	91.8	3.7	−3.0
'False'	72.5	–	–

[a]Based on an eight-choice forced paradigm including an open response category, chance performance = 12.5%.
[b]Represents the perceived intensity of emotion features present in the signal when identified as the intended emotion.
[C]Reflects evaluation of stimuli along an 11-point positive–negative continuum by a separate group of decoders.

Experimental Design

The experiment was composed of 480 trials representing an equal ratio of true and false facial targets, each posed by an equal ratio of female and male encoders. Trials were constructed by separately pairing each auditory prime with 'true' and 'false' facial expressions posed by an encoder of the same gender (although not always the same encoder). Whereas 'false' trials were constructed by combining each prime type with each of the 40 nonsense faces (6 primes $\times$ 40 items $=$ 240), prosody-face correspondences underlying 'true' trials were factorially manipulated to consider all combinations of prime and target values (6 primes $\times$ 4 targets $\times$ 10 items $=$ 240). A critical subset of 40 'true' trials were characterized by *congruent* pairings, where the prosody and face conveyed the same emotion; for this subset of trials only, the perceived intensity of the vocal versus the facial exemplar of each emotion was controlled by combining tokens that varied by no more than .5 on the 5-point rating continuum, avoiding potential asymmetries in the strength of evaluative meanings associated with matched prosody-face events (Fazio et al., 1986; Kitayama & Howard, 1994). The majority of "true" face trials involved prosody-face pairings characterized by an *incongruent* emotional relationship (4 prime emotions $\times$ 3 incongruent target emotions $\times$ 10 items $=$ 120). Forty *neutral* trials were composed of a prosodically-neutral utterance matched with facial tokens of the four target emotions, and 40 "filler" trials consisted of the 300 Hz pure tone stimulus matched with each of the 40 "true" emotional faces. Note that strategic processing of facial targets was mitigated by this experimental design through the low proportion of trials in which prosody-face features were actually related (.08 based on emotion congruence, or .17 based on either valence or arousal congruence) (Neely, Keefe, & Ross, 1989). To provide initial data on the timecourse of nonverbal priming effects, the interval between onset of the vocal prime and onset of the facial decision target was systematically varied within all conditions involving 'true' and 'false' facial targets. Half of the trials presented the facial stimulus at a *short* (300 ms) SOA and half presented the face at a *long* (1000 ms) SOA. The prime-target SOA for filler trials was set at 300 ms, 1000 ms, and at the offset of the vocal stimulus to further vary expectations about the timing of facial targets.

Procedure

The 480 trials were pseudo-randomly separated into 12 blocks of 40 trials for presentation. Each vocal prime appeared eight times in the experiment in unique combinations with facial targets (four times each with true and false targets). Each true and false emotional face appeared six times in the experiment. All blocks contained a relatively equal ratio of true and false facial targets and a roughly proportionate distribution of trials from each priming condition. No facial stimulus appeared twice in the same block, although a single repetition of vocal primes in the same block occurred occasionally (separately by at least five trials). Pure tone auditory fillers appeared an average of seven times per block. Short and long SOA trials were randomly intermixed within blocks.

Subjects were tested individually in a dimly lit, sound-attenuated room during two 30-min sessions. The experiment was run as follows: prosodic and facial stimuli were presented from a portable computer with a 37.5 cm color monitor and millisecond timer, controlled by SuperLab software (Cedrus Corporation). Prosodic primes were played through high-quality stereo headphones with manual volume adjustment. Facial stimuli were simultaneously viewed on the computer monitor from a distance of approximately 45 cm. Each trial consisted of: a centrally-displayed visual fixation marker of 350 ms; a 500 ms pause; the vocal stimulus; and the facial target, appearing at the short or long SOA. Participants were instructed to ignore the vocal stimulus (which would sound like ''made up words'') and to decide whether the face conveyed a recognizable or *true* emotional expression by pressing the YES or NO button on a Cedrus 6-button response box. The vocal prime was always presented in its entirety or until a key press was detected by the computer in response to the face, at which time both the facial and vocal stimulus was arrested and a 2000 ms pause was initiated before moving to the next trial. Both accuracy and speed of the response were recorded.

To familiarize subjects with the procedure and the concept of 'true' versus 'false' emotional faces, testing always began with two training blocks. In an initial practice block, subjects rendered a decision about five true and five false emotional faces without auditory accompaniment. A second practice block consisted of the same facial decision targets coupled with prosodic stimuli in a manner that approximated the actual experiment. Following practice, six blocks were assigned to each testing session and randomized for presentation order using a Latin square design. Sessions were counterbalanced for order of initial presentation over subjects and separated by at least one week to mitigate stimulus repetition effects. At no time during the practice nor the experiment did the examiner direct subjects' attention to specific affective/emotional qualities of the vocal or facial stimuli.

Results

Table 2 presents the mean error and latency scores obtained at each SOA as a function of the emotional value of the prosody and of the face. Overall, most subjects could reliably categorize facial expressions as 'true' or 'false' representations of emotion as indexed in a unique manner by the FADT: the group error rate for the 48 subjects was 16.8% (±11.5) for ''YES'' trials and 7.4% (±6.1) for ''NO'' trials. It is noteworthy that *false* faces, which will not be considered in subsequent analyses, were associated with relatively low error rates and average latencies, when compared to the different *true* emotional faces, and that responses to *false* faces varied little as a function of the prosodic prime value or SOA (review Table 2). These overall findings establish the participants' sensitivity to the emotional-symbolic value of faces while highlighting an elevated tendency for certain subjects to incorrectly reject certain 'true' expressions of emotion through this paradigm (see discussion below).

Data analysis (repeated measures ANOVA) was performed separately on error and latency scores and concentrated on decisions about 'true' facial expressions to establish that congruent emotional features of a prosodic and facial stimulus facilitate processing measures in a detectable manner across channels. Subsequent analyses sought to qualify the effect of nonverbal emotion priming and to probe alternative accounts for patterns in the data. Analyses of error measures always considered data for all 48 participants, whereas analyses of latency measures were confined to a subset of 33 participants (17 female, 16 male), whose accuracy for 'true' face targets was sufficiently reliable for these analyses, with no greater than 25% errors corresponding to 'true' face targets overall (Wurm et al., 2001).[2] Latency analyses considered only correct decisions about 'true' face targets, after the data were combed for outliers and extreme values by eliminating latencies over 2000 ms or shorter than 300 ms (.6% of all values) and by setting latencies more than two standard deviations from the conditional mean based on individual subject distributions at values equal to 2 SD in the appropriate direction (5.5% of total values) (Fazio et al., 1986; Neely, Keefa, & Ross, 1989). To focus discussion on the most reliable statistical effects, interpretation of ANOVA results referred to a more conservative confidence level of $p < .01$ considering the corrected p values for tests involving repeated measures with more than 1 df (Max & Onghena, 1999). Tukey *post hoc* comparisons were employed to elaborate all significant main and interactive effects.

TABLE 2

Mean Error Rates (%) and Response Times (ms, in parentheses) for 'True' Facial Targets as a Function of Characteristics of the Prosodic Prime and Face Target, Separately by SOA

	Emotion of prosody				
Emotion of face	Happiness	Pleasant surprise	Sadness	Anger	Total
SOA = 300 ms					
Happiness	5.8 (633)	10.8 (681)	5.9 (655)	4.2 (635)	6.7 (651)
Pleasant Surprise	11.3 (700)	13.8 (702)	14.1 (721)	11.5 (663)	12.7 (697)
Sadness	24.0 (763)	17.7 (747)	15.8 (733)	19.6 (777)	19.3 (755)
Anger	26.0 (761)	27.9 (775)	40.8 (785)	24.6 (773)	29.8 (774)
All 'true' faces	16.8 (714)	17.6 (726)	19.2 (724)	15.0 (712)	17.1 (719)
All 'false' faces	7.5 (702)	6.1 (696)	8.0 (707)	7.2 (688)	7.2 (698)
SOA = 1000 ms					
Happiness	2.1 (619)	2.9 (608)	4.2 (631)	7.6 (655)	4.2 (628)
Pleasant Surprise	21.7 (756)	13.8 (667)	15.6 (746)	17.7 (691)	17.2 (715)
Sadness	12.5 (695)	18.2 (711)	16.3 (710)	15.0 (692)	15.5 (702)
Anger	26.7 (767)	28.8 (731)	18.7 (787)	20.0 (763)	23.6 (762)
All 'true' faces	15.8 (709)	15.9 (679)	13.7 (719)	15.1 (700)	15.1 (702)
All 'false' faces	6.7 (668)	8.3 (654)	7.3 (682)	9.7 (666)	8.0 (668)

Note. Corresponding measures for 'false' faces are shown for comparison purposes.

Emotion Priming Across Nonverbal Channels

Our main analysis tested the question of whether congruent emotional attributes of prosody-face stimuli reliably enhance performance involving the emotional processing system over trials involving discrepant combinations of emotions, furnishing evidence of shared associative ties for these separate events. In addition, these analyses sought to reveal whether emotion congruity effects were greater for certain emotions represented by our target faces and/or according to the prime-target SOA. For both error and latency scores, conditions in which the prosody and face represented the same emotion (*congruent*) were compared to the average scores for the three conditions in which prosodic information conflicted with the face emotion identity (*incongruent*). Separate $2 \times 4 \times 2$ ANOVAs with repeated factors of Prosody-face relationship (congruent, incongruent), Face target value (*happiness, pleasant surprise*, *anger, and sadness*), and prime-target SOA (short, long) were then performed on each measure. Table 3 summarizes the error and latency data according to the prosody-face emotion relationship.

Analysis of *error scores* indicated that the ability to accurately judge the emotional-symbolic value of a face revolved significantly around the emotional relationship of the prosody to the face, $F(1, 47) = 10.24$, $p = .002$, $r = .42$. Overall, subjects could more precisely render an acceptability judgement about the face when it was preceded by emotionally congruent (M errors = .14) rather than incongruent (M errors= .17) prosodic information present in a spoken utterance. For *latency scores*, there were again reliable indications that facial targets primed by a congruent emotional prosody ($M = 700$ ms) facilitated decisions about their emotional acceptability over conditions in which prosodic cues biased an incongruent emotion ($M = 714$ ms), $F(1, 32) = 7.27$, $p = .01$, $r = .43$. The overall benefits of encountering emotionally congruent prosody-face events on both error and latency performance in the experiment are exemplified by Figure 2.

For both decision errors and latencies, the ANOVAs failed to uncover evidence that emotion congruity effects varied significantly as a function of specific target emotions within the dataset (all interactions involving Face and Prosody-face relationship, ps > .01). In addition, there were no statistically reliable indications that distinctions in the prime-target SOA moderated the extent of emotion congruence priming for these analyses (ps for the interactions >.72). Not surprisingly, the prime-target SOA was a significant independent factor in both error and latency patterns, as subjects made more accurate and faster facial affect decisions at the long rather than the

TABLE 3

Mean Error Rates (%) and Response Times (ms, in parentheses) to the Four 'True' Facial Target Emotions According to the Prosody-Face Priming Relationship, Separately by SOA

	Prosody-face relationship			
Emotion of face	Congruent	Incongruent	Neutral	Priming (Incon-con)
SOA = 300 ms				
Happiness	5.8 (633)	7.0 (657)	7.5 (612)	+1.2 (+24)
Pleasant surprise	13.8 (702)	12.3 (695)	23.8 (706)	−1.5 (−7)
Sadness	15.8 (733)	20.4 (762)	18.3 (714)	+ 4.6 (+29)
Anger	24.6 (773)	31.6 (773)	35.8 (832)	+7.2 (0)
Total	15.0 (710)	17.8 (722)	21.4 (716)	+2.8 (+12)
SOA = 1000 ms				
Happiness	2.1 (619)	4.9 (631)	3.8 (597)	+2.8 (+12)
Pleasant surprise	13.8 (667)	18.3 (731)	22.9 (743)	+4.5 (+64)
Sadness	16.3 (710)	15.2 (699)	18.8 (739)	−1.1 (−11)
Anger	20.0 (763)	24.7 (762)	25.0 (706)	+4.7 (−1)
Total	13.1 (690)	15.8 (706)	17.6 (696)	+2.7 (+16)

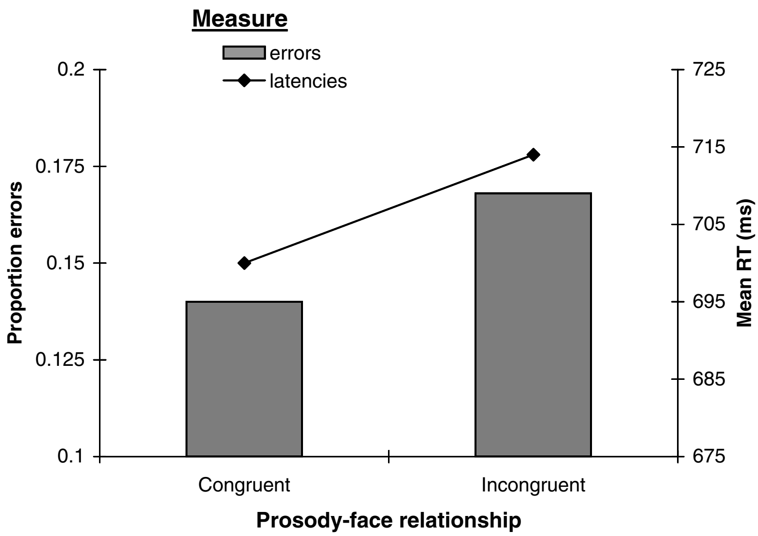

Figure 2. Influence of prosody-face emotion congruity on facial affect decision errors and response times corresponding to 'true' emotional face targets.

short SOA, $F_{errors}(1, 47) = 6.64$, $p = .01$, $r = .35$; $F_{latencies}(1, 32) = 9.08$, $p = .005$, $r = .47$.

Of interest to understanding the impact of different face targets in the FADT, each analysis yielded a prominent main effect for Face emotion, $F_{errors}(3, 141) = 25.10$, $p<.001$; $F_{latencies}(3, 96) = 38.61$, $p < .001$. Irrespective of the prosody relationship, error and latency tendencies were dictated strongly and systematically by the emotional identity of the face; facial affect decisions were slowest and most error prone when evaluating facial expressions of *anger,* and quickest and virtually without error when presented with *happiness*. Accuracy/speed of judgements for *sadness* and *pleasant surprise* assumed an intermediate position within this set for each dependent measure (review Table 3). Performance measures explained by specific face targets were largely stable across SOAs, although *for errors* only, a significant interaction of Face × SOA was observed: *anger* faces were selectively influenced by the prime-target SOA and were significantly less error-prone at the long ($M = .22$) compared to the short ($M = .28$) SOA, $F(3, 141) = 4.76$, $p = .007$. No significant interactions were yielded by the analysis of response latencies for these data (all ps > .01).

Influence of Affective Features of Emotional Displays on Priming

Given evidence that emotional prosody primes decisions about a congruent emotional face in both accuracy and speed, a second pair of analyses was conducted to gauge the possibility that similarities in valence and/or arousal characteristics of prime-target events were critical to understand congruence effects witnessed in the experiment. These analyses also served to disentangle effects in the main analysis related to incongruent trials which reflected two displays that were actually *related* along one of these affective dimensions (although not by emotion). These supplementary findings will prove useful in the future development and interpretation of the FADT by elucidating some of the additional factors that may impact on performance in this new investigative paradigm.

For these analyses, distinctions in SOA—which did not mediate the congruity effect for either performance measure—were collapsed within levels of prosody and face to focus the discussion on effects of greatest theoretical importance. Recall that in a very general manner, each emotion could be recoded according to a unique two-factor representation of the valence and arousal features embedded in the prosody or face, where *happiness:* positive valence-low arousal; *pleasant surprise:* positive valence-high arousal; *sadness:* negative valence-low arousal; *anger:* negative valence-high arousal. Data representing each dependent measure were therefore submitted to separate $2 \times 2 \times 2 \times 2$ ANOVAs involving factorial combinations of Prosody Valence (positive and negative), Prosody Arousal (low, high), Face Valence (positive, negative), and Face Arousal (low, high) as repeated measures. It was anticipated that if the two-way interactions of prosody/face valence or prosody/face arousal were statistically valid without higher-order interactions in the data, these dimensions may be instrumental in explaining cross-modal bias in earlier analyses. However, the emergence of a higher-order (especially four-way) interaction would support the contention of greater emotion-specificity in this form of emotion priming.

For *errors*, main effects in the data were due to Face Valence, $F(1, 47) = 31.75$, $p < .001$, $r = .63$, and Face Arousal, $F(1, 47) = 38.40$, $p < .001$, $r = .67$. As well, the four-way effect of Prosody Valence × Prosody Arousal × Face Valence × Face Arousal was statistically reliable, $F(1, 47) = 8.33$, $p = .006$, $r = .39$. *Post hoc* inspection of the interaction, illustrated in Figure 3, confirmed that face stimuli defined by two-factor combinations of valence and arousal yielded selectively fewer errors when presented with a prosodic stimulus bearing the same combined set of features, consistent with the argument that neither valence nor arousal

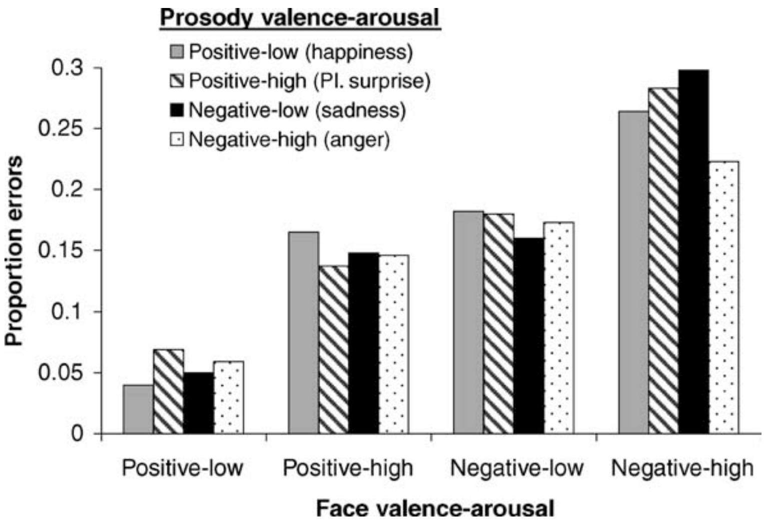

Figure 3. Four-way interaction of face valence, face arousal, prosody valence, and prosody arousal on errors for 'true' emotional face targets when stimuli representing each of the four emotions were recoded, where: *happiness*:positive valence-low arousal; *pleasant surprise*:positive valence-high arousal; *sadness*:negative valence-low arousal; *anger*: negative valence-high arousal.

characteristics of the stimuli strongly biased priming effects in a general, independent manner. The data also highlighted the gross differences in error tendencies for different target faces described earlier (*anger> sadness= pleasant surprise> happiness*), which again could not be reduced to the individual influences of either valence or arousal dimensions of the targets. No two- or three-way interactions attained statistical significance from this analysis.

For *latencies*, the ANOVA again produced main effects for Face Valence, $F(1, 32) = 56.04$, $p <.001$, $r = .80$, and Face Arousal, $F(1, 32) = 28.89$, $p < .001$, $r = .69$. A single interaction was observed involving Prosody Arousal and Face Arousal, $F(1, 32) = 24.65$, $p < .001$, $r = .66$. *Post-hoc* inspection of the interaction revealed that decisions about high (rather than low) arousal faces took consistently longer irrespective of the degree of prosody arousal; however, when high arousal faces were encountered, these decisions were significantly facilitated by a congruent, high arousal prosodic stimulus (see Figure 4). Decisions about low arousal faces were unaffected by low or high arousal features of a prosodic prime

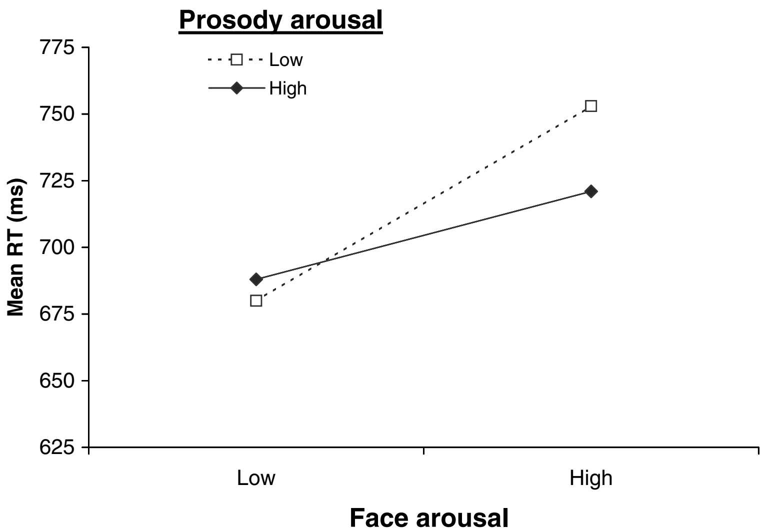

Figure 4. Impact of prosody-face arousal on the speed of correct decisions about 'true' emotional face targets.

based on decision times. There were no higher-order interactions for this analysis, including the four-way effect (p=.59), implying that the arousal component of emotional stimuli assumed a notable and presumably independent role in directing the *speed* of priming latencies across nonverbal channels.[3]

Discussion

Researchers interested in how humans interpret emotional attributes of nonverbal signals have frequently assumed that decoding of speech prosody and facial expressions of emotion rely on overlapping cognitive processing mechanisms (Borod, 2000; Bowers et al., 1993; Rapcsak et al., 1993). These assumptions are supported by limited behavioral evidence of this relationship during on-line processing of nonverbal signals (de Gelder et al., 1999; de Gelder & Vroomen, 2000; Massaro & Egan, 1996) and little specificity of what representational detail is actually shared by these two communicative systems. The present study afforded a preliminary but unprecedented opportunity to register influences of emotional prosody on processing of related and unrelated facial expressions of emotion in a context that did not rely on explicit decisions about underlying affective components of either stimulus type. In this investigation, thorough perceptual characterization of vocal and facial stimuli posed by a common set of 'encoders' (Pell, 2002) ensured that emotion attributes of greatest theoretical interest were rigorously defined for the stimuli, augmenting the strength of initial conclusions about cross-modal connections between prosody and face.

Prosody-Face Interactions in Emotional Processing

The central aim of the study was to illuminate how implicit activations of an unattended *prosodic* stimulus representing a related or unrelated emotion to a face guide cognitive operations leading to recognition of the face as a true emotion. There were clear indications that the prime-target relationship informed both error and latency responses to facial expressions, establishing firmly that emotional prosody biases on-line processing of an adjacent emotional face (de Gelder & Vroomen, 2000; Massaro & Egan, 1996). Specifically, error and latency patterns in the experiment showed highly robust and complementary effects of *cross-modal emotion priming*: when a prosodic prime was emotionally related rather than unrelated to a facial stimulus, emotional-semantic elaboration of the target face was significantly more reliable and efficient. This new evidence of ''nonverbal emotion priming'', which generalized to four distinct and purportedly 'basic' human emotions (Ekman, 1992; Izard, 1992), implies that mechanisms devoted to the nonverbal decoding of emotional prosody and emotion faces are critically bound in some manner, perhaps owing to shared representation of units activated by displays encountered in these separate but frequently conjoined communicative channels (Borod, 2000; Bowers et al., 1993; de Gelder & Vroomen, 2000; Massaro & Egan, 1996). More broadly, evidence of emotional bias between nonverbal stimuli adds to descriptions of affective congruity effects during information processing derived from a variety of sources and for different stimulus types (e.g., Bargh, Chaiken, Raymond, & Hymes, 1996; Fazio et al., 1986; Hermans De Mouwer, & Eelen, 1994; Klauer, Robnagel, & Musch, 1997; Wentura, 1999), extending this evidence to a context restricted to nonverbal stimuli.

In accordance with our hypothesis that nonverbal displays would activate modal or 'prototypical' information about how emotions are expressed in these channels, prosody-face interactions uncovered by our main analysis appeared to be dictated by the discrete emotional characteristics associated with the displays (recall that stimuli were normed and characterized on the basis of their emotion identity). As such, current findings emphasize the central importance of discrete emotion *identities* or 'emotional response categories' to the perceptual analysis and recognition of nonverbal signals (Bower, 1981; Hansen & Shantz, 1995; Niedenthal & Halberstadt, 1995; Niedenthal, Halberstadt, & Setterlund, 1997). The key role of discrete emotion qualities in explaining congruity effects was implied further by subsequent analyses that recoded emotional stimuli according to their valence and arousal components, where there was little evidence that patterns in the performance measures could be reduced to the operation of one or both of these affective qualities in isolation (especially for accuracy where performance was highly differentiated by emotion). One exception to a "specific-emotions" explanation of the priming noted here was that the *arousal* component of nonverbal displays assumed a prominent role in guiding the *speed* of accurate decisions about target faces in certain conditions; when listeners were exposed to a prosodic stimulus for a relatively long duration (i.e., long SOA), response latencies to the face were significantly quicker when prosodic cues were related rather than conflicting in arousal, a pattern that did not depend on additional features of nonverbal displays (i.e., latency patterns could not be further differentiated according to valence or individual emotions defined by combinations of arousal and valence). Pending much further data to clarify the potentially separable influences of discrete emotion qualities and general arousal (or valence) on nonverbal emotion processing, one may speculate that the emotion identity of nonverbal displays was instrumental in producing the congruity effect in the FADT for response accuracy, but that once a "true" emotion was recognized, the speed of these judgements was highly sensitive to underlying conflicts in the arousal of paired emotional displays. The role of the arousal component of nonverbal displays—which is likely to reflect contributions of physiological cues marking the intensity of these expressions *and* relative changes in the psychophysiological status of the decoder in response to these markers (e.g., Scherer, 1989)—would seem to merit a more concerted evaluation in future administrations of the FADT. In particular, rigorous attention to matching prosody-face events according to the degree of subjective arousal experienced by decoders when presented these stimuli, rather than the perceived *intensity* of the emotional meanings identified, could be usefully undertaken in future renditions of this approach.

Irrespective of how (and when) different associations of the stimuli may have biased emotional processing, it is unlikely that prosody-face interactions observed in this administration of the FADT can be explained strictly in terms of passive spreading activation of emotional details about the prosody to the face in some form of associative memory structure (e.g., Bargh et al., 1992; Hermans et al., 1996). A role for attentional factors in understanding congruity effects in our data is dictated by the experimental design. Based on past findings, strategic processing of unambiguous facial displays such as those presented here is known to engender diminished semantic contributions of an adjoining prosodic stimulus (Massaro & Egan, 1996), except in task designs that allow decoders to engage in deliberate processing of prosodic cues to develop strong expectancies about the significance of the target face. In the present experiment, subjects were provided little incentive to allocate controlled processing resources to affective evaluations of the prime, owing to the low relatedness proportion of prime-target trials (Hutchison, Neely, & Johnson, 2001; Neely et al., 1989) and explicit instructions to attend strictly to the face.

Thus, while the significance of prosodic cues was undeniably evaluated to some extent during the on-line procedure—and in a seemingly involuntary manner—given the strong attentional bias to the face in the paradigm, it is likely that (weakly activated) information about the prosodic stimulus influenced target processing *retrospectively*, by means of a comparative procedure that impaired decisions when prosody-face activations in memory conflicted in critical ways (incongruent trials) or which allowed decisions about the target to proceed relatively normally when activations for the two events overlapped (congruent trials) (see Massaro & Egan, 1996). This explanation may also account for the observation that congruity effects were stable throughout the study but relatively small in *magnitude* (results of a forthcoming study reveal that emotion congruity effects are much larger in magnitude when the prosody-face relatedness proportion is increased, due to greater incentive for subjects to allocate attention to features of the prime, see Pell, under review). Obviously, these initial conclusions await further data that clarify how automatic and attentional factors guide affective congruity effects during nonverbal processing, a debate that is prominent in the wider literature on affective priming (see Fazio, 2001; Klauer & Musch, 2002; Musch & Klauer, 2001).

Impact of Target Expressions on Facial Affect Decisions

The sensitivity of the FADT to emotionally meaningful faces allows for commentary on how different emotional expressions influenced analysis of face targets independent of the prosody-face relationship. Given the focus of this evaluative paradigm on attention to the *face*, it is not surprising that parameters of these events served a dominant role in explaining performance outcomes in the experiment (Massaro & Egan, 1996; Niedenthal & Halberstadt, 1995) with little evidence that variations in prosody alone affected the dependent measures overall.

A consistent finding was that errors and response times varied in a systematic and complementary manner according to the identity of the face under scrutiny: these measures were always lowest for faces expressing *happiness* and greatest for faces expressing *anger*, with targets depicting *pleasant surprise* and *sadness* assuming an intermediate (but not identical) position in most analyses. These primary influences of the emotional target identity on subsequent processing were stable throughout the experiment and at both the short and long SOA (discussed below). The apparent ease by which *happiness* faces were accepted as 'true' exemplars of emotion was especially marked, with mean overall error rates that were three to five times lower, and response times that were 70–130 ms shorter on average, than for the other three target emotions (cf. de Gelder & Vroomen, 2000; Kirouac & Dore, 1984). Processing of *anger* target faces was also highly discrepant but in the opposite direction, eliciting extremely high error rates (26.7% versus 5.5% for *happiness* overall) and prolonged response times across conditions. In fact, it is curious that recognition of *anger* faces evaluated under unique on-line demands of the FADT was selectively less accurate than established for these stimuli offline prior to the experiment (Pell, 2002).

Thus, error and latency patterns in the FADT demonstrate that unambiguous tokens of facial emotion do not engage the emotion processing system in a uniform manner (see De Houwer & Eelen, 1998; Halberstadt, Niedenthal, & Kushner, 1995; Kitayama, 1996; Niedenthal, 1990; Pratto & John, 1991 for related discussion of emotion asymmetries in studies of affective priming). Within the limited sample of emotions tested, a more general observation that can be made is that facial expressions associated with a negative appraisal or those associated with relatively high arousal were always more prone to error and took longer to accept as a 'true' emotion than faces bearing a positive evaluation or low emotional arousal. These observations fit long-standing claims that perceptual events are encoded with respect to their pleasantness/unpleasantness (Bargh, 1992; Frijda, 1993; LeDoux, 1989) and their level of emotional arousal (Kitayama & Howard, 1994; Öhman & Soares, 1994; Robinson, 1998; Scherer, 1986), and that such activations are likely present in the context of assigning emotional significance to an impinging face stimulus. Indeed, it is noteworthy that when exaggerated "YES" error tendencies witnessed for eleven of our 48 subjects were examined across target categories, these participants rejected the acceptability of nearly half of all negative faces (leading to the exclusion of these individuals from latency analyses) but committed very few errors when presented with positive-valenced emotional faces. These patterns reinforce claims regarding the pronounced disruptive effect of encountering negative information, especially *anger*, on subsequent processes (Isen, 1984; Niedenthal, 1990; Pratto, 1994) and shows that these evaluations are indexed to some extent by the FADT.[4]

An explanation for why *anger* faces were associated with notably greater interference in accuracy and speed of rendering a facial affect decision is that these stimuli were strongly conducive to an evaluation of *threat*, leading to narrowed and sustained attentional focus to these tokens which impaired subsequent stages of task processing (Dimberg & Öhman, 1996; Niedenthal, 1990; Williams et al., 1996). Automatic attentional biases that capture and sometimes direct attention away from immediate task goals have been shown for a range of negative stimuli, but are triggered potently by faces conveying salient cues to threat such as *anger* or *fear* (Dimberg & Öhman, 1996; Mathews & Mackintosh, 1998; Mogg & Bradley, 1999; Öhman, 1992; van Honk, Tuiten, de Haan, van den Hout, & Stam, 2001). Deployment of attentional resources to negative information in *anger* faces may have inhibited/delayed schematic processes needed to classify these targets as 'true' representations of emotion. Alternatively, it is possible that interference produced by potentially threatening stimuli was introduced in part at so-called "post-lexical" stages of producing a "YES" response. According to this account, heightened perceptual focus to "negative" detail in threatening faces may have promoted response path interference on many trials, whereby strong "negative" evaluations primed a "NO" response in the paradigm (see Klauer & Musch, 2002; Wentura, 2000 for recent discussions). If contributions at a post-lexical stage account for some of the patterns in the current data, clearly the potency of *anger* expressions amplified these response tendencies over faces of *sadness* in the FADT (review Figure 3). Identifying the extent to which target features may have biased judgmental tendencies at late stages of response generation in the FADT remains an important issue, which can be usefully informed by *inverting* the target-response relationship in the task (such that "true" faces generate a "NO" response), as accomplished recently for word stimuli by Wentura (2000).

The FADT as a Sensitive Index of Emotional "Acceptability"

Establishing the utility of the FADT as a future approach to tap on-line activations of nonverbal displays in a variety of subject groups constituted a further aim of this report. Recall that the FADT requires subjects to visually attend to a static facial stimulus with the goal of determining its emotional-symbolic value or acceptability as a 'true' communicative exemplar of emotion, mimicking certain properties of the visual lexical decision paradigm (Swinney, 1979). It was hypothesized that facial affect decisions about a visual facial configuration would index those procedures that activate stored detail of how 'true' emotions are encoded in the face, at least sufficient to yield familiarity with the communicative value of the target display. Critically, these properties would not be associated with facial "grimaces" or 'false' face targets for which underlying associations do not exist, revealing a differentiation of responses to valid and invalid facial representations of emotion in the experiment as a whole.

While overall accuracy patterns in the experiment highlighted that some participants were highly prone to err in conditions involving negative face targets, it should be emphasized that subjects experienced little difficulty understanding the concept of 'true' versus 'false' displays of facial emotion following appropriate instructions and practice (Rapcsak et al., 1993). Of all participants entered into the study, the communicative value of approximately 85% of 'true' facial expressions of emotion was appropriately recognized and well in excess of 90% of 'false' targets were correctly rejected, showcasing a reliable capacity for humans to judge the emotional-symbolic nature of facial expressions, akin to procedures that enable accurate "lexical decisions" or "grammaticality" judgements in areas of shared linguistic competence. In fact, 'false' facial expressions were rarely interpreted as emotionally relevant despite the impression that many nonsense faces bore explicit affective qualities of some nature (review Figure 1). The fact that 'true' and 'false' portrayals of emotion could be robustly differentiated overall, with evidence that 'true' emotional targets were uniquely sensitive to affective manipulations in the experiment, justify the rationale for eliciting "facial affect decisions" as an effective means for tapping cognitive-regulatory processes triggered by communicatively *valid* facial expressions of emotion. These data call for continued development of the FADT by researchers in different settings who may modify the technique in theoretically motivated ways, contributing to a more sophisticated database on how humans process nonverbal displays in real time.

In asserting the vast potential of the FADT to future research, no assumptions are made that decoders who take part experience the target emotions of interest or any other feeling states while performing the task, nor that such a process is necessarily required to retrieve associative knowledge about communicative displays (cf. Adolphs, 2002). Note as well that in manipulating the prime-target SOA within the experiment, the current design does not advocate any strong claims about possible *automatic* influences of prosody on face processing (e.g., Shiffrin & Schneider, 1977) since emotionally-inflected utterances in this procedure continued to be presented following onset of the facial target until a button press was registered. A test of automatic activations of emotional prosody on a concurrent face could be accomplished by *arresting* presentation of the auditory stimulus at the short (300 ms) onset of the facial target in future applications of the FADT, yielding more precise conclusions about *how much* prosodic information is needed to bias emotional processing of a related face (Pell, under review).

Further on-line testing will permit increasingly sophisticated inferences about the cognitive nature of social interactions involving emotional prosody and emotional faces, and how representational detail activated by these different sources of information interact and are combined in the comprehension process. The manner in which ''congruence'' was defined and analyzed in this experiment, and the specificity of many of our effects to single target emotions (such as *anger*), allow for the opinion that nonverbal emotion priming is tied to the discrete identities or 'emotion response categories' associated with both prosody and face stimuli (Bower, 1981; Hansen & Shantz, 1995; Niedenthal & Halberstadt, 1995). Nonverbal signals constitute primary channels for communicating emotion to conspecifics (Ekman, 1992; Izard, 1992) and these displays may be decoded in reference to 'prototypical' or schematic knowledge about differentiated emotion states and their expression (Etcoff & Magee, 1992; Massaro & Egan, 1996; Young et al., 1997). By eliciting facial affect decisions in highly controlled processing environments, it would appear that investigators can index the presumably automatic evaluations of valence and arousal components of nonverbal displays, but also emotion-specific knowledge that has adapted for interpreting nonverbal signals (Hansen & Shantz, 1995; Niedenthal & Halberstadt, 1995) and which is partly shared across nonverbal channels.

Notes

1. Three facial and five prosodic stimuli that did not meet entry criteria were included to ensure balanced representation of tokens across conditions as a function of encoder sex. Criteria for the entry of vocal stimuli reflect the fact that perceptual raters commonly

arrive at a reduced consensus when required to assign discrete emotion labels to prosodic as opposed to facial stimuli (Pell, 2002; Scherer et al., 2001; Wurm et al., 2001).

2. Eleven of these subjects failed to meet the accuracy criterion (< 25% errors for ''YES'' trials) for inclusion in latency analyses due to an excessive bias to reject the validity of 'true' emotional faces overall: for this group, $M = 33.5\%$ errors, $SD = 8.3$, range = 25.1–50.4. Four additional participants met the accuracy criterion for inclusion in latency analyses but each produced no correct decisions in one of the prosody-face conditions, yielding an empty cell (three of these occurred in the *sadness–anger* prime-target condition and one occurred for *happiness–sadness*). Data for these subjects could not be entered in the within-subjects factorial design for latencies.
3. In light of the number of subjects who were excluded from latency analyses due to errors, the two ANOVAs performed on the latency measures were rerun on the full sample of 48 subjects, excluding only subjects who had empty cells and could not be entered into the full factorial analysis (two and five subjects, respectively). The main $2 \times 4 \times 2$ ANOVA produced the three main effects of prosody-face Relationship, SOA, and Face which all remained highly significant, all Fs > 8.6, all ps < .005. No significant interactions were again observed for this analysis. The secondary $2 \times 2 \times 2 \times 2$ ANOVA run on the data coded according to the valence/arousal dimensions of the Prosody and Face replicated the main effects for Face Valence and Face Arousal and the significant interaction of Face Arousal × Prosody Arousal, all Fs > 33.4, all ps < .001. A main effect of Prosody Arousal, which contributes to the interaction, was also significant for this analysis when it was rerun on the larger sample, $F(1, 42) = 7.59$, $p = .01$, $r = .39$.
4. As noted by an anonymous reviewer, differences in performance measures attributed to the *valence* of face targets may have been influenced somewhat by systematic variations in the brightness/luminosity of positive versus negative targets, a factor that was not carefully controlled in this study. The potential role of this factor in understanding the Face main effect awaits future clarification, although differences in the luminosity of face targets would have no bearing on how attributes of the prosodic stimulus biased processing of different face targets, i.e., the interaction of prosody and face**,** which revolved necessarily around emotional aspects of these stimuli.

References

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*, 169–177.

Bargh, J. (1988). Automatic information processing: Implications for communication and affect. In L. Donohew, H. Sypher, & E. T. Higgins (Eds.), *Communication, social cognition and affect* (pp. 9–32). Hillsdale, NJ: Erlbaum.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce processing effects. *American Journal of Psychology, 105*, 181–199.

Bargh, J., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*, 893–912.

Bargh, J., Chaiken, S., Raymond, P., & Hymes, C. (1996). The automatic evaluation effect: Unconditional automatic attitude activation with a pronunciation task. *Journal of Experimental Social Psychology, 32*, 104–128.

Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology, 13*, 581–608.

Bimler, D., & Kirkland, J. (2001). Categorical perception of facial expressions of emotion: Evidence from multidimensional scaling. *Cognition and Emotion, 15*, 633–658.

Borod, J. (2000). Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition and Emotion, 14*, 193–211.

Borod, J. C., Cicero, B. A., Obler, L. K., Welkowitz, J., Erhan, H. M., Santschi, C., Grunwald, I. S., Agosti, R. M., & Whalen, J. R. (1998). Right hemisphere emotional perception: Evidence across multiple channels. *Neuropsychology, 12*, 446–458.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36*, 129–148.

Bowers, D., Bower, R., & Heilman, K. (1993). The nonverbal affect lexicon: Theoretical perspectives from neuropsychological studies of affect perception. *Neuropsychology, 7*, 433–444.

de Gelder, B., Bocker, K. B. E., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: Early interaction revealed by human brain responses. *Neuroscience Letters, 260*, 133–136.

de Gelder, B., Pourtois, G., Vroomen, J., & Bachoud-Levi, A. C. (2000). Covert processing of faces in prosopagnosia is restricted to facial expressions: Evidence from cross-modal bias. *Brain and Cognition, 44*, 425–444.

de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion, 14*, 289–311.

De Houwer, J., & Eelen, P. (1998). An affective variant of the Simon Paradigm. *Cognition and Emotion, 12*, 45–61.

Dimberg, U., & Öhman, A. (1996). Beyond the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion, 20*, 149–182.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Ekman, P., & Friesen, W. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologist's Press.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*, 86–88.

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

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition and Emotion, 15*, 115–141.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology, 50*, 229–238.

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affects. In R. M. Sorrentino, & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 167–203). New York, NY: Guilford.

Frijda, N. H. (1993). The place of appraisal in emotion. *Cognition and Emotion, 7*, 357–387.

Halberstadt, J., Niedenthal, P., & Kushner, J. (1995). Resolution of lexical ambiguity by emotional state. *Psychological Science, 6*, 278–282.

Hansen, C. H., & Shantz, C. A. (1995). Emotion-specific priming: Congruence effects on affect and recognition across negative emotions. *Personality and Social Psychology, 21*, 548–557.

Hermans, D., De Houwer, J., & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition and Emotion, 8*, 515–533.

Hermans, D., De Houwer, J., & Eelen, P. (1996). Evaluative decision latencies mediated by induced affective states. *Behaviour Research and Therapy, 34*, 483–488.

Hutchison, K. A., Neely, J. H., & Johnson, J. D. (2001). With great expectations, can two "Wrongs" prime a "Right"? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1451–1463.

Isen, A. (1984). Toward understanding the role of affect in cognition. In R. S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 179–236). Hillsdale, NJ: Erlbaum.

Izard, C. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review, 99*, 561–565.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press.

Kirouac, G., & Dore, F. (1984). Accuracy and latency of judgment of facial expressions of emotion. *Perceptual Motor Skills, 57*, 683–686.

Kitayama, S. (1996). Remembrance of emotional speech: Improvement and impairment of incidental verbal memory by emotional voice. *Journal of Experimental Social Psychology, 32*, 289–308.
Kitayama, S., & Howard, S. (1994). Affective regulation of perception and comprehension: Amplification and sematic priming. In P. M. Niedenthal, & S. Kitayama (Eds.), *The heart's eye: Emotional influences in perception and attention* (pp. 41–65). San Diego: Academic Press.
Klauer, K. C., & Musch, J. (2002). Goal-dependent and goal independent effects of irrelevant evaluations. *Personality and Social Psychology Bulletin, 28*, 802–814.
Klauer, K. C., Robnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 246–255.
LeDoux, J. (1989). Cognitive-emotional interactions in the brain. *Cognition and Emotion, 3*, 267–289.
Massaro, D., & Egan, P. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review, 3*, 215–221.
Mathews, A., & Mackintosh, B. (1998). A cognitive model of selective processing in anxiety. *Cognitive Therapy and Research, 22*, 539–560.
Max, L., & Onghena, P. (1999). Some issues in the statistical analysis of completely randomized and repeated measures designs for speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research, 42*, 261–270.
Mogg, K., & Bradley, B. P. (1999). Orienting of attention to threatening facial expressions presented under conditions of restricted awareness. *Cognition and Emotion, 13*, 713–740.
Murphy, S., & Zajonc, R. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology, 64*, 723–739.
Musch, J., & Klauer, K. C. (2001). Locational uncertainty moderates affective congruency effects in the evaluative decision task. *Cognition and Emotion, 15*, 167–188.
Neely, J. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In Besner (Ed.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1003–1019.
Niedenthal, P. M. (1990). Implicit perception of affective information. *Journal of Experimental Social Psychology, 26*, 505–527.
Niedenthal, P. M., & Halberstadt, J. B. (1995). The acquisition and structure of emotional response categories. *The Psychology of Learning and Motivation, 33*, 23–64.
Niedenthal, P. M., Halberstadt, J. B., & Setterlund, M. B. (1997). Being happy and seeing "happy": Emotional state mediates visual word recognition. *Cognition and Emotion, 11*, 403–432.
Öhman, A. (1987). The psychophysiology of emotion: An evolutionary-cognitive perspective. *Advances in Psychophysiology, 2*, 79–127.
Öhman, A. (1992). Orienting and attention: Preferred preattentive processing of potentially phobic stimuli. In B. Campbell, H. Haynes, & R. Richardson (Eds.), *Attention and information processing in infants and adults: Perspectives from human and animal research* (pp. 263–295). Hillsdale, NJ: Erlbaum.
Öhman, A., & Soares, J. (1994). Unconscious anxiety: Phobic reponses to masked stimuli. *Journal of Abnormal Psychology, 103*, 231–240.
Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory and Cognition, 9*, 225–236.
Pell, M. D. (2002). Evaluation of nonverbal emotion in face and voice: Some preliminary findings on a new battery of tests. *Brain and Cognition, 48*, 499–504.

Pell, M. D. (under review). Prosody-face interactions in emotional processing as revealed by the facial affect decision task.

Pell, M. D., & Baum, S. R. (1997). The ability to perceive and comprehend intonation in linguistic and affective contexts by brain-damaged adults. *Brain and Language, 57*, 80–99.

Pratto, F. (1994). Consciousness and automatic evaluation. In P. Niedenthal, & S. Kitayama (Eds.), *The heart's eye* (pp.115–143). New York: Academic Press.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*, 380–391.

Rapcsak, S., Comer, J., & Rubens, A. (1993). Anomia for facial expressions: Neuropsychological mechanisms and anatomical correlates. *Brain and Language, 45*, 233–252.

Robinson, M. (1998). Running from William James' bear: A review of preattentive mechanisms and their contributions to emotional experience. *Cognition and Emotion, 12*, 667–696.

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*, 143–165.

Scherer, K. R. (1989). Vocal correlates of emotional arousal and affective disturbance. In H. Wagner & A. Manstead (Eds.), *Handbook of social psychophysiology* (pp. 165–197). Oxford, England: Wiley.

Scherer, K. R. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion, 7*, 1–41.

Scherer, K. R., Banse, R., & Wallbott, H. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*, 76–92.

Scherer, K. R., & Wallbott, H. G. (1994). Evidence for universality and cultural variation of differential emotion response patterning. *Journal of Personality and Social Psychology, 66*, 310–328.

Schirmer, A., Kotz, S., & Friederici, A. (2002). Sex differentiates the role of emotional prosody during word processing. *Cognitive Brain Research, 14*, 228–233.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attention and a general theory. *Psychological Review, 84*, 127–190.

Swinney, D. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior, 18*, 645–659.

van Honk, J., Tuiten, A., de Haan, E., van den Hout, M., & Stam, H. (2001). Attentional biases for angry faces: Relationships to trait anger and anxiety. *Cognition and Emotion, 15*, 279–297.

Wentura, D. (1999). Activation and inhibition of affective information: Evidence for negative priming in the evaluation task. *Cognition and Emotion, 13*, 65–91.

Wentura, D. (2000). Dissociative affective and associative priming effects in the lexical decision task: Yes versus no responses to word targets reveal evaluative judgement tendencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 456–469.

Williams, J., Matthews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120*, 3–24.

Wurm, L. H., Vakoch, D. A., Strasser, M. R., Calin-Jageman, R., & Ross, S. E. (2001). Speech perception and vocal expression of emotion. *Cognition and Emotion, 15*, 831–852.

Young, A., Rowland, D., Calder, A., Etcoff, N., Seth, A., & Perrett, D. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition, 63*, 271–313.