UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Temporal Characteristics of Categorical Perception of Emotional Facial Expressions

Permalink
https://escholarship.org/uc/item/0nw113h7

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN
1069-7977

Authors
Suzuki, Atsunobu
Shibui, Susumu
Shigemasu, Kazuo

Publication Date
2004

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Temporal Characteristics of Categorical Perception of Emotional Facial Expressions

Atsunobu Suzuki (suzuki@bayes.c.u-tokyo.ac.jp)
Department of Cognitive and Behavioral Science,
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

Susumu Shibui (shibui@bayes.c.u-tokyo.ac.jp)
Department of Cognitive and Behavioral Science,
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

Kazuo Shigemasu (kshige@bayes.c.u-tokyo.ac.jp)
Department of Cognitive and Behavioral Science,
3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan

Abstract

Perceptual processing of emotional facial expressions occurs very quickly, even without awareness. To determine whether the fast processing of facial expressions is categorical, we studied temporal characteristics of categorical perception (CP) of facial expressions. We investigated the effect of shortening stimulus duration on participant performance with respect to identifying and discriminating morphed facial expressions. The results of two experiments showed that CP was attenuated or even disappeared when facial stimuli were presented for as briefly as 50-75 milliseconds. These findings indicate that CP is irrelevant to the fast perceptual processing of facial expressions.

Facial expressions of emotion are processed very quickly, even without awareness (Morris, Ohman, & Dolan, 1998; Whalen et al., 1998). It has been proposed that this fast emotional processing provides a 'dirty' image of the external world, enabling organisms to detect salient stimuli immediately (Adolphs, 2002; LeDoux, 1996). What is still not understood, however, is how this fast and dirty processing of facial expressions works.

A hallmark of facial expression recognition is its categorical nature, that is, facial expressions are recognized as belonging to discrete categories of emotion, so-called *basic emotions* (e.g., Ekman, 1992; Izard, 1992). Reports on categorical perception (CP) of facial expressions are thought to provide strong evidence that people process facial expressions categorically (Calder et al., 1996; DeGelder, Teunisse, & Benson, 1997; Etcoff & Magee, 1992; Young et al., 1997). Previous studies have confirmed that recognizing facial expressions fulfills the following features of CP: (a) in identifying a stimulus within a continuum extending from one category to another, the rate of categorizing the stimulus changes abruptly at a boundary (category boundary); and (b) in discriminating a pair of stimuli that differ by a constant physical amount, discrimination is superior for pairs straddling the category boundary (between-category pairs), as compared to pairs falling within one category (within-category pairs). It has been argued that the perceptual system transforms the information continuously received from a given facial expression into categorical information corresponding to the most likely emotion (Etcoff & Magee, 1992).

In previous research, however, facial stimuli were presented for a relatively long period ($\geqq$750 ms), so it remains unclear whether the fast perceptual processing of facial expressions is categorical in nature. Our goal was to investigate the effects of shortening stimulus duration on CP of facial expressions and to examine whether the fast perceptual processing is categorical.

Experiment 1

In Experiment 1, participants identified facial expressions from three continua extending from anger to happiness, from happiness to fear, and from fear to anger. We investigated how shortening stimulus exposure duration (750 ms, 150 ms, 50 ms) affected the abrupt change in the rate of identification at the category boundary.

Method

Participants Twelve undergraduates and postgraduates participated (8 men, 4 women; 20-24 years old).

Facial Stimuli Three sets of eight gray-scale images of facial expressions were used. Each set consisted of a continuum extending either from anger to happiness, from happiness to fear, or from fear to anger. The original (endpoint) facial expressions were posed photographs of a Japanese woman, and interpolated (morphed) expressions between the two continuum endpoints were created with software for facial expression processing (Information-technology Promotion Agency, Japan (IPA), 1998).

The morph transformation started with the delineation of two endpoint expressions; landmarks were placed manually on critical positions of each image. There were 759 landmarks in total, 88 points of which were placed manually on corresponding positions of each image: for the head, 4 points; for the outline, 28 points; for the eyes, 5 x 2 points; for the eyebrows, 4 x 2 points; for the nose, 4 points; for the

mouth, 6 points; for the neck, 13 points; and finally for the hairline, 15 points. An intermediate face was then created with linear interpolation between point-to-point pixel intensity values, yielding a weighted blend of both facial configuration and texture of the two endpoint faces. Six intermediate faces were generated for each continuum, so that eight images from one endpoint to the other were spaced with a 14.3% gap.

Procedure and Design Participants performed the identification task (three sessions) by viewing stimuli presented on a 17-inch CRT monitor. Each trial began with an 800-ms presentation of a fixation point, followed by a blank interval of 300 ms, and then a facial stimulus from one continuum for 750 ms, 150 ms, or 50 ms. After a 300-ms blank interval, participants were asked to decide which endpoint category the face expressed. The continuum was manipulated across sessions: from anger to happiness (AH session), from happiness to fear (HF session), and from fear to anger (FA session). The order of the three sessions was counterbalanced across participants. Each session began with 24 training trials, followed by five blocks of 48 experimental trials, corresponding to two repetitions of the 24 combinations of faces (8) and stimulus durations (3). The order of the face-duration sets was fully randomized in each repetition.

Results & Discussion

Figure 1 displays the overall percentage with which a given endpoint expression was identified at each stimulus duration, for each continuum. As an estimate of the category boundary, the point at which two endpoints were identified with equal probability was computed by applying a logit model to each identification rate curve (Table 1). In general, the category boundary lay near the 50% morph. More importantly, Table 1 also indicates the slope of the logistic curve at the category boundary as a measure of the identification rate change at the boundary. Multiple comparisons (z tests) with Bonferroni's method (α=0.017) revealed that the slope at 50 ms was significantly, or marginally significantly, smaller than at 150 ms and at 750 ms for the anger-happiness and happiness-fear continua (AH session, 50 ms vs. 150 ms, z=2.37, p=0.018, 50 ms vs. 750 ms, z=3.67, p<0.001, 150 ms vs. 750 ms, z=1.48, p=0.139; HF session, 50 ms vs. 150 ms, z=2.69, p=0.007, 50 ms vs. 750 ms, z=4.07, p<0.001, 150 ms vs. 750 ms, z=1.74, p=0.082). No slope change was observed for the fear-anger continuum (all ps>0.35).

The decreased slopes of the identification rate curves for the anger-happiness and happiness-fear continua at 50 ms indicate that the CP of facial expressions is attenuated with the briefest stimulus exposure duration. The lack of slope change for the fear-anger continuum may be due to the originally weak categorical perception between fear and anger, which reflects their remarkable similarity with respect to the dimensional information of pleasure and arousal (Russell, 1997). Indeed, the slope for the fear-anger continuum was the smallest of the three continua at 750 ms. Likewise, there are previous reports that have failed to detect clear CP for the fear-anger continuum (Calder et al., 1996; DeGelder et al., 1997).

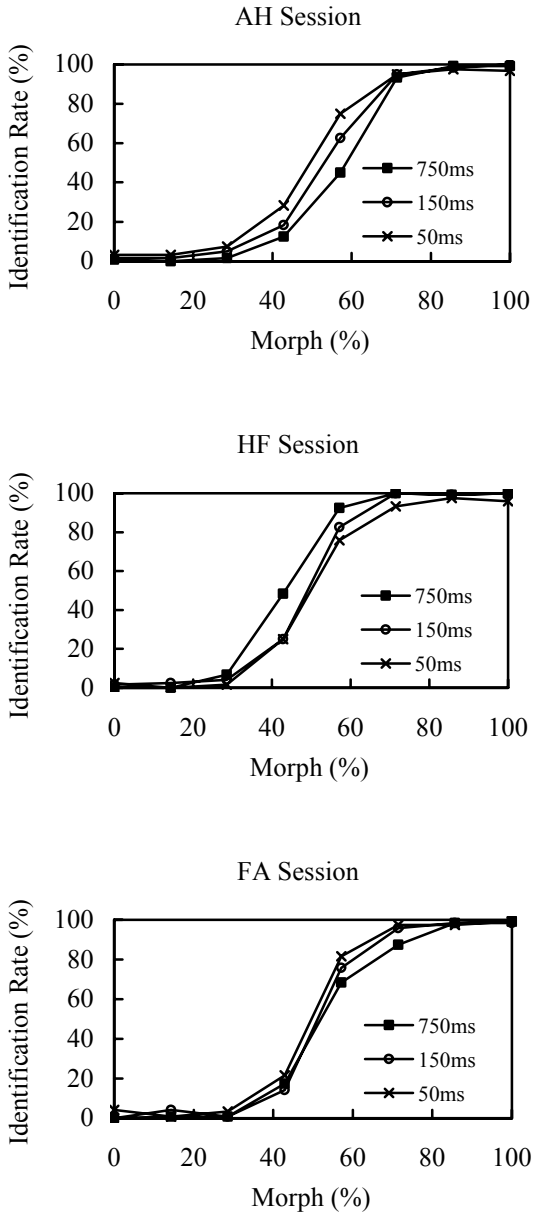


Figure 1: Overall percentage with which a given endpoint expression (top: happiness, middle: fear, bottom: anger) was identified at each stimulus duration, for each session. *Morph* represents a percentage of happiness (top), fear (middle), or anger (bottom) for a given face.

Table 1: Estimates of the position of the category boundary and the slope at the boundary.

Session	Duration	Boundary (%)	Slope
AH	750 ms	56.9	14.5
	150 ms	52.5	12.5
	50 ms	49.0	10.0
HF	750 ms	43.3	18.1
	150 ms	47.9	14.9
	50 ms	51.2	11.3
FA	750 ms	53.9	13.3
	150 ms	51.8	13.8
	50 ms	49.2	12.6

Note *Boundary* represents a percentage of happiness (AH Session), fear (HF Session), or anger (FA Session) for a morphed face.

Experiment 2

Experiment 1 revealed that the abrupt change in identification rate at the category boundary was attenuated at the briefest stimulus exposure duration, indicating that CP of facial expressions did not reflect fast perceptual processing. To further examine temporal characteristics of CP, Experiment 2 investigated whether superior discrimination for between-category pairs could be observed with the brief stimulus exposure duration when compared to discrimination for within-category pairs. Because a more accurate discrimination performance for between-category pairs is regarded as the key indicator of CP (Harnad, 1987), its disappearance at the brief stimulus duration provides crucial evidence that categorical perception does not occur rapidly. In Experiment 2, participants engaged in both discrimination and identification tasks of facial expressions from the happiness-fear continuum. We selected the happiness-fear continuum from the three continua used in Experiment 1 because it did not incorporate any gross changes in physical features (e.g., from open to closed mouth), which might obscure CP (Calder et al., 1996).

Method

Participants Twenty-three undergraduates participated (12 men, 11 women; 18-24 years old).

Facial Stimuli One set of 11 gray-scale images of facial expressions was used. The set consisted of a continuum extending from happiness to fear. The endpoint expressions and the morphing procedure to create interpolated faces were the same as in Experiment 1. Nine intermediate faces were generated so that the 11 images from happiness to fear were spaced with a 10% gap.

Procedure and Design Participants performed two successive tasks (two sessions each) by viewing stimuli presented on a 17-inch CRT monitor.

XAB discrimination task The sequence of the stimuli presented in each trial was as follows: (1) a fixation point for 450 ms; (2) a blank interval for 300 ms; (3) a facial stimulus 'X' for 150 ms or 75 ms (target); (4) a black-and-white checker-pattern for 150 ms (backward mask); (5) a blank interval for 750 ms; (6) and finally, two facial stimuli 'A' and 'B', positioned horizontally to the right and left of center, displayed until the participants responded (reference). Participants were asked to decide which reference was identical to the target.

The backward mask was used to restrict visual access to the target over the controlled stimulus exposure duration (Enns & DiLollo, 2000). Because the backward mask might degrade target perception, we used 75 ms as the brief exposure duration, which is somewhat longer than the 50 ms tested in Experiment 1. We also used 150 ms as the sole longer exposure duration interval, because we speculated that task difficulty might differ too much between 75 ms and 750 ms. Reference stimuli were spaced with a 20% gap, resulting in nine possible pairs. For each pair, there were four presentation orders; (X,A,B) = (A,A,B), (A,B,A), (B,A,B), (B,B,A). Target duration was manipulated across sessions, and the order of the two durations was counterbalanced across participants. Each session contained 10 training trials and two blocks of 36 experimental trials representing all combinations of pairs (9) and presentations (4). The order of the pair-presentation sets was fully randomized in each block.

Identification task Stimulus sequence in each trial was the same as in the XAB discrimination task, with the exception that reference stimuli were removed. Participants were asked to decide whether the target stimulus expressed fear or happiness. The order of the two sessions across which the target duration was manipulated was the same as in the XAB discrimination task. Each session contained 10 training trials and two blocks of 44 experimental trials corresponding to four repetitions of the 11 faces. The order of the faces was fully randomized in each repetition.

Results & Discussion

Identification Task Figure 2 presents the overall percentage of trials in which "fear" was identified at each stimulus exposure duration in the identification task. Application of a logit model to each identification rate curve revealed that the category boundary lay at 52.4% (150 ms) and 55.2% (75 ms). Contrary to the results of Experiment 1, logit models also showed that the slope at 150 ms (17.9) was somewhat smaller than at 75 ms (19.4), though this difference did not reach significance (z=1.03, p=0.303). As the slope was attenuated in proportion to the decreased duration in Experiment 1, these data suggest that the duration difference between 150 ms and 75 ms was insufficient to cause a significant slope change.

XAB Discrimination Task Figure 3 presents the overall correct response rate for each stimulus exposure duration in

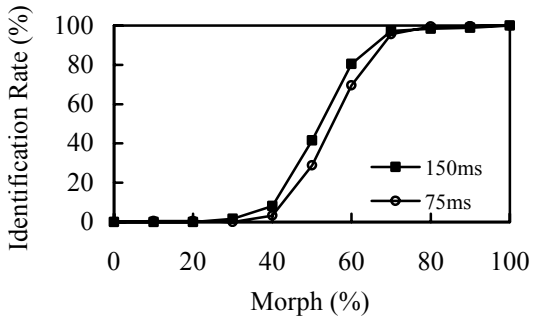


Figure 2: Overall percentage with which "fear" was identified at each stimulus duration. *Morph* represents a percentage of "fear" for a given face.

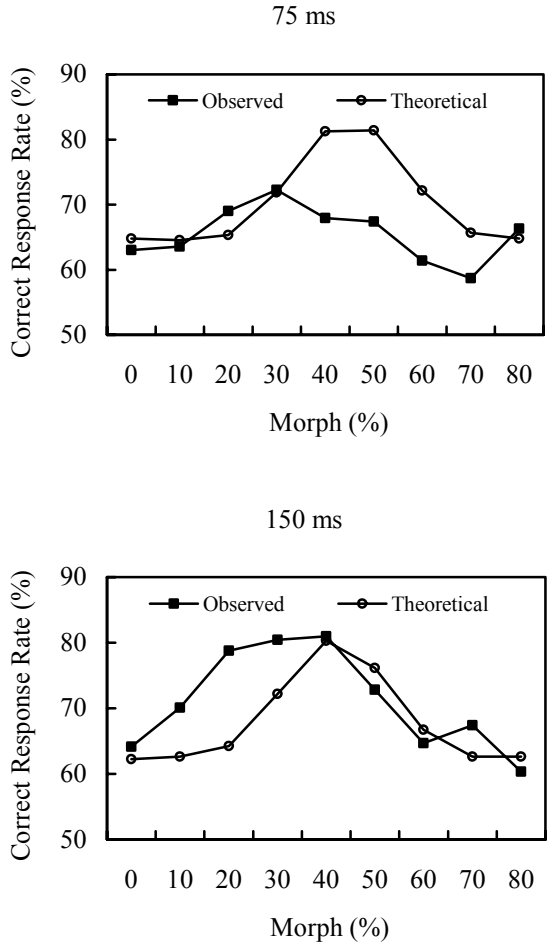


Figure 3: Overall correct response rate in the XAB discrimination task for each duration. *Observed*=observed rate. *Theoretical*=theoretical rate predicted from CP. *Morph* represents a percentage of "fear" for the less fearful face of a given pair.

the XAB discrimination task. The data from the identification task revealed that identifying between-category pairs was 40%-60% and 50%-70% at both durations. The peak in discrimination performance at the category boundary was found only at the 150-ms stimulus exposure duration.

We calculated the mean correct rate for between-category pairs (40%-60%, 50%-70%) and within-category pairs (the remaining seven pairs) for each duration, and conducted a 2 x 2 ANOVA of the mean correct rate with the two factors of pair (between-category or within-category) and stimulus duration (150 ms or 75 ms). A significant main effect of pair was found ($F(1,22)=7.79$, $p=0.011$, $MSE=0.008$), indicating better discrimination for between-category pairs (72%) than for within-category pairs (67%). The main effect of duration was also significant ($F(1,22)=9.74$, $p=0.005$, $MSE=0.011$), indicating worse performance at 75 ms (66%) than at 150 ms (73%). There was no significant interaction between the two factors ($p>0.25$), suggesting that discrimination of between-category pairs was more accurate than that for within-category pairs, at both durations. However, post-hoc paired *t* tests revealed that superior discrimination for between-category pairs was significant only at the 150-ms exposure duration (150 ms, $M=8\%$, $t(22)=2.42$, $p=0.024$; 75 ms, $M=3\%$, $t(22)=1.18$, $p=0.252$). Post-hoc comparisons also revealed that poorer discrimination at 75 ms was significant for both between-category and within-category pairs (between-category, $M=9\%$, $t(22)=2.36$, $p=0.027$; within-category, $M=4\%$, $t(22)=2.66$, $p=0.014$).

To indicate CP in the discrimination task, previous studies (Calder et al., 1996; Young et al., 1997) have reported correlations between observed and theoretical performances, predicted from CP. Calder and colleagues (Calder et al., 1996) assumed that discrimination between two facial expression stimuli depends on two cues: first, the physical difference between the stimuli, which is constant for any pair, regardless of their expressions; and second, the expression categories of the stimuli. To estimate the contribution of the first non-categorical cue, they computed the mean observed discrimination for the two pairs, placed at both ends of the continuum. To estimate the contribution of the second categorical cue, they calculated the difference in identification rates for the two relevant stimuli observed in the identification task, and multiplied the obtained difference by 0.25 (a constant). Calder et al. (1996) argued that summing the two estimates measures theoretical performance in the discrimination task.

Figure 3 also presents theoretical performance, based on the same formula as Calder et al. (1996), along with observed performance. The fit between observed and theoretical performance looks better at 150 ms than at 75 ms; indeed, their correlation was significant only at the 150-ms exposure duration (150 ms, $r=0.667$, $t(7)=2.37$, $p=0.050$; 75 ms, $r=0.362$, $t(7)=1.03$, $p=0.338$).

Analyses revealed that the superior discrimination performance for between-category pairs disappeared at 75 ms but was present at 150 ms, providing decisive evidence of null CP at the brief stimulus duration. In response to objections claiming that observing only one continuum is insufficient, studies using neuroimaging (Morris et al.,

1996; Morris, Friston, et al., 1998; Whalen et al., 1998; Wright et al., 2001) and developmental data (Kotsoni, DeHaan, & Johnson, 2001) have indicated that happiness and fear are the most distinct categories, and that the lack of CP between the two expressions is important. Critics might also insist that the disappearance of CP at 75 ms is a floor effect; however, discrimination for any stimulus pair at 75 ms was above chance (all $ps<0.01$), indicating that residual perception of differences in some visual properties was present. Moreover, the 75-ms duration was sufficient for accurate identification of near-endpoint facial expressions (see Figure 2). Thus, categorical perception did not occur, even with such well-preserved visual processing.

General Discussion

The present experiments revealed that categorical perception of facial expressions was attenuated and even disappeared with brief stimulus exposure durations. The abrupt change in identifying facial expressions at the category boundary was weakened at the 50-ms stimulus duration (Experiment 1), and the more accurate discrimination observed for between-category pairs than for within-category pairs was eliminated at the 75-ms stimulus duration (Experiment 2). Our findings indicate that the fast perceptual processing of emotional facial expressions is not categorical.

It has been postulated that the fast processing of facial expressions involves subcortical structures, specialized in the detection of *salient* stimuli (Adolphs, 2002). However, categorical perception refers to the processing of *ambiguous* stimuli or the transformation of an indistinctive facial configuration into a distinctive emotion category. Such fine analysis of emotional information may involve cortical structures, implying slower processing (LeDoux, 1996).

Etcoff and Magee (1992) have claimed that their observation of CP in facial expression recognition rejects the possibility that categorical processing of facial expressions is performed by higher conceptual and linguistic systems. There are data, however, demonstrating that verbal interference eliminated CP of colors and facial expressions, suggesting the essential role of verbal coding over visual coding (Roberson & Davidoff, 2000). Medin and Barsalou (1987) have argued that it is important to distinguish between *sensory perception categories* (SP categories), categories related to low-level perceptual experiences, and *generic knowledge categories* (GK categories), categories related to high-level knowledge representation. Shibui and colleagues (Shibui, Yamada, Sato, & Shigemasu, 2001) found that discrimination accuracy for within-category pairs of facial expressions was proportional to their semantic distance, and they inferred that facial expressions belong to cognitive GK categories. Our findings are consistent with the view that categorical processing of facial expressions is performed by higher cognitive systems.

Some researchers have postulated that CP of facial expressions supports the notion of basic emotions (DeGelder et al., 1996; Ekman, 1994), such that each emotion possesses an innate and specialized neuro-cognitive system. However, categorical perception is also known to occur for non-emotional visual categories such as identity (Beale & Keil, 1995; Levin & Beale, 2000), race (Levin & Angelone, 2002), and even familiar objects (Newell & Bulthoff, 2002). Categorical perception of facial expressions may thus reflect visual information processing that is common to general object recognition rather than specific to emotional content.

References

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

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57,* 217-239.

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

DeGelder, B., Teunisse, J. P., & Benson, P. J. (1997). Categorical perception of facial expressions: categories and their internal structure. *Cognition and Emotion, 11,* 1-23.

Ekman, P. (1992). Are there basic emotions? *Psychological Review, 99,* 550-553.

Ekman, P. (1994). Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. *Psychological Bulletin, 115,* 268-287.

Enns, J. T., & DiLollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345-352.

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

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: a critical over view. In S. Harnad (Ed.), *Categorical perception: the groundwork of cognition* (pp. 1-25). Cambridge: Cambridge University Press.

Information-technology Promotion Agency, Japan. (1998). Software for Facial Image Processing System for Human-like 'Kansei' Agent [Computer Software]. Retrieved from http://www.tokyo.image-lab.or.jp/aa/ipa/

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

Kotsoni, E., DeHaan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception, 30,* 1115-1125.

LeDoux, J. (1996). *The emotional brain: the mysterious underpinnings of emotional life.* New York: Simon & Schuster Inc.

Levin, D. T., & Angelone, B. L. (2002). Categorical perception of race. *Perception, 31,* 567-578.

Levin, D. T., & Beale, J. M. (2000). Categorical perception occurs in newly learned faces, other-race faces, and inverted faces. *Perception and Psychophysics, 62,* 386-401.

Medin, D. L., & Barsalou, L. W. (1987). Categorization processes and categorical perception. In S. Harnad (Ed.), *Categorical perception: the groundwork of cognition* (pp. 1-25). Cambridge: Cambridge University Press.

Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in

processing emotional facial expressions. *Brain, 121,* 47-57.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383,* 812-815.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature, 393,* 467-470.

Newell, F. N., & Bulthoff, H. H. (2002). Categorical perception of familiar objects. *Cognition, 85,* 113-143.

Roberson, D., & Davidoff, J. (2000). The categorical perception of colors and facial expressions: the effect of verbal interference. *Memory and Cognition, 28,* 977-986.

Russell, J. A. (1997). Reading emotions from and into faces: resurrecting a dimensional-contextual perspective. In J. A. Russell, & J. M. Fernandez-Dols (Ed.), *The psychology of facial expression* (pp. 295-320). Cambridge: Cambridge University Press.

Shibui, S., Yamada, H., Sato, T., & Shigemasu, K. (2001). The relationship between the categorical perception of facial expressions and semantic distances. *The Japanese Journal of Psychology, 72,* 219-226.

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18,* 411-418.

Wright, C. I., Fischer, H., Whalen, P. J., McInerney, S., Shin, L. M., & Rauch, S. L. (2001). Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. *NeuroReport, 12,* 379-383.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition, 63,* 271-313.