Foundations of Human Centered Computing: Facial Expression and Emotion*

Jeffrey F. Cohn
University of Pittsburgh
Department of Psychology, Pittsburgh, PA 15260
jeffcohn@cs.cmu.edu

Abstract

Many people believe that emotions and subjective feelings are one and the same and that a goal of human-centered computing is emotion recognition. The first belief is outdated; the second mistaken. For human-centered computing to succeed, a different way of thinking is needed.

Emotions are species-typical patterns that evolved because of their value in addressing fundamental life tasks (Ekman, 1992a). Emotions consist of multiple components that may include intentions, action tendencies, appraisals, other cognitions, central and peripheral changes in physiology, and subjective feelings. Emotions are not directly observable, but are inferred from expressive behavior, self-report, physiological indicators, and context. I focus on expressive behavior because of its coherence with other indicators and the depth of research on the facial expression of emotion in behavioral and computer science. In this paper, among the topics I include are approaches to measurement, timing or dynamics, individual differences, dyadic interaction, and inference. I propose that design and implementation of perceptual user interfaces may be better informed by considering the complexity of emotion, its various indicators, measurement, individual differences, dyadic interaction, and problems of inference.

1 Introduction

How can computers recognize human emotions? Is this even the correct question? By emotion, people often think of subjective feelings, but emotions are more than that and subjective feeling is in no sense essential. There is no *sin qua non* for emotion. Emotions are species-typical patterns consisting of multiple components that may include intentions, action tendencies, appraisals, other cognitions, neuromuscular and physiological changes, expressive behavior, and subjective feelings. None of these is necessary or sufficient. In human-human interaction, intentions and action tendencies often are more important than what an individual may be feeling. People may or may not be aware of what they're feeling, and feelings often come about some time late in the temporal unfolding of an emotion.

A goal of human-centered computing is computer systems that can unobtrusively perceive and understand human behavior in unstructured environments and respond appropriately. Much work has strived to recognize human emotions. This effort is informed by the importance of emotion to people's goals, strivings, adaptation, and quality of life (Ekman, 2003; Lazarus, 1991) at multiple levels of organization, from intra-personal to societal (Keltner & Haidt, 1999). Efforts at emotion recognition, however, are inherently flawed unless one recognizes that emotion – intentions, action tendencies, appraisals and other cognitions, physiological and neuromuscular changes, and feelings – is not an observable. Emotion can only be inferred from context, self-report, physiological indicators, and expressive behavior (see Figure 1). The focus of the current paper is on expressive behavior, in particular facial expression, and approaches to measurement, feature selection, individual differences, interpersonal regulation, and inference.

Facial expression is a useful place to begin when thinking about foundations of human computing. Facial expression has been a subject of keen study in behavioral science for more than a hundred years(Darwin, 1872/1998; Ekman & Rosenberg, 2005), and within the past 10 years considerable progress has been made in automatic analysis of facial expression from digital video input (Pantic & Patras, 2006; Pantic & Rothkrantz, 2000; Tian, Cohn, & Kanade, 2005).

Facial expression correlates moderately with self-reported emotion (Ekman & Rosenberg, 2005), pain (Prkachin, 1992), craving (Sayette et al., 2003) and emotion-related central and peripheral physiology (Davidson, Ekman, Saron, Senulis, & Friesen, 1990; Levenson, Ekman, & Friesen, 1990). Facial expression and self-reported emotion have similar underlying dimensions (e.g., positive and negative

---

* A previous version of this paper was originally published in the Proc. ACM Int'l Conf. Multimodal Interface 2006 (Copyright © ACM Press).

Cohn, J.F. (2007). Foundations of human-centered computing: Facial expression and emotion. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'07),* Hyderabad, India, xxx-xxx.

affect) (Watson & Tellegen, 1985) and serve interpersonal functions by conveying communicative intent, signaling affective information in social referencing, and contributing to the regulation of social interaction (Cohn & Elmore, 1988; Schmidt & Cohn, 2001). Cultural differences in how and when to express emotion emerge in infancy (Malatesta & Haviland, 1982; Oster et al., 1996). As a measure of trait affect and socialization, stability in facial expression emerges early in life (Cohn & Campbell, 1992). By adulthood, stability is moderately strong, comparable to what has been found for self-reported emotion (Cohn, Schmidt, Gross, & Ekman, 2002). Expressive changes in the face are a rich source of cues about intra- and interpersonal indicators and functions of emotion (Gottman, Levenson, & Woodin, 2001; Keltner & Haidt, 1999).

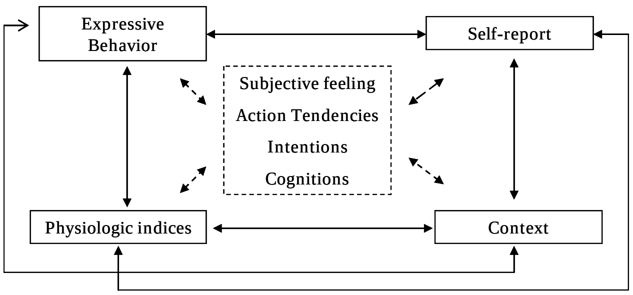


Figure 1. Components and indicators of emotion. Solid boxes represent observables, dashed boxes latent variables. Solid arrows indicate observable correlations among indicators. Large correlations among multiple indicators indicate greater coherence among indicators. Dashed arrows represent inferential paths. Paths between emotion components are omitted.

Here, I present key issues to consider in designing interfaces that approach the naturalness of face-to-face interaction. These include approaches to measurement, types of features, individual differences, dyadic interaction, and inference.

2 Approaches to Measurement

Two major approaches are sign- and message judgment (Cohn, Ambadar, & Ekman, In press). In message judgment, the observer's task is to make *inferences* about something underlying the facial behavior, such as emotion or personality. In measuring sign vehicles, the task is to *describe* the surface of behavior, such as when the face moves a certain way. As an example, upon seeing a smiling face, an observer with a judgment-based approach would make judgments such as "happy," whereas an observer with a sign-based approach would code the face as having an upward, oblique movement of the lip corners. Message judgment implicitly assumes that the face is an emotion "read out." Sign-based measurement is agnostic and leaves inference to higher-order decision making.

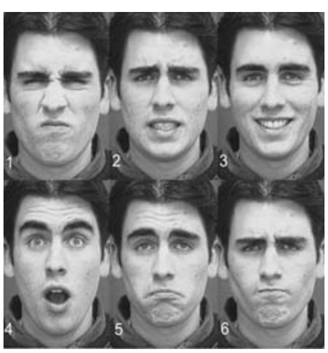


Figure 2. Emotion-specified expressions: disgust, fear, joy, surprise, sadness, and anger.

2.1 Message Judgment

Message judgment approaches define facial expressions in terms of inferred emotion. Of the various descriptors, those of Ekman have been especially influential. Ekman (Ekman, 1992b) proposed six "basic emotions." They are joy, surprise, sadness, disgust, fear, and anger. Each was hypothesized to have universally recognized and displayed signals, universal elicitors, specific patterns of physiology, rapid, unbidden onset, and brief duration, among other attributes. Since then, some additional emotions, such as embarrassment and contempt, have been added. Examples of facial expressions for the initial six basic emotions are shown in Figure 2. Most research in automatic recognition of facial expression (Pantic & Rothkrantz, 2003; Pantic, Sebe, Cohn, & Huang, 2005) and much emotion research in psychology (Keltner & Ekman, 2000) has concentrated on one or more of these six emotions. This list, however, was never intended as exhaustive of human emotion. Rather, it was proposed in terms of conformity with the criteria noted.

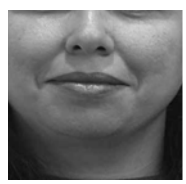


Figure 3. Example of masking smile (AU 12+14+15).

An especially important class of expressions is those that include traces of contradictory emotion expression. Masking smiles (Ekman, Friesen, & O'Sullivan, 1988), in which smiling is used to cover up or hide an underlying emotion are the best known. An example is shown in Figure 3. Signs of contempt (AU 14) and sadness (AU 15) can be seen along with the smile (AU 12). Negative emotion is believed to "leak" through the dominant positive expression.

2.2 Sign Measurement

Cohn & Ekman (Cohn & Ekman, 2005) review manual methods for labeling facial actions. Of the various methods, the Facial Action Coding System (FACS) (Ekman & Friesen, 1978; Ekman, Friesen, & Hager, 2002) is the most comprehensive, psychometrically rigorous, and widely used (Cohn, Ambadar, & Ekman, In press; Ekman & Rosenberg, 2005). Using FACS and viewing video-recorded facial behavior at frame rate and slow motion, users can manually label nearly all possible facial expressions, which are decomposed into action units (AUs). Action units, with some

Cohn, J.F. (2007). Foundations of human-centered computing: Facial expression and emotion. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'07)*, Hyderabad, India, xxx-xxx.

qualifications, are the smallest visually discriminable facial movements. By comparison, other systems are less thorough (Malatesta, Culver, Tesman, & Shephard, 1989), fail to differentiate between some anatomically distinct movements (Oster, Hegley, & Nagel, 1992), consider as separable movements that are not anatomically distinct (Oster, Hegley, & Nagel, 1992), and often assume a one-to-one mapping between facial expression and emotion (Cohn, Ambadar, & Ekman, In press; Cohn & Ekman, 2005).

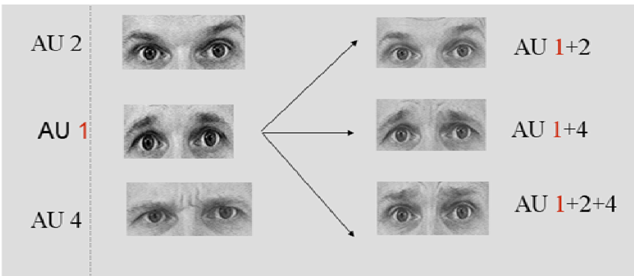


Figure 4. Examples of individual action units and action unit combinations. AU 1+2 is an additive combination. AU 1+4 and AU 1+2+4 are non-additive, comparable to co-articulation effects in speech.

The most recent version of FACS specifies 9 action units in the upper face, 18 in the lower face, 11 for head position and movement, nine for eye position and movement, and additional descriptors for miscellaneous actions, gross body movement, and supplementary codes.

Action units may occur singly or in combinations. Action unit combinations may be additive or non-additive. In additive combinations, the appearance of each action unit is independent; whereas in non-additive combinations they modify each other's appearance. Non-additive combinations are analogous to co-articulation effects in speech, in which one phoneme modifies the sound of ones with which it is contiguous. An example of an additive combination in FACS is AU 1+2, which often occurs in surprise (along with eye widening, AU 5) and in the brow-flash greeting (Eibl-Eibesfeldt, 1989). The combination of these two action units raises the inner (AU 1) and outer (AU 2) corners of the eyebrows and causes horizontal wrinkles to appear across the forehead. The appearance changes associated with AU 1+2 are the product of their joint actions.

An example of a non-additive combination is AU 1+4, which often occurs in sadness (Darwin, 1872/1998) (see Figure 4). When AU 1 occurs alone, the inner eyebrows are pulled upward. When AU 4 occurs alone, they are pulled together and downward. When AU 1 and AU 4 occur together, the downward action of AU 4 is modified. An example is shown in Figure 4. The result is that the inner eyebrows are raised and pulled together. This action typically gives an oblique shape to the brows and causes horizontal wrinkles to appear in the center of the forehead, as well as other changes in appearance. Automatic recognition of non-additive combinations presents similar complexity to that of co-articulation effects in speech. Failure to account for non-additive combination in automatic recognition exploits the correlation among AUs and can lead to inflated estimates of algorithm performance.

2.3 Reliability

The reliability of manually labeled images is a critical concern for machine learning algorithms. If ground truth is contaminated by 20-30% error, which is not uncommon, that is a significant drag on algorithm performance. For both message judgment and sign-based approaches, similar concerns arise. Using AUs as an example, at least four types of reliability (i.e., agreement between observers) are relevant to the interpretation of substantive findings. These are reliability for occurrence/non-occurrence of individual AUs, temporal precision, intensity, and aggregates. Most research in automatic facial expression analysis has focused on occurrence/non-occurrence (Pantic & Rothkrantz, 2000; Tian, Cohn, & Kanade, 2005).

Temporal precision refers to how closely observers agree on the timing of action units, such as when they begin or end. This level of reliability becomes important when examining features such as response latency and turn taking (see Section 5). Action unit intensity becomes important for questions such as whether facial expression is influenced by audience effects (Fridlund et al., 1990). Several groups have found, for instance, that people tend to smile more intensely in social contexts than when they are alone (Cohn & Schmidt, 2004; Fridlund et al., 1990). A related question is whether two measurement systems have concurrent validity for continuous measures of intensity. Our research group recently examined inter-system precision for intensity by comparing Automatic Facial Image Analysis (AFA v.4) with continuous ratings of affective intensity by human observers. Lip-corner displacement in spontaneous smiles was measured from video by AFA. Human observers made continuous ratings of affective intensity using a joy-stick like device. We found high concurrent validity between the two methods (see Figure 5 for an example) (Ibanez, Messinger, Ambadar, & Cohn, 2006; Messinger et al., 2006).

3 Dynamics

Both the configuration of facial features and the timing of facial actions are important in emotion expression and recognition. The configuration of facial actions (whether emotion-specified expressions or individual action units) in relation to emotion, communicative intent, and action tendencies has been a major research topic. Less is known about the timing of facial actions, in part because manual measurement of timing is coarse and labor intensive. We know, however, that people are highly sensitive to the timing of facial actions (Edwards, 1998) in social settings. Slower facial actions, for instance, appear more genuine (Krumhuber & Kappas, 2005), as do those that are more synchronous in their movement (Frank & Ekman, 1997). Especially subtle facial expressions become visible only when motion information is available to the perceiver (Ambadar, Schooler, & Cohn, 2005). Rapid responses to perception of facial expression can be detected within 0.5 seconds using facial EMG (Dimberg, Thunberg, &

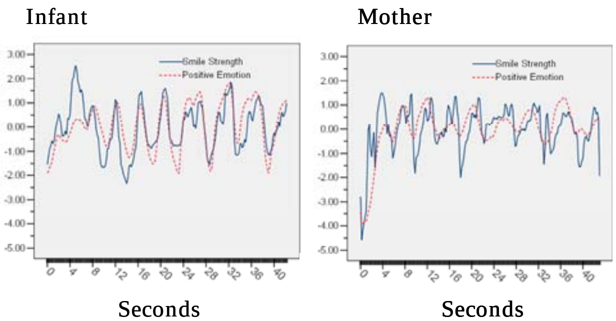
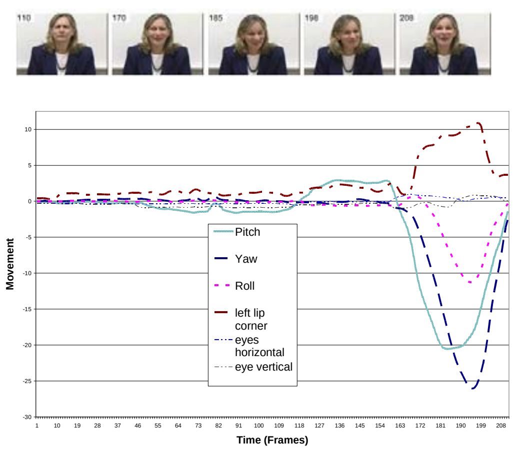


Figure 5. Time series for AFA-measured lip-corner displacement and human-observer based ratings of positive affect in a mother-infant dyad. Data series for human observers are shifted by about ½ second to adjust for human reaction time. (Ibanez, Messinger, Ambadar, & Cohn, 2006)

Grunedal, 2002). Recently, automatic facial image analysis has shown strong concurrent validity with facial EMG (Cohn & Kanade, in press), which suggests that it has similar capability.

Dynamics is especially important to inferences about communicative intention. Using automatic facial image

A



B

Figure 6. Multimodal coordination of head motion, lip-corner displacement, and gaze in smiles of embarrassment. A: Selected frames from image sequence depicting embarrassment. B: Corresponding time series. Reprinted with permission from (Cohn et al., 2004). (©2004 IEEE)

analysis to quantify the timing of facial actions, research by the CMU/Pitt group found that dynamic features discriminated between deliberate and spontaneous smiles with 89% accuracy (Cohn & Schmidt, 2004). Adding duration and amplitude to the classifier increased accuracy to 93%. Comparable findings were recently reported by (Valstar, Pantic, Ambadar, & Cohn, 2006). Using similar features, amusement, embarrassment, and polite smiles were discriminated with 83% accuracy (Kanade, Hu, & Cohn, 2005), which is comparable to that of human judges.

Recent work suggests that multimodal coordination of facial expression, head motion, and gesture is a defining feature of embarrassment (Keltner, 1995). An example is illustrated in Figure 6. Note that head pitch is closely coordinated with smile intensity. As the head pitches down, smile intensity increases, decreasing again only as the head comes back to frontal. For human-computer interaction, dynamic features are important to empirically based inferences about the meaning of otherwise similar facial actions, such as lip corner raise in smiling.

4 Individual Differences

As noted above, stable individual differences in facial expression emerge early in development and by adulthood represent 25% or more of the variation in emotion expression (Cohn, Schmidt, Gross, & Ekman, 2002; Moore, Cohn, & Campbell, 1997). Individual differences include reaction range for positive and negative affect and specific emotions and the probability of conforming to display rules. Display rules are culturally specific prescriptions for when and how to show emotion in various contexts. Sources of individual differences in emotion expression include temperament, personality, gender, socialization, and cultural background (Camras & Chen, 2006; Matsumoto & Willingham, 2006; Oster et al., 1996). In some cultures, for instance, children learn not to express anger; whereas in others, anger is considered important to self expression. Among traditional Japanese, for instance, anger is less likely to be shown outside the family than in the U.S. (Markus & Kitayama, 1991). As another example, European-American and Chinese-American couples differ in proportion of positive and negative expressions, but not in autonomic reactivity or self-

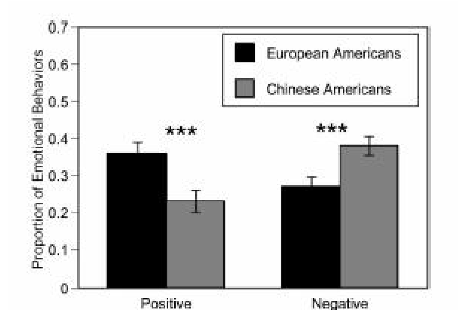


Figure 7. Cultural differences in emotional expression between European-American and Chinese-American couples. Observations were made while they discussed conflicts in their relationship. (Tsai, Levenson, & McCoy, 2006). (©2004 APA)

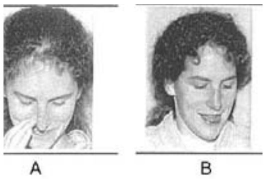


Figure 8. Some expressions appear in all or almost all cultures. Others are culture specific (A and B, respectively). Examples here are for embarrassment. From (Haidt & Keltner, 1999). (©1998 Taylor & Francis)

reported emotion, when discussing conflicts in their relationship (Tsai, Levenson, & McCoy, 2006). Emotion expressions may also be ritualized and culture specific. The tongue-bite display communicates embarrassment/shame in parts of India and south Asia but not the U.S. (see Figure 8). Within a given culture, individual differences in facial expression of all sources are strong enough to serve as a biometric (Cohn, Schmidt, Gross, & Ekman, 2002). An important implication for perceptual computing is that inferences about emotion will become more reliable when individual differences are taken into account.

5 Dyadic Interaction

Synchrony or coherence refers to the extent to which individuals are moving together in time with respect to one or more continuous output measures, such as affective valence or level of arousal. Reciprocity refers to the extent to which behavior of one individual is contingent on that of the other. Both synchrony and reciprocity have proven informative in studies of marital interaction, social development, and social psychology. Figure 9 shows an example taken from mother-infant interaction (Ibanez, Messinger, Ambadar, & Cohn, 2006; Messinger et al., 2006). Facial features and head motion were tracked automatically by the CMU/Pitt automated facial image analysis system version 4 (Cohn & Kanade, in press). The time series plot shows displacement of mother and infant lip-corners during smiles. Note that while partners tend to cycle together, there is a pattern of non-stationarity in which mother and infant take turns in leading the dyad into shared smiling, which is indicated by mother and infant time series increasing together. An important advantage of measures derived from interaction analysis is that they are largely outside of people's awareness and are difficult to manipulate intentionally.

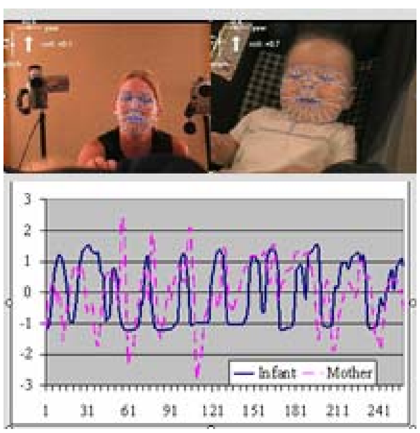


Figure 9. Example of interaction analysis. Synchrony and reciprocity of smiling between mother and infant. Source: (Ibanez, Messinger, Ambadar, & Cohn, 2006; Messinger et al., 2006).

Coordinated interpersonal timing (CIT) is the extent to which participants in a social interaction match the duration of interpersonal pauses or floor switches (Jaffe, Beebe, Feldstein, Crown, & Jasnow, 2001). Floor switches are pauses that occur between the time when one person stops speaking and another begins. Coordination of floor switches follows an inverted U-shaped function in relation to affective intensity. Mid-range values are associated with optimal affective involvement and interpersonal attraction. CIT has been studied most often with respect to vocal timing, but applies equally to facial expression and other modalities. CIT is impaired in clinical depression, with switching pauses becoming longer, more variable, and less predictable (Zlochower & Cohn, 1996).

In behavioral science, time- and frequency domain analyses have emphasized issues of quasi-periodicity in the timing of expressive behavior and bidirectional influence with respect to amplitude (Cohn & Tronick, 1988). Lag-sequential and related hidden Markov modeling have been informative with respect to the dynamics of discrete actions and individual and dyadic states (Cohn & Tronick, 1987). Recent work with dampened oscillator models considers regulation of changes in velocity and acceleration (Chow, Ram, Boker, Fujita, & Clore, 2005). Most approaches assume that time series are stationary. This assumption may not always hold for behavioral data. Boker (Boker, Xu, Rotondo, & King, 2002) identified "symmetry breaks," in which the pattern of lead-lag relationships between partners abruptly shifts. Failure to model these breaks may seriously compromise estimates of mutual influence.

6 Conclusion

Emotions are species-typical patterns that evolved because of their value in addressing fundamental life tasks (Ekman, 1992b). They are central to human experience, yet largely beyond the comprehension of contemporary computer interfaces. Human-centered computing seeks to enable computers to unobtrusively perceive, understand, and respond appropriately to human emotion, to do so implicitly, without the need for deliberate human input. To achieve this goal, it is argued that we forgo the notion of "emotion recognition" and adopt an iterative approach found in human-human interaction. In daily life, we continually make inferences about other people's emotions – their intentions, action tendencies, appraisals, other cognitions, and subjective feelings – from their expressive behavior, speech, and context. The success of human-centered computing depends in part on its ability to adopt an iterative approach to inference. Computing systems are needed that can automatically detect and dynamically model a wide range of multimodal behavior from multiple persons, assess context, develop representations of individual differences, and formulate and test tentative hypotheses though the exchange of communicative signals. Part of the challenge is that the computer becomes an active agent, in turn influencing the very process it seeks to understand. Human emotions are moving targets.

Acknowledgments

Portions of this work were supported by NIMH grant R01 MH51435 and NSF grant HSD-0527444.

References

[Ambadar et al., 2005] Ambadar, Z., Schooler, J., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics to interpreting subtle facial expressions. *Psychological Science, 16*, 403-410.

[Boker et al., 2002] Boker, S. M., Xu, M., Rotondo, J. L., & King, K. (2002). Windowed cross–correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods, 7*(1), 338-355.

[Camras & Chen, 2006] Camras, L. A., & Chen, Y. (2006). Culture, ethnicity, and children's facial expressions: A study of European American, mainland Chinese, Chinese American, and adopted Chinese girls. *6*(1), 103–114.

[Chow et al., 2005] Chow, S. M., Ram, N., Boker, S. M., Fujita, F., & Clore, G. C. (2005). Emotion as a thermostat: representing emotion regulation using a damped oscillator model. *Emotion, 5*(2), 208-225.

[Cohn et al., In press] Cohn, J. F., Ambadar, Z., & Ekman, P. (In press). Observer-based measurement of facial expression with the Facial Action Coding System. In J. A. Coan & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment. Oxford University Press Series in Affective Science*. New York, NY: Oxford University.

[Cohn & Campbell, 1992] Cohn, J. F., & Campbell, S. B. (1992). Influence of maternal depression on infant affect regulation. In D. Cicchetti & S. L. Toth (Eds.), *Developmental perspectives on depression* (pp. 103-130). Rochester, New York: University of Rochester Press.

[Cohn & Ekman, 2005] Cohn, J. F., & Ekman, P. (2005). Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. In J. A. Harrigan, R. Rosenthal & K. Scherer (Eds.), *Handbook of nonverbal behavior research methods in the affective sciences* (pp. 9-64). New York: Oxford.

[Cohn & Elmore, 1988] Cohn, J. F., & Elmore, M. (1988). Effects of contingent changes in mothers' affective expression on the organization of behavior in 3-month old infants. *Infants Behavior and Development, 11*, 493-505.

[Cohn & Kanade, In press] Cohn, J. F., & Kanade, T. Iin press). Use of automated facial image analysis for measurement of emotion expression. In J. A. Coan & J. B. Allen (Eds.), *The handbook of emotion elicitation and assessment*. New York, NY: Oxford.

[Cohn et al., 2004] Cohn, J. F., Reed, L. I., Moriyama, T., Xiao, J., Schmidt, K. L., & Ambadar, Z. (2004). *Multimodal coordination of facial action, head rotation, and eye motion.* Paper presented at the Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea.

[Cohn & Schmidt, 2004] Cohn, J. F., & Schmidt, K. L. (2004). The timing of facial motion in posed and spontaneous smiles. *International Journal of Wavelets, Multiresolution and Information Processing, 2*, 1-12.

[Cohn et al., 2002] Cohn, J. F., Schmidt, K. L., Gross, R., & Ekman, P. (2002). *Individual differences in facial expression: Stability over time, relation to self-reported emotion, and ability to inform person identification.* Paper presented at the International Conference on Multimodal User Interfaces, Pittsburgh, PA.

[Cohn & Tronick, 1987] Cohn, J. F., & Tronick, E. Z. (1987). Mother infant interaction: The sequence of dyadic states at three, six, and nine months. *Developmental Psychology, 23*, 68 77.

[Cohn & Tronick, 1988] Cohn, J. F., & Tronick, E. Z. (1988). Mother-Infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. *Developmental Psychology, 34*(3), 386-392.

[Darwin, 1872/1998] Darwin, C. (1872/1998). *The expression of the emotions in man and animals (3rd Edition)*. New York, New York: Oxford University.

[Davidson et al., 1990] Davidson, R. J., Ekman, P., Saron, C. D., Senulis, J. A., & Friesen, W. V. (1990). Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. *Journal of Personality and Social Psychology, 58*(2), 330-341.

[Dimberg et al., 2002] Dimberg, U., Thunberg, M., & Grunedal, S. (2002). Facial reactions to emotional stimuli: Automatically controlled emotional responses. *Cognition and Emotion, 16*(4), 449–471.

[Edwards, 1998] Edwards, K. (1998). The face of time: Temporal cues in facial expressions of emotion. *Psychological Science, 9*(4), 270-276.

[Eibl-Eibesfeldt, 1989] Eibl-Eibesfeldt, I. (1989). *Human ethology*. NY, NY: Aldine de Gruvier.

[Ekman, 1992a] Ekman, P. (1992a). Are there basic emotions? *Psychological Review, 99*(3), 550-553.

[Ekman, 1992b] Ekman, P. (1992b). An argument for basic emotions. *Cognition and Emotion, 6*(3/4), 169-200.

[Ekman, 2003] Ekman, P. (2003). *Emotions revealed*. New York, NY: Times.

[Ekman & Friesen, 1978] Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

[Ekman et al., 2002] Ekman, P., Friesen, W. V., & Hager, J. C. (Eds.). (2002). *Facial action coding system*: Research Nexus, Network Research Information, Salt Lake City, UT.

[Ekman et al., 1988] Ekman, P., Friesen, W. V., & O'Sullivan, M. (1988). Smiles when lying. *Journal of Personality and Social Psychology, 54*(3), 414-420.

[Ekman & Rosenberg, 2005] Ekman, P., & Rosenberg, E. (Eds.). (2005). *What the face reveals* (2nd ed.). New York: Oxford.

[Frank & Ekman, 1997] Frank, M. G., & Ekman, P. (1997). The ability to detect deceit generalizes across different types of high-stakes lies. *Journal of Personality and Social Psychology, 72*(6), 1429-1439.

[Fridlund et al., 1990] Fridlund, A. J., Sabini, J. P., Hedlund, L. E., Schaut, J. A., Shenker, J. J., & Knauer, M. J. (1990). Audience effects on solitary faces during imagery: Displaying to the people in your head. *Journal of Nonverbal Behavior, 14*(2), 113-137.

[Gottman et al., 2001] Gottman, J., Levenson, R., & Woodin, E. (2001). Facial expressions during marital conflict *Journal of Family Communication, 1*(1), 37-57.

[Haidt & Keltner, 1999] Haidt, J., & Keltner, D. (1999). Culture and facial expression: Open-ended methods find more expressions and a gradient of recognition. *Emotion and Cognition, 13*(3), 225-266.

[Ibanez et al., 2006] Ibanez, L., Messinger, D., Ambadar, Z., & Cohn, J. F. (2006). *Automated measurement of infant and mother interactive smiling.* Paper presented at the American Psychological Society.

[Jaffe et al., 2001] Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. (2001). Rhythms of dialogue in early infancy. *Monographs of the Society for Research in Child Development, 66*(2, Serial No. 264).

[Kanade et al., 2005] Kanade, T., Hu, C., & Cohn, J. F. (2005). *Facial expression analysis.* Paper presented at the IEEE International Workshop on Modeling and Analysis of Faces and Gestures, Beijing, China.

[Keltner, 1995] Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement and shame. *Journal of Personality and Social Psychology, 68*(3), 441-454.

[Keltner & Ekman, 2000] Keltner, D., & Ekman, P. (2000). Facial expression of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbooks of emotions* (second ed., pp. 236-249). New York: Guilford.

[Keltner & Haidt, 1999] Keltner, D., & Haidt, J. (1999). Social functions of emotions at multiple levels of analysis. *Cognition and Emotion, 13*(5), 505-522.

[Krumhuber & Kappas, 2005] Krumhuber, E., & Kappas, A. (2005). Moving smiles: The role of dynamic components for the perception of the genuineness of smiles. *Journal of Nonverbal Behavior, 29*, 3-24.

[Lazarus, 1991] Lazarus, R. S. (1991). *Emotion and adaptation*. NY: Oxford.

[Levenson et al., 1990] Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology, 27*(4), 363-384.

[Malatesta et al., 1989] Malatesta, C. Z., Culver, C., Tesman, J. R., & Shephard, B. (1989). The development of emotion expression during the first two years of life. *Monographs of the Society for Research in Child Development, 54*(219).

[Malatesta & Haviland, 1982] Malatesta, C. Z., & Haviland, J. M. (1982). Learning display rules: The socialization of emotion expression in infancy. *Child Development, 53*, 991-1003.

[Markus & Kitayama, 1991] Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.

[Matsumoto & Willingham, 2006] Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: spontaneous expressions of medal winners of the 2004 Athens olympic games. *Journal of Personality & Social Psychology, 91*(5), 568–581.

[Messinger et al., 2006] Messinger, D. S., Chow, S. M., Koterba, S., Hu, C., Haltigan, J. D., Wang, T., et al. (2006). *Continuously measured smile dynamics in infant-mother interaction*.Unpublished manuscript, Miami.

[Moore et al., 1997] Moore, G. A., Cohn, J. F., & Campbell, S. B. (1997). Mothers' affective behavior with infant siblings: Stability and change. *Developmental Psychology., 33*, 856-860.

[Oster et al., 1996] Oster, H., Camras, L. A., Campos, J., Campos, R., Ujiee, T., Zhao-Lan, M., et al. (1996). The patterning of facial expressions in Chinese, Japanese, and American infants in fear- and anger- eliciting situations. Poster presented at the International Conference on Infant Studies, Providence, RI.

[Oster et al., 1992] Oster, H., Hegley, D., & Nagel, L. (1992). Adult judgments and fine-grained analysis of infant facial expressions: Testing the validity of a priori coding formulas. *Developmental Psychology, 28*(6), 1115-1131.

[Pantic & Patras, 2006] Pantic, M., & Patras, I. (2006). Dynamics of facial expressions: Recognition of facial actions and their temporal segments from profile image sequences. *IEEE Tansactions on Systems, Man, and Cybernetics, Part B, 36*(2), 443-449.

[Pantic & Rothkrantz, 2000] Pantic, M., & Rothkrantz, M. (2000). Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 1424-1445.

[Pantic & Rothkrantz, 2003] Pantic, M., & Rothkrantz, M. (2003). Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE, 91*, 1371-1390.

[Pantic et al., 2005] Pantic, M., Sebe, N., Cohn, J. F., & Huang, T. S. (2005). *Affective multimodal human-computer interaction.* Paper presented at the ACM International Conference on Multimedia.

[Prkachin, 1992] Prkachin, K. M. (1992). The consistency of facial expressions of pain. *Pain, 51*, 297-306.

[Sayette et al., 2003] Sayette, M., Wertz, J., Martin, C. S., Cohn, J. F., Perrott, M., & Hobel, J. (2003). Effects of smoking opportunity on cue-elicited urge: A facial coding analysis. . *Experimental and Clinical Psychopharamacology, 11*, 218-227.

[Schmidt & Cohn, 2001] Schmidt, K. L., & Cohn, J. F. (2001). Human facial expressions as adaptations: Evolutionary perspectives in facial expression research. *Yearbook of Physical Anthropology, 116*, 8-24.

[Tian et al., 2005] Tian, Y., Cohn, J. F., & Kanade, T. (2005). Facial expression analysis. In S. Z. Li & A. K. Jain (Eds.), *Handbook of face recognition* (pp. 247-276). New York, New York: Springer.

[Tsai et al., 2006] Tsai, J. L., Levenson, R. W., & McCoy, K. (2006). Cultural and temperamental variation in emotional response. *Emotion, 6*(3), 484-497.

[Valstar et al., 2006] Valstar, M., Pantic, M., Ambadar, Z., & Cohn, J. F. (2006, November). *Spontaneous vs. posed facial behavior.* Paper presented at the ACM International Conference on Multimodal Interfaces, Banff, Canada.

[Watson & Tellegen, 1985] Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219-235.

[Zlochower & Cohn, 1996] Zlochower, A. J., & Cohn, J. F. (1996). Vocal timing in face-to-face interaction of clinically depressed and nondepressed mothers and their 4-month-old infants. *Infant Behavior and Development, 19*, 373-376.