Production of Emotional Facial Expressions in European American, Japanese, and Chinese Infants

Linda A. Camras  
DePaul University

Joseph Campos and Rosemary Campos  
University of California, Berkeley

Kazuo Miyake  
University of Hokkaido

Harriet Oster  
New York University

Tatsuo Ujiie  
Fukushima University

Lei Wang and Zhaolan Meng  
Peking University

Over 25 years ago, Ekman, Sorenson, and Friesen (1969) conducted a landmark study demonstrating that preliterate New Guinea tribespeople identified a number of emotional facial expressions similar to participants in three literate cultures: Japan, Brazil, and the United States. These findings stood in strong contrast to earlier anthropological descriptions of numerous cross-cultural differences in adult expressive behavior (e.g., Birdwhistell, 1970; LaBarre, 1947). To resolve this apparent conflict, Ekman (1972) proposed that a species-specific set of innate emotions and emotional facial expressions provide the foundation for cultural differences that are created over the course of development through socialization processes.

The evidence for universally recognized facial expressions in adults suggests that a basic repertoire of emotional expressions in infants might also be universal (Camras, Oster, Campos, Miyake, & Bradshaw, 1992; Oster, 1997). However, according to some developmental researchers (e.g., Fogel, Stevenson, & Mesinger, 1992), early socialization processes might produce cross-cultural differences even in infant expressive behavior. Furthermore, ethnic differences in infants' innate emotional reactivity might also contribute to dissimilarities observed in emotional communication (Freedman, 1974; Kagan, Kearsley, & Zelazo, 1978). These considerations highlight the importance of empirical studies focusing on infant facial expression to clarify the origins and development of cross-cultural differences in emotional expression.

Comparisons between European American and Asian infants might be particularly fruitful because of widely held beliefs about differences in emotional expressivity between adults in these cultures. Several early anthropological researchers (LaBarre, 1947; Klíneberg, 1938) described differences between Asians and Europeans or Americans in both the overall quantity of expressive behavior they produced and the situations in which they used specific facial expressions (e.g., smiling). Furthermore, differences between American and Japanese college students in their public displays of emotion have been systematically documented in a more recent study (Ekman, 1972). Although several investigations have examined the expressive...
behavior of Asian and American infants, none have focused specifically on their facial behavior. Nevertheless, these studies may shed light on the question of whether Asians and Americans differ in emotional expression during infancy as well as adulthood.

In one of the earliest cross-cultural investigations of emotional reactivity, Freedman (1974) reported that European American neonates showed more reactivity and distress during infant testing procedures than did Japanese infants who, in turn, showed more reactivity and distress than Chinese American neonates. Looking at older Japanese infants, Lewis and his colleagues (Lewis, 1989; Lewis, Ramsay, & Kawakami, 1993) have obtained similar results. For example, Lewis et al. found that 2- to 6-month-old Japanese infants showed a less intense initial response to inoculation and were more quickly soothed than European American infants. Similarly, Camras et al. (1992) reported that 5-month-old Japanese infants took slightly longer to become distressed during an arm restraint procedure than did European American infants.

However, not all studies have shown that Japanese infants are less reactive than European American infants. For example, Freedman's findings for Japanese neonates were not replicated in Kosawa's (1980, cited in Azuma, 1982) study of newborns. Camras et al. found that by 12 months of age, Japanese and European American infants no longer differed in either latency or intensity of their response to the arm restraint procedure. Furthermore, during the Strange Situation procedure, Japanese infants have been found to display more distress than American infants (Miyake, Chen, & Campos, 1985; Takahashi, 1990). Thus studies comparing Japanese to European American infants during stressful procedures have produced disparate results.

During less stressful interactions (i.e., face-to-face play), heterogeneous findings have also been obtained. For example, some studies (Caudill & Weinstein, 1969; Fogel, Toda, & Kawai, 1988) have found American infants to display more positive emotion than Japanese infants, whereas others (Caudill & Frost, 1974; Otaki, Durrett, Richards, Nyquist, & Pennebaker, 1986) have obtained no significant differences. In reviewing the literature on infant activity, Bornstein (1989) also found that results were highly divergent across studies. Such varied findings have led some researchers (e.g., Fogel, Stevenson, & Messinger, 1992) to propose that differences between European American and Japanese infants occur in response to environmental factors (e.g., pregnancy and childbirth practices, socialization processes) and are mitigated when Japanese families adopt more Westernized practices. Nonetheless, this issue is far from being resolved.

Relatively few studies have focused on the behavior of Chinese or Chinese American infants. As previously described, Freedman (1974) reported that during infant testing procedures, Chinese American neonates were less negatively reactive than both Japanese and European American neonates. Consistent with this report, Kagan et al. (1994) found that 4-month-old Chinese infants cried, fretted, and vocalized less than European American infants. However, at older ages, Chinese Americans fretted and cried more than European Americans during a maternal separation procedure and were more hesitant about approaching unfamiliar peers (Kagan et al., 1978). Regarding positive emotion, some researchers have found Chinese and Chinese American infants to smile less than European American infants during the first half year (Kisilevsky et al., 1997; Kuchner, 1989). However, Kagan and his colleagues (Kagan et al., 1978, 1994) found differences in smiling only in older infants.

Despite these heterogeneous findings, most researchers seem to agree that there are some behavioral differences between Asian and European American infants, especially at older ages. The question of whether genetic factors, prenatal nutrition, or birthing practices could contribute to these differences is controversial (Fogel et al., 1992; Freedman, 1974; Kagan et al., 1978). In contrast, investigators generally concur that postnatal cultural influences (e.g., socialization processes) might play an important role in producing cross-cultural differences in infants' behavior. Chinese and Japanese mothers have been found to differ from American mothers in ways that might clearly be expected to affect their infants' emotional and expressive development. For example, both Chinese and Japanese mothers are reported to be very indulgent toward their infants and less concerned with fostering autonomy and independence than are American mothers (Caudill & Weinstein, 1969; Chao, 1994; Chen & Miyake, 1986; Doi, 1973; Ho, 1986; Miyake, Campos, Bradshaw, & Kagan, 1986; Fogel, Stevenson, & Messinger, 1992; Stevenson, Chen, & Lee, 1992). Consistent with these cultural beliefs and goals, Asian infants are more constrained in their self-initiated motoric activities than are American infants because of maternal carrying and clothing practices (Caudill & Weinstein, 1969; Dong & Pang, 1995). Overall, during face-to-face interactions, Japanese and American mothers are equally responsive to their infants. However, Japanese mothers rely more on nonverbal cues and indirect verbal messages than European American mothers (Fogel et al., 1988) and direct their infant's attention less often toward the external environment (Bornstein et al., 1992). Similarly, Chinese American mothers have been observed to introduce less change and novelty during face-to-face interactions, to manipulate their infants' attention less than European American mothers (Kuchner, 1989), and to encourage less motoric activity (Papousek & Papousek, 1991).

Regarding the socialization of affect expression, Chinese families have traditionally fostered emotional restraint in their children (Bond & Hwang, 1986). With regard to infants, Chinese American mothers have been described as neither valuing nor encouraging the expression of positive affect as much as do European American mothers (Kagan et al., 1978; Kuchner, 1989). Descriptions of Japanese mothers' attitudes toward emotional expression have focused on negative affect more than positive affect. In the home, Japanese mothers are reported to rarely display overt expressions of anger toward their infants and to try to minimize infant crying (Doi, 1973; Lebra, 1976; Miyake et al., 1986). Japanese mothers value a quiet infant, possibly because the structure of Japanese housing limits auditory privacy. Thus, both Chinese and Japanese mothers are described as discouraging rather than encouraging emotional expression in their infants.

Direct comparisons between Chinese and Japanese mothers have not been made within a single research investigation.
Nonetheless, both Chinese and Japanese mothers are considered members of collectivist cultures, and fairly similar pictures of their attitudes about emotional expression have emerged in the literature. Thus, current descriptions of cross-cultural differences and similarities in maternal emotion socialization practices might lead one to expect that Japanese and Chinese infants would be similar to each other in emotional expression and would be significantly different from European American infants.

The present study is part of a larger investigation designed to increase our understanding of cross-cultural similarities and differences in emotional expression in infancy. American, Japanese, and Chinese infants participated in several emotion-eliciting laboratory procedures. We are assessing infants’ responses using multiple measures, including both facial and nonfacial actions and observers’ judgments of the infants’ affects (Camras et al., 1997). Herein, we present a set of initial comparisons of the infants’ facial behavior. Our measures were selected to represent previously described differences between Asians and European Americans in adult expressivity as well as in infant affect. Thus, we included quantitative measures of facial activity that could serve as indexes of overall expressivity (i.e., expressive ‘‘dynamics,’’ Thompson, 1990). We also examined smiling and latency to distress as have several previous studies (e.g., Camras et al., 1992; Lewis et al., 1993; Kagan et al., 1994).

Last, we examined additional individual facial movements to identify other expressive differences with potential emotional significance.

Among American infants, laboratory procedures have been widely and successfully used to induce and to study infants’ emotions. For example, Campos and his colleagues (Hiatt, Campos, & Emde, 1979; Stenberg & Campos, 1990) have used an arm restraint procedure, a vanish toy procedure, and a stranger approach procedure to elicit anger–frustration, surprise–interest, and fear, respectively. For the present study, we pilot tested these and other paradigms to identify those that would induce strong emotional reactions in European American, Japanese, and Chinese infants at the same age. On the basis of this pilot testing, we selected three procedures (gentle arm restraint, vanishing toy, and growing toy gorilla) that appeared appropriate for 11-month-old infants. Our study is unique in several important respects. First, we included infants from two different Asian cultures rather than comparing European American infants to either Asian Americans or to only one Asian group. Second, we observed infants in several emotion-eliciting procedures (two reported herein), including some that have previously been used only with American infants. Third, we used a fine-grained anatomically based system, BabyFACS (Oster & Rosenstein, 1996), to code the infants’ facial behavior.

Method

Participants

The participants were 11-month-old European American infants from Berkeley, California (n = 24), Japanese infants from Fukushima, Japan (n = 24), and Han Chinese infants from Beijing, People’s Republic of China (n = 24). Each cultural group was equally divided by sex. However, one male Chinese infant was dropped from the data analyses because technical difficulties precluded coding his facial behavior.

We selected 11 months as our infant age through pilot testing that evaluated the effectiveness of our procedures in eliciting affective reactions in same-age infants from all three cultures. Participants were screened free prematurely but not for parity. However, virtually all Chinese infants were only children because of China’s one-child-family policy. The infants’ families were unpaid and took part on a voluntary basis. They were recruited from urban neighborhoods surrounding the universities at which data collection took place (University of California, Berkeley; Fukushima University, Fukushima City, Japan; Peking University, Beijing, People’s Republic of China). American and Japanese infants came from primarily middle-class families in which parents were college educated. Chinese infants came from relatively advantaged families within Beijing, although their economic standards did not match those of the American and Japanese participants. Chinese parents were typically high school graduates, although some had university training.

Procedures

Each baby participated in three procedures, two of which were analyzed for this report (arm restraint and growing gorilla presentation). During all procedures, the infants were seated in a high chair (standard across cultures). Mothers sat in a chair on the infant’s right side facing perpendicular to the baby. Mothers were instructed to remain passive during stimulus presentation. All infants wore lightweight clothing, allowing them equivalent freedom of movement.

In the arm restraint procedure (designed to elicit anger–frustration), a female experimenter first spoke to the infant in a friendly manner for approximately 10 s. Then, she gently grasped the infant’s wrists and held them immobile on the high chair tray for up to 3 min. Experimenters were instructed to release the infant’s wrists before the time limit if the infant showed 7 s of continuous crying. The procedure was terminated early for 5 American, 4 Japanese, and 1 Chinese infant. In the vanishing toy procedure (designed to elicit surprise), a small barking toy dog that the infant was watching appeared to vanish instantaneously. Data for this procedure have not yet been coded and will be reported elsewhere.

The growling gorilla procedure (designed to elicit fear) involved a disembodied toy gorilla head that could be remotely activated to emit loud, unpleasant growling noises while its eyes lighted up and its lips moved. To start the procedure, the experimenter (who was hidden behind a curtain) placed the gorilla head on a table approximately 120 cm from the infant. After 10 s, the gorilla was remotely activated to emit six growls. After the growls (lasting approximately 15 s), the experimenter moved the silent gorilla head 15 cm closer to the infant. This sequence was repeated up to three times. Again, the experimenter was instructed to terminate the procedure early if the infant showed 7 s of continuous crying. This occurred for 3 American and 4 Japanese infants.

Presentation order was balanced within each sex and culture grouping. To minimize carryover effects, 3- to 5-min play periods were interposed between procedures. During these play periods, the mother first comforted the infant, if necessary. Then, the mother, research assistant, or both engaged the infant in brief play with an attractive but not unusual toy. The subsequent testing procedure was initiated only when the infant was determined to be calm and undistressed.

All procedures were videotaped with two Super-VHS Panasonic camcorders, one showing only a close-up view of the infant’s face, the other showing a wide-angle view of the infant’s full body and the experimental context.

Facial Expression Coding

The infants’ facial behavior (as recorded on the close-view videotape) was coded with Oster and Rosenstein’s (1996) BabyFACS (Baby Facial
Action Coding System), a comprehensive, anatomically based system for coding facial expressions in infants and young children. BabyFACS is a modification of Ekman and Friesen's (1978) Facial Action Coding System (FACS). As in the adult FACS, the basic coding units in BabyFACS are discrete, minimally distinguishable actions of the facial muscles (termed action units or AUs). Thus, BabyFACS describes infant facial behavior without making a priori assumptions about the emotions being expressed. Because considerable debate exists as to the emotional interpretation of many infant facial expressions (Camras, 1992; Izard et al., 1995; Oster, Hegley, & Nagel, 1992), an objective system such as BabyFACS is necessary to allow for investigation of cross-cultural similarities and differences in facial behavior, independent of assumptions about meaning.

Coding was performed with the James Long Video Coding System (VCS; Long, 1996). This system includes a hardware component that interfaces a Panasonic 1960 or 1970 videotape recorder/player with an International Business Machines compatible PC and lays a vertical interval time code (VTTC) onto a VHS-format videotape. The system also includes a software component that creates a time-coded data file in which the temporal onset and offset of coding units are automatically recorded.

To use the VCS descriptive statistics programs, we organized the BabyFACS AUs in modules. Each module (e.g., eyebrow actions) contained a set of mutually exclusive and exhaustive subcategories representing anatomically related facial actions or facial action combinations (e.g., brows raised, brows knit, brows raised and contracted). No AUs were omitted, and AUs that could physically co-occur (such as brow raising, cheek raising, smiling) were placed in different modules. Therefore, all possible facial actions and combinations of actions were comprehensively and unambiguously coded (see Oster & Rosenstein, 1996, BabyFACS Manual, for further details on this Modular Coding System). However, because the full Modular Coding System (like BabyFACS itself) includes a large number of minimally distinct facial actions, several analyses presented herein are based on a predetermined clustering of the facial actions codes. To be specific, we grouped some action units together to increase data manageability while retaining distinctions that previous researchers have found relevant to emotional expression (e.g., Ekman & Friesen, 1975; Izard, Dougherty, & Hembree, 1983; Oster & Ekman, 1978; Oster, Hegley, & Nagel, 1992). For example, nose wrinkle (AU 9) and upper lip raise (AU 10) were grouped together because they often co-occurred in the data, and both have been related to disgust. This set of facial movement variables (see Appendix) could still generate more than 50,000 combinations of facial muscle action codes.

Because facial coding is extremely labor intensive (approximately 1 hour of coding time per minute of active facial behavior), four predesignated coding intervals were identified for each infant. For each procedure, a baseline episode and a stimulus episode were demarcated, representing the infant's facial behavior immediately before and immediately after the onset of the high-impact emotion stimulus. For the arm restraint procedure, the baseline coding episode was the 10-s interval of friendly infant–experimenter interaction that occurred immediately before the infant's wrists were grasped. The stimulus coding episode was the first 20 s of the restraint procedure. For the growling gorilla procedure, the baseline episode was the 10-s interval during which the toy gorilla head was first silently placed before the infant. The 20-s stimulus coding episode began when the gorilla head began to emit its first set of growls. Baseline episodes were selected to be nonaversive versions of the emotion-eliciting procedures involving the same stimulus objects (female experimenter, toy gorilla head). Baseline and stimulus episodes differed in duration because pilot testing had shown that infants typically did not attend to the silent gorilla head for more than 10 s.

Coding was performed by two research assistants who were trained and certified to use Ekman and Friesen's (1978) FACS and who were also trained in the BabyFACS coding procedure. Reliability was assessed for the specific materials coded in this study by having both coders score approximately 25% of the infants in each culture. Kappa statistics for the coding modules of the modular system exceeded .62 with the exception of the eyelid action module, which was .57 for the Japanese babies.

**Results**

**Expressive Dynamics**

Chinese infants were consistently found to be less expressive than the European American and Japanese infants, who typically did not differ from each other. This pattern of results was obtained in a 3 (culture) X 2 (infant sex) X 2 (procedure) X 2 (episode: baseline vs. stimulus) multivariate analysis of variance (MANOVA), using three dependent measures reflecting the quantity of facial behavior produced by the infants. These measures were (a) facial movement time, that is, the percentage of the coding interval during which the infant produced facial movement (i.e., during which one or more facial actions were coded), (b) expression variability (i.e., the number of different facial configurations produced during the coding interval), and (c) expression lability (i.e., the number of times the infants' facial configuration changed during the coding interval). A facial configuration was a set of simultaneously produced facial actions; two configurations might differ from each other by either one or several action units (e.g., AU 1 + 2 + 12 vs. AU 1 + 2 + 15 ±17). To achieve a better approximation to the normal distribution, the facial movement scores were transformed from raw percentages to log transformation scores before being entered into the MANOVA (Winer, 1971). The expression variability and lability scores were transformed into rate measures (number of changes per second) to equate for differences in the duration of episodes.

The analysis yielded significant multivariate effects for culture, F(6, 126) = 12.38, p < .001, procedure, F(3, 63) = 9.18, p < .001, episode, F(3, 63) = 14.50, p < .001, Culture X Episode, F(6, 126) = 3.74, p < .002, and Procedure X Episode, F(3, 63) = 12.38, p < .001. The univariate F tests for each dependent variable were inspected (see Table 1), and post hoc Tukey tests (p < .05) were performed to identify significant differences in infants' use of these facial movements (see Tables 2 and 3 for means and summary of significant results).

**Table 1**

**Significant Univariate Effects for Expressive Dynamics Variables**

<table>
<thead>
<tr>
<th>Source</th>
<th>dF</th>
<th>Facial movement time</th>
<th>Expression variability</th>
<th>Expression lability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>6, 126</td>
<td>18.84***</td>
<td>23.55***</td>
<td>25.95***</td>
</tr>
<tr>
<td>Procedure</td>
<td>3, 63</td>
<td>ns</td>
<td>21.30***</td>
<td>24.95***</td>
</tr>
<tr>
<td>Culture X Episode</td>
<td>6, 126</td>
<td>8.48***</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Procedure X Episode</td>
<td>3, 63</td>
<td>ns</td>
<td>27.44</td>
<td>34.40</td>
</tr>
</tbody>
</table>

*p < .05. ***p < .001.
Results showed that American and Japanese infants produced facial movement during a greater proportion of the coding interval than did Chinese infants (71%, 82%, and 53% of their coding intervals, respectively). Post hoc comparisons additionally showed that Japanese infants produced even more facial movement than American infants during the baseline episodes, although they did not differ from the American infants during the stimulus episodes. Overall, facial movement time was longer during stimulus episodes than during baseline episodes.

Both expression variability and expressive lability were greater for American and Japanese infants than for Chinese infants. Overall, infant expressive behavior was more variable and labile during the arm restraint procedure than during the growling gorilla procedure. However, post hoc comparisons showed this difference to be significant for only the stimulus episodes of these procedures. During the arm restraint procedure, expression variability and lability was greater during the stimulus episode than during the baseline episode. However, during the growling gorilla procedure, expression variability and lability were greater during the baseline episode than during the stimulus episode. Given that facial movement time did not differ across procedures, the overall pattern of results suggests that facial expressions produced during the growling gorilla stimulus episode changed less and were held longer than those produced during the arm restraint stimulus episode.

**Latency to Cry Face**

Data analysis showed that American infants cried significantly sooner than Chinese infants, with Japanese infants falling between the other two groups (latency = 50.9 s, 56.2 s, and 77.0 s for American, Japanese, and Chinese babies, respectively), $F(2, 65) = 4.02, p < .03$, for culture (Tukey tests, $p < .05$). In addition, girls cried more quickly than boys (52.5 s vs. 65.1 s), $F(1, 65) = 4.13, p < .05$, for gender. Whereas only 2 babies showed very brief cry faces during a baseline episode, 40% of the babies ($n = 29$) showed cry faces during the arm restraint procedure, and over 30% ($n = 22$) showed cry faces during the growl procedure. Infants who did not cry were assigned a latency score that was equivalent to the maximum possible duration of the episode.

**Smiles**

Smiles were grouped into two categories: (a) Duchenne smiles, that is, smile mouths accompanied by cheek raise (AU 6; produced by contraction of the outer portion of orbicularis oculi, the muscle encircling the eye) and (b) non-Duchenne smiles, that is, smile mouths not accompanied by cheek raising. Duchenne smiles are hypothesized by Ekman, Davidson, and Friesen (1990) to reflect genuine positive emotion, whereas non-Duchenne smiles are hypothesized to be nonenjoyment (sometimes purely social) smiles.

A 3 (culture) X 2 (gender) X 2 (procedure) X 2 (episode) MANOVA was performed with scores equaling log values of the percentage of coding time during which each smile type was produced. The analysis yielded significant effects for culture, $F(4, 128) = 5.85, p < .001$, gender, $F(2, 64) = 4.19, p < .05$, procedure, $F(2, 64) = 8.31, p < .001$, episode, $F(2, 64) = 8.08, p < .001$, Culture X Episode, $F(4, 128) = 2.84, p < .05$, Gender X Episode, $F(2, 64) = 3.35, p < .05$. Follow-up univariate $F$ tests showed that the main effects and the Culture X Episode interaction held only for the Duchenne smiles, $F(2, 65) = 12.13, p < .001$, for culture, $F(1, 65) = 8.31, p < .005$, for gender, $F(1, 65) = 16.01$ and 15.89, $p < .001$, for procedure and episode, respectively, $F(2, 65) = 5.77, p < .005$, for Culture X Episode. The Gender X Episode interaction held only for non-Duchenne smiles, $F(1, 65) = 6.41, p < .02$.

American and Japanese infants produced significantly more Duchenne smiles than Chinese infants. However, these cultural differences reached significance only for the baseline episodes (Tukey test, $p < .05$) during which most of the smiles occurred. Infants produced more Duchenne smiles during the arm restraint procedure than during the gorilla procedure. In addition, boys produced significantly more non-Duchenne smiles than did girls during the baseline episodes.

**Individual Facial Action Variables**

Initial data inspection showed that 21 of the 22 facial action variables (as presented in the Appendix) were produced by at least one infant. For this analysis, orbicularis oculi contractions involving only the inner muscle strands (i.e., AU 7) were excluded because they do not meet the criteria for Duchenne smiles.

---

**Table 2**

*Scores for Expressive Dynamics Variables*

| Variable | United States | | Japan | | China |
|----------|---------------|---|---|---|---|---|---|---|
|          | Arm base | Arm stim | Gor base | Gor stim | Arm base | Arm stim | Gor base | Gor stim |
| Facial movement time | 0.61 | 0.81 | 0.59 | 0.83 | 0.78 | 0.85 | 0.79 | 0.83 |
| Variability | 0.82 | 1.36 | 1.10 | 0.60 | 1.01 | 1.22 | 0.86 | 0.67 |
| Lability | 1.12 | 1.75 | 1.52 | 0.86 | 1.24 | 1.57 | 1.03 | 0.85 |

*Note.* Arm = arm restraint procedure; Base = baseline episode; Gor = growling gorilla procedure; Stim = stimulus episode.

*Percentage of coding interval.  Number of changes per second.*
least one baby from each culture. One variable (AU 15 ± 17, i.e., lip corners depressed with or without chin raised) was recorded for only the Japanese infants but has previously been observed in American infants (Oster & Rosenstein, 1996). The data thus indicate substantial commonality in the repertoires of the American, Japanese, and Chinese infants.

To further examine the infants' facial behavior, those facial variables (as designated in the Appendix) that occurred most often in the data were identified using two criteria drawn from Shaver, Schwartz, Kirson, and O'Connor's (1987) studies of prototypic emotion features. The criteria were (a) the facial variable was produced by at least 20% of the infants from any culture during one (or more) of the coding intervals (i.e., arm restraint baseline episode, arm restraint stimulus episode, gorilla baseline episode, or gorilla stimulus episode) and (b) the facial variable's mean duration was at least 20% of the coding interval. Each variable could be produced alone or in combination with other facial movements. Using these criteria, we identified seven commonly produced facial variables: (a) smile mouth (irrespective of accompanying cheek movement), (b) mild-to-intense cry mouth, (c) brow raised, (d) brow lowered, (e) brow raised and contracted, (f) cheek raised and/or lower eyelid raised (by contraction of orbicularis oculi), and (g) nasolabial furrow deepening with or without nose wrinkle and/or upper lip raised. Regarding the interpretation of these actions, smile mouths and cry mouths are associated with positive and negative emotion, respectively. In infants, lowered brows and nasolabial furrow deepening are found exclusively in expressions of negative affect (Oster & Rosenstein, 1996; Oster, 1997). Although the interpretation of individual facial movements in adults is ambiguous and context dependent (Ekman, 1979; see Discussion section below), lowered brows can occur as a component of adult prototypic anger or disgust expressions, whereas nasolabial furrow deepening can be a component of sadness (Ekman & Friesen, 1978). Raised brows can be a component of adult prototypic surprise expressions but may also occur in other contexts in both adults (Ekman, 1979) and infants (e.g., object mouthing; Camras, Lambrecht, & Michel, 1996). Raised and contracted brows can be a component of prototypic adult fear expressions. They have not been frequently observed in babies (Camras, 1991).

To compare infants' use of these facial actions, a 3 (culture) × 2 (infant sex) × 2 (procedure) × 2 (episode) MANOVA was performed using scores for the seven facial actions as dependent variables. These scores were log values of the percentage of coding time during which each movement variable was produced. The analysis yielded significant overall effects for culture, $F(14, 118) = 7.84, p < .001$, procedure, $F(7, 59) = 3.89, p < .002$, episode, $F(7, 59) = 13.13, p < .001$, and also a significant Culture × Episode interaction, $F(14, 118) = 4.64, p < .001$. The univariate $F$ tests for each dependent variable were inspected (see Table 4), and post hoc Tukey tests ($p < .05$) were performed to identify significant differences in infants' use of these facial movements (see Tables 5 and 6 for means and summary of significant results).

Results showed that American infants produced more smile mouths than did Japanese infants, who themselves produced more smile mouths than the Chinese infants. These cultural differences reached significance in the baseline episodes during which most of the smiling occurred. In addition, smile mouths occurred more during the arm restraint procedure than during the growling gorilla procedure.

Regarding the infant negative affect variables, American infants produced more mild-to-intense cry mouths than Chinese infants and more lowered brows than both Japanese and Chinese infants. However, these differences reached significance in only the stimulus episodes (during which most negative affect occurred). Japanese infants also produced more cry mouths and nasolabial furrow deepening than Chinese infants during the stimulus episodes. There were no cultural differences for brows raised and contracted. All four facial variables were produced more during stimulus than during baseline episodes. There were no significant differences across procedures.2

Table 3
Direction of Significant Effects for Expressive Dynamics Variables

<table>
<thead>
<tr>
<th>Source</th>
<th>Facial movement time</th>
<th>Variability</th>
<th>Lability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>A = J &gt; C</td>
<td>A = J &gt; C</td>
<td>A = J &gt; C</td>
</tr>
<tr>
<td>Procedure</td>
<td>Stim &gt; Base</td>
<td>Arm &gt; Gor</td>
<td>Arm &gt; Gor</td>
</tr>
<tr>
<td>Episode</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture × Episode</td>
<td>Base: J &gt; A &gt; C</td>
<td>Arm: Stim &gt; Base</td>
<td>Gor: Base &gt; Stim</td>
</tr>
<tr>
<td>Procedure × Episode</td>
<td></td>
<td>Gor: Stim &gt; Base</td>
<td>Stim: Arm &gt; Gor</td>
</tr>
</tbody>
</table>

Note. A = American; J = Japanese; C = Chinese; Arm = arm restraint procedure; Gor = growling gorilla procedure; Stim = Stimulus episode; Base = Baseline episode.

2 Because significant procedure effects might be expected for facial movements that can be components of anger—frustration versus fear in adults (e.g., brow lower vs. brow raised and contracted), we conducted an additional MANOVA to explore the possibility that such effects were
Japanese infants raised their brows more than American infants during both baseline and stimulus episodes. During stimulus episodes, Chinese infants also used this movement significantly more than American infants. Brow raises were produced more often during the growling gorilla procedure than during the arm restraint procedure.

American and Japanese infants showed raised cheeks and/or raised lower eyelids significantly more than did Chinese infants. Results also showed that this facial action code occurred more during the arm restraint procedure than during the growling gorilla procedure and more during stimulus episodes than baseline episodes.

Discussion

This study found cross-cultural differences both in the infants’ expressive dynamics and in the durations of specific facial movements. Using detailed coding of facial behavior, we confirmed earlier reports that Asians and European Americans differ in emotional expression and that such differences originate in infancy. However, our findings are also significant because they contradict some prevalent but largely unsubstantiated beliefs about the nature of these differences. First, they demonstrate that two Asian cultural groups may differ significantly from each other in emotional expressivity. Second, they show that European Americans may differ significantly from some Asian groups (e.g., Chinese infants) but not from others (e.g., Japanese infants).

These data have important implications for cross-cultural research on the socialization of emotion. Although we cannot yet distinguish among competing views regarding the hereditary versus experiential origins of the differences we obtained, our study suggests several potentially profitable lines of further investigation. For example, future studies might compare the facial behavior of Asian and Asian American infants whose families have adopted European American values and socialization practices. In addition, future research might focus on identifying differences between Chinese and Japanese mothers that might produce the dissimilarities we found in their infants’ expressive behavior. That is, significant differences should be looked for in the socialization practices of cultures that are currently considered to share a collectivist orientation. Such studies are critical to our eventual understanding of the relative contribution of innate and sociocultural factors to the development of emotional reactivity and emotional expression.

Expressive Dynamics

Regarding expressive dynamics, our results showed specifically that Chinese infants produced less facial movement than did American and Japanese infants and were less labile and variable in the expressive behavior that they did produce. To some extent, the Chinese infants thus conformed to early descriptions of Asians’ lesser overall expressivity, although even these infants did produce considerable facial activity (i.e., during 53% of the coding interval). Of particular interest, only minimal differences were found between the American and Japanese infants for the expressive dynamics variables. Although our expressive dynamics measures did not distinguish between positive and negative reactivity, we were able to examine these separately in our analyses of crying and smiling.

Smiling

To some extent, our results for smiling replicate the findings of several past studies comparing European American infants to both Chinese (Kagan et al., 1978) and Japanese (Fogel et al., 1988) infants. As in these studies, we found that European Americans smiled more than Asian infants. However, when Duchenne smiles (i.e., smile mouths accompanied by cheeks raised) were examined, American and Japanese infants did not significantly differ, and both groups produced significantly more

---

Table 4

<table>
<thead>
<tr>
<th>Facial action variables</th>
<th>Smiling</th>
<th>Mild-to-intense cry mouth</th>
<th>Brow raise &amp; lower contract</th>
<th>Cheek, lid raise</th>
<th>Nasolabial deepen ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>ddfs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture</td>
<td>14, 118</td>
<td>14.47***</td>
<td>4.69*</td>
<td>5.55**</td>
<td>11.26***</td>
</tr>
<tr>
<td>Procedure</td>
<td>7, 59</td>
<td>9.37**</td>
<td>ns</td>
<td>6.16*</td>
<td>ns</td>
</tr>
<tr>
<td>Episode</td>
<td>7, 59</td>
<td>15.98***</td>
<td>52.86***</td>
<td>ns</td>
<td>30.73***</td>
</tr>
<tr>
<td>Culture × Episode</td>
<td>14, 118</td>
<td>5.76**</td>
<td>4.99*</td>
<td>5.97**</td>
<td>12.34***</td>
</tr>
</tbody>
</table>

Note. Cheek, lid raise = cheek raise, lower lid raise; Nasolabial deepen ± = nasolabial furrow deepen ± nose wrinkle, upper lip raised.

* p < .05. ** p < .01. *** p < .001.

---

not found because of emotional "carry-over" from the first-presented to the second-presented procedure. For this analysis, we examined only the first-presented procedure for each participant. The set of significant results involving procedure replicated the pattern obtained for the larger analysis save that an additional Culture × Procedure × Episode effect was obtained for smiling.
Table 5

Duration of Often-Produced Individual Facial Action Variables: Mean Percentages of Coding Interval

<table>
<thead>
<tr>
<th>Facial action</th>
<th>United States</th>
<th>Japan</th>
<th>China</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Arm base</td>
<td>Arm stim</td>
<td>Gor base</td>
</tr>
<tr>
<td>Smile mouth</td>
<td>33.7</td>
<td>14.3</td>
<td>19.8</td>
</tr>
<tr>
<td>Mild-to-intense cry mouth</td>
<td>0.0</td>
<td>21.6</td>
<td>0.1</td>
</tr>
<tr>
<td>Brow raise</td>
<td>13.7</td>
<td>7.8</td>
<td>21.5</td>
</tr>
<tr>
<td>Brow lower</td>
<td>0.3</td>
<td>25.9</td>
<td>2.6</td>
</tr>
<tr>
<td>Brow raise &amp; contract</td>
<td>0.4</td>
<td>4.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Cheek raise, lower lid raise</td>
<td>31.3</td>
<td>44.6</td>
<td>21.3</td>
</tr>
<tr>
<td>Nasolabial deepen ± nose wrinkle, upper lip raise</td>
<td>0.0</td>
<td>11.0</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Note. Arm = arm restraint procedure; Base = baseline episode; Gor = growling gorilla procedure; Stim = stimulus episode.

Duchenne smiles than did the Chinese infants. Although evidence supporting the distinction between Duchenne and non-Duchenne smiles is robust for adults (Ekman, 1992; Ekman et al., 1990), there is some controversy about whether only Duchenne smiles indicate genuine positive emotion in infancy (Dickson, Walker, & Fogel, 1997; Fox & Davidson, 1988; Mesinger, Fogel, & Dickson, 1997). Nonetheless, Duchenne smiles are generally acknowledged to have special social significance. Thus, our results for smiling demonstrate that Asian cultural groups may differ from each other significantly in the expression of positive emotion and underscore the value of empirical comparisons using more than one Asian culture.

Crying

The results for our crying measures again showed that differences between European American and Chinese infants were greater than differences between European American and Japanese infants. American infants cried significantly sooner and produced significantly more mild-to-intense cry mouths than the Chinese infants. Japanese infants fell between the two other groups on both measures and differed significantly from only the Chinese infants in the duration of cry mouth movements. No significant differences between European American and Japanese infants were obtained for either measure.

These findings are consistent with Kagan et al.'s (1994) finding that 4-month-old European American babies cry more than Chinese babies during infant testing procedures. They are also partly consistent with Freedman's (1974) investigation of neonates, in which he reported that Japanese infants were more irritable than Chinese infants but less irritable than European American infants. However, in our study, Japanese infants differed significantly from only the Chinese infants and not from American infants. Our findings also differ from Lewis et al.'s (1993) investigation of 2- to 6-month-old Japanese versus European American infants. As described earlier, Lewis et al. found that Japanese infants scored lower than European American infants on several (though not all) of his negative reactivity measures. Methodologically, our study differed from Lewis's investigation in several respects, including infant age, emotion-

Table 6

Direction of Significant Effects for Often-Produced Individual Facial Action Variables

<table>
<thead>
<tr>
<th>Source</th>
<th>Smile mouth</th>
<th>Mild-to-intense cry mouth</th>
<th>Brow raise</th>
<th>Brow lower</th>
<th>Brow raise &amp; contract</th>
<th>Cheek, lid raise</th>
<th>Nasolabial deepen ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture</td>
<td>A &gt; J &gt; C</td>
<td>A &gt; C</td>
<td>J &gt; A</td>
<td>A &gt; J = C</td>
<td>A = J &gt; C</td>
<td>J &gt; C</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>Arm &gt; Gor</td>
<td>Gor &gt; Arm</td>
<td>Stim &gt; Base</td>
<td>Stim &gt; Base</td>
<td>Stim &gt; Base</td>
<td>Stim &gt; Base</td>
<td></td>
</tr>
</tbody>
</table>

Note. Cheek, lid raise = cheek raise, lower lid raise; Nasolabial deepen ± = nasolabial furrow deepen ± nose wrinkle, upper lip raised; A = American; J = Japanese; C = Chinese; Arm = arm restraint procedure; Gor = growling gorilla procedure; Base = baseline episode; Stim = stimulus episode.
eliciting procedures, specific measurement of distress, and, possibly, distress intensity. Nevertheless, divergence between the two studies suggests that differences between European American and Japanese infants in negative expressivity are not robust.

Cultural Differences in Other Facial Actions

Cultural differences were also found in the infants’ production of several other facial movement variables (see Table 6). For example, during the stimulus episodes, American infants produced more lowered brows than either Japanese or Chinese infants, whereas Japanese infants produced more midface actions (i.e., nasolabial furrow deepening with or without nose wrinkle and/or upper lip raised) than did the Chinese infants. In infants, lowered brows (as opposed to knit brows, see Appendix) and midface actions can be components of cry faces and distaste expressions (Oster & Rosenstein, 1996; Rosenstein & Oster, 1988). However, because individual facial actions are not invariably related to specific, discrete emotions in adults, we do not assume that lowered brows or nasolabial furrow deepening reflect specific emotions (e.g., anger or sadness) in infants. Nevertheless, because they are components of negative, rather than positive, affect expressions in infants, cultural differences in these movements may reflect differences in the intensity or the nature of emotional responsiveness to aversive stimulation.

Despite the preceding caveats, we have considered the possibility that lowered brows was selectively associated with anger—frustration in our study. Thus, the cultural differences we found for lowered brows might reflect differences in the experience of a discrete negative emotion. However, as yet, no evidence for this proposal has been produced in our investigation. Although they produced more lowered brows than their Asian counterparts, American infants were not judged to be more angry (or frustrated) by judges who rated their emotions on the basis of nonfacial cues (as described in Camras et al., 1997). Furthermore, lowered brows was not produced more often in the arm restraint procedure (intended to elicit anger—frustration) than in the growling gorilla procedure (intended to elicit fear).

An alternative possibility is that lowered brows may occur with a variety of different negative emotions or during states of relatively undifferentiated negative affect (Camras, 1991; Oster, Hegley, & Nagel, 1992). In the present study, American infants cried more and were rated as being more distressed than the Asian infants. Possibly, the greater use of lowered brows by American infants was related to their greater degree of distress and crying (i.e., undifferentiated negative affect) rather than to differences in their specific anger reactions. In addition, according to the dynamical systems view of emotional expression, production of specific facial components (e.g., lowered brows) may be related to nonemotional contextual factors, such as the intensity and the type of sensory input to the infant, the action demands of the situation, and the infant’s initial motoric state (see Camras, 1991, 1992; Camras, Lambrecht, & Michel, 1996, for further discussion). In summary, the results of this study demonstrated cross-cultural differences in infants’ tendency to produce some negative facial affect components. Further investigations must continue to investigate the question of whether these differences are related to discrete negative emotions or to nonemotion factors and, also, whether they generalize to other emotion situations and whether they significantly influence observers’ judgments and responses to the infant.

Cross-cultural differences were also found for infants’ production of raised brows, with Japanese infants using this movement significantly more than American infants. In expression recognition studies, raised brows is identified as a component of both surprise (Ekman & Friesen, 1975; Izard et al., 1983) and interest (Izard et al.) facial expressions—although, again, they are not invariably associated with these emotions (Ekman, 1979). In the present study, perhaps Japanese infants produced more brow raises because they were indeed more surprised or interested than American infants (especially in the growling gorilla procedure, during which brow raises most often occurred). Providing weak support for this hypothesis, Japanese infants were rated (nonsignificantly) higher in surprise than American infants by judges observing their nonfacial behavior during the growling gorilla procedure (see Camras et al., 1997). However, the ratings for surprise were low overall and did not differ across cultures for the arm restraint procedure. In addition, interest ratings were virtually identical for the American, Japanese, and Chinese babies.

As an alternative possibility, infants from all three cultures may have been equally interested and surprised by our procedures but may have differed in their production of raised brows for other reasons. In previous studies of expression production, raised brows has been found to occur with visual search (Burton Jones & Konner, 1971), upward head and gaze movements (Michel, Camras, & Sullivan, 1992), and appetitive actions involving mouth opening (i.e., bringing an object to the mouth; Camras, Lambrecht, & Michel, 1996). Such motor actions might clearly be produced by surprised or interested babies but would not be uniquely associated with these emotions. In the present study, perhaps Japanese infants produced more of these head, gaze, and mouth actions than did American babies, thus accounting for their significantly greater production of raised brows. Such findings would again be consistent with a dynamical systems view of emotional expression (Camras, 1991, 1992; Camras, Lambrecht, & Michel, 1996; Michel et al., 1992). Further research involving fine-grained coding of both facial and nonfacial behavior is required to assess this possibility.

A last cross-cultural difference involved American and Japanese infants’ greater production of orbicularis oculi movements (i.e., AU 6,7; cheek raise and/or lower eyelid raise) relative to Chinese infants. As noted above, cheek raises (AU 6, produced by contraction of the outer portion of orbicularis oculi) distinguish Duchenne from non-Duchenne smiles. However, cheek raising is also a component of cry faces. Lower eyelid raising (AU 7, produced by contraction of the inner portion of orbicularis oculi) occurs almost exclusively in negative affect expressions. Thus, our finding that both of these movements (alone or in combination) were shown more by American and Japanese infants can probably be explained by their greater production of both smiles and cry-face expressions relative to the Chinese infants.

Differences Across Procedures and Episodes

Beyond these cultural differences in facial behavior, some differences across procedures and episodes were also observed
expressions of distress. Lewis's findings indicate that measures than did American infants in response to inoculation, although et al.'s (1993) cross-cultural study of American and Japanese infants. This possibility was suggested by the results of Lewis's findings that Japanese and Chinese infants may differ in their expressive behavior. Such studies might also include a measure of parents' educational levels relative to the parents of the Chinese infants. Future studies must determine whether demographically diverse samples of American, Japanese, and Chinese infants also differ in their expressive behavior. Such studies might proceed to systematically examine its socialization correlates, thus turning a previously presumed source of "error variance" into an important source of information about the relationship between socialization practices and the development of emotional expression.

Conclusions

In summary, the present investigation contributes to our understanding of cross-cultural differences in emotional expression. Regarding overall expressivity, our results indicated that differences between European American and Chinese infants are more robust than differences between European American and Japanese infants and that Japanese and Chinese infants may
differ significantly from each other. Future studies must continue to explore the influence of innate and sociocultural factors on infant expressiveness along several lines to identify the source of these differences.

Our study also demonstrated cultural differences in infants’ use of some specific facial action variables. However, our data did not suggest that these facial movements were associated with specific, discrete negative emotions in infants. Future research must explore the possibility that infant emotional expression does not always involve the production of prototypic adult-like configurations for discrete emotions. Instead, infants often may produce global expressions of positive and negative affect that sometimes incorporate additional movements, depending on situational or contextual factors rather than on the particular discrete emotion.

References


Blurton Jones, N., & Konner, M. (1971). An experiment on eyebrow-like configurations for discrete emotions. Instead, infants often with specific, discrete negative emotions in infants. Future re-


(Appendix follows)
## Appendix

### Facial Action Variables Used in Data Analyses

<table>
<thead>
<tr>
<th>Facial action module</th>
<th>Verbal label for facial action variable</th>
<th>BabyFACS action units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smile mouth</td>
<td>Smile</td>
<td>12</td>
</tr>
<tr>
<td>Cry mouth</td>
<td>Mild-to-intense cry mouth</td>
<td>20a/b+25/26/27 or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20c/d/e+26d/e/27</td>
</tr>
<tr>
<td></td>
<td>Horizontal cry mouth</td>
<td>20c/d/e+25/26a/b</td>
</tr>
<tr>
<td></td>
<td>Horizontal lip stretch</td>
<td>20</td>
</tr>
<tr>
<td>Modulated negative mouth</td>
<td>Horseshoe mouth</td>
<td>15c/d/e+17a/b/c</td>
</tr>
<tr>
<td></td>
<td>Pout/lower lip out/kidney-shaped mouth</td>
<td>17±8,9,10,11,15,20,24</td>
</tr>
<tr>
<td>Eyebrow</td>
<td>Brow oblique</td>
<td>1±3,4</td>
</tr>
<tr>
<td></td>
<td>Brow knit</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Brow lowered</td>
<td>4±3</td>
</tr>
<tr>
<td></td>
<td>Brow raised</td>
<td>1±2 or 2</td>
</tr>
<tr>
<td></td>
<td>Brow raised and contracted</td>
<td>1+2±3, 4 or 2±3, 4</td>
</tr>
<tr>
<td>Eyelid</td>
<td>Normal, drooped</td>
<td>None or 41</td>
</tr>
<tr>
<td></td>
<td>Widened</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Closed</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Squint</td>
<td>44</td>
</tr>
<tr>
<td>Eye orbital</td>
<td>Cheek raise, lower lid raise</td>
<td>6,7</td>
</tr>
<tr>
<td>Midface</td>
<td>Nose wrinkle, upper lip raised</td>
<td>9,10</td>
</tr>
<tr>
<td></td>
<td>Nasolabial furrow deepen ± nose wrinkle, upper lip raised</td>
<td>11±9, 10</td>
</tr>
<tr>
<td>Mouth orbital</td>
<td>Lips tight/press</td>
<td>23/24</td>
</tr>
<tr>
<td></td>
<td>Other orbital actions</td>
<td>18/22/28/32/33</td>
</tr>
<tr>
<td>Lips/chin</td>
<td>Lip corner depress ± chin raise</td>
<td>15±17</td>
</tr>
<tr>
<td></td>
<td>Other lip/chin actions</td>
<td>14/16/17/21</td>
</tr>
</tbody>
</table>

**Note.** Alternative action units (AUs) are indicated by /; Optional AUs are indicated by ±. Commas designates an “and/or” relationship. Intensity of the facial actions are indicated by small letters (a through e). Further details of the coding scheme are found in BabyFACS (Baby Facial Action Coding System; Oster & Rosenstein, 1996) and may be requested from the authors at the Department of Psychology, Newark University, 6 Washington Place, New York, New York 10003 or through electronic mail addressed to osterh@psych.nyu.edu.

* The cry mouth movement codes were scored only when other negative components were also present in the brows, eyes, cheeks, or midface region (i.e., AUs 3, 4, 6, 7, 9, 10 and/or 11). Categorization of the cry mouth was based on a 3 X 3 matrix specifying intensity ranges for horizontal lip stretching (AU 20) and vertical mouth opening (AUs 25/26/27; see Oster & Rosenstein, 1996, for details). Horizontal lip stretch alone was scored when AU 20 occurred in a context other than one of the cry mouths. The modulated negative mouth codes represent distinctive infant facial configurations that occur in contexts suggesting an effort to regulate or modulate negative affect and/or competing tendencies to cry and to inhibit crying (see Oster & Rosenstein, 1996, for details). This code was used when AU 15 occurred but did not create the distinctive "horseshoe" configuration usually because of the co-occurrence of other actions or the minimal intensity of AU 15. The code was used when AU 17 occurred but did not create the distinctive "pout or lower lip out" configuration usually because of its minimal intensity or the co-occurrence of other actions.

Received December 16, 1996
Revision received October 10, 1997
Accepted October 23, 1997