

Research Article

Sequential Dynamics of Culturally Moderated Facial Expressions of Emotion

David Matsumoto,¹ Bob Willingham,² and Andres Olide¹¹San Francisco State University and ²Center for Psychological Studies, Berkeley, California

ABSTRACT—*There is consensus that when emotions are aroused, the displays of those emotions are either universal or culture-specific. We investigated the idea that an individual's emotional displays in a given context can be both universal and culturally variable, as they change over time. We examined the emotional displays of Olympic athletes across time, classified their expressive styles, and tested the association between those styles and a number of characteristics associated with the countries the athletes represented. Athletes from relatively urban, individualistic cultures expressed their emotions more, whereas athletes from less urban, collectivistic cultures masked their emotions more. These culturally influenced expressions occurred within a few seconds after initial, immediate, and universal emotional displays. Thus, universal and culture-specific emotional displays can unfold across time in an individual in a single context.*

Although research has demonstrated that facial expressions of emotion are both universal and culture-specific, surprisingly little research has examined cultural differences in actual (not self-reported) emotionally expressive behaviors (Ekman, 1972; Matsumoto & Kupperbusch, 2001; Waxer, 1985; the most well-known of these studies—Ekman, 1972—was not published in a peer-reviewed journal). In addition, no study to date has examined exactly what aspects of culture are associated with emotional expressions. The consensus in the field is that when emotions are aroused, the displays of those emotions are either universal or culture-specific, depending on context. For example, in Ekman's (1972) study, American and Japanese participants' emotional expressions were similar when they viewed

stressful stimuli alone, but were different when they viewed the stressful stimuli with the experimenter. Ekman and Friesen (1969) coined the term *cultural display rules* to account for these differences and postulated that the regulation of expressions by cultural display rules results in six possible modes of expression when emotions are aroused: Emotion may be expressed naturally (expression), the display may express more emotion than is actually felt (amplification), the display may express less emotion than is actually felt (deamplification), no emotion may be shown (neutralization), the expression of the emotion felt may be blended with other expressions (qualification), and another emotion entirely may be displayed (masking).¹

We investigated the idea that an individual's emotional displays in a given context can be both universal and culturally variable as they unfold over time, because of the dual neural control of expressive displays. We expected this to be true especially when intense emotions are aroused in a context that dictates regulation of emotional behaviors. When emotions are aroused, impulses emanating from subcortical areas of the brain initiate the emotion-response system, which produces expressive behavior (LeDoux, 2000; Panksepp, 2008). At the same time, the facial nerve receives stimulation from cortical areas under voluntary control (Matsumoto & Lee, 1993; Rinn, 1991), and this stimulation is probably associated with display rules (Ekman & Friesen, 1969). Thus, when intense emotions are aroused in a social context, the initial and immediate emotional responses may be universal and originate from the subcortex; these may appear first because the neural connections from the subcortex are lower and closer to the facial nerve than those from the cortex. Culturally influenced emotional displays may occur subsequently, after there has been sufficient time for the additional neural work necessary for regulatory needs (i.e., display rules) to drive expressive behavior from the cortex.

Address correspondence to David Matsumoto, Department of Psychology, San Francisco State University, 1600 Holloway Ave., San Francisco, CA 94132, e-mail: dm@sfsu.edu.

¹A seventh expressive mode, simulation, occurs when individuals portray an emotion even though they are experiencing no emotion.

In fact, a previous study evaluated the initial, spontaneous expressions of Olympic athletes when they had just won or lost a match for a medal at the Olympic Games and demonstrated that these facial reactions were universal (and their configurations matched those that have been demonstrated to express emotion in studies of production and judgment of facial expression; Matsumoto & Willingham, 2006). The athletes' subsequent expressions, however, were not analyzed. We conducted such an analysis in the study reported here, examining two important questions: Did these subsequent emotional displays exhibit cultural differences, and, if so, when did those differences first appear?

Examining cultural differences requires operationalizing meaningful aspects of culture and empirically linking these cultural variables to the data (Matsumoto & Yoo, 2006). Given that the Olympic athletes came from many different cultures around the world, it was possible to select from a broad range of cultural variables that might be related to the athletes' expression regulation. We focused on two broad domains. One consisted of country demographics: population density, affluence, and religion. These variables were selected because they have been theoretically (Triandis, 1972, 1995) and empirically (Georgas, van de Vijver, & Berry, 2004) related to culture, ways of life, and culturally based behavior regulation. High population densities, for example, may influence expression regulation because of the social impact of emotions on individuals in the immediate environment; thus, we predicted that higher population density would be associated with greater regulation. Affluence provides individuals with greater resources to survive without reliance on others, lessening the need for social modulation and conformity; thus, we predicted that affluence would be associated with less regulation and greater expression.

The second domain consisted of cultural values. We focused on Hofstede's (2001) five cultural dimensions: individualism versus collectivism (IC), power distance, uncertainty avoidance, masculinity, and time orientation. These dimensions appear to broadly capture cultural differences in a wide range of psychological variables. In particular, IC has been related to cultural norms for expression (Matsumoto, Yoo, Fontaine, et al., 2008; Matsumoto, Yoo, Nakagawa, et al., 2008); individualistic cultures are associated with greater endorsement of expression, whereas collectivistic cultures are associated with greater expression regulation.

We built on our previous work (Matsumoto & Willingham, 2006), examining changes in the same Olympic athletes' expressions after their initial reactions and classifying the expressions according to expressive styles. We then examined the relationship between these expressive styles and the cultural and demographic variables. We hypothesized that the expression of emotion would be negatively related to population density, but positively related to affluence and IC. In addition, we examined how quickly after the initial reactions the subsequent expressions occurred. Expressive styles requiring greater modification of the original response

(neutralization, qualification, and masking) should require more neurocognitive work than expressive styles requiring less modification (expression or deamplification). Thus, we hypothesized that the former expressive styles would be associated with greater elapsed time from first to subsequent expressions than the latter expressive styles.

METHOD

The Setting

Expressions were recorded during the judo competition at the 2004 Athens Olympic Games. The stadium contained two competition areas measuring 8 m × 8 m; both were located on elevated platforms. A 3-m safety area bordered three of the outside edges of each competition area, and a 4-m safety area separated the two competition areas. The second author, who is a professional photographer and was the official photographer of the International Judo Federation, was situated between the two competition areas on the same side as the technical officials, opposite the main spectator seating and the main television cameras.

In judo, there are seven weight categories each for men and women; on each of 7 days, one weight category for each sex was contested. Except for gold-medal matches, competitions were held concurrently in the two competition areas (two bronze medals, and two fifth-place standings, are awarded in judo, and matches are held simultaneously); the photographer took shots from both competitions, alternating between them depending on the action and the athletes competing.² Final, gold-medal matches occurred one at a time, so the photographer was able to focus all attention on those matches.

Athlete Participants

The pool of athletes comprised the 84 gold, silver, bronze, and fifth-place winners of the judo competition at the 2004 Athens Olympic Games. They represented 35 countries from six continents³ and constitute the most culturally diverse sample in which spontaneous expressions in a highly charged, emotional event have been examined.

Photographic Equipment and Shooting

The photographer took action shots during the contests. For the purposes of this study, however, he also took shots of the athletes after match completion. Judo matches can end in one of five ways: Time runs out; an athlete is thrown cleanly with speed,

²Because the photographer had to alternate between athletes and competition areas, and because the athletes' actions were dynamic (so that at times they turned away from the camera), we believe that this study may underestimate the total number of expressions actually produced.

³In this data set, we know of only 1 athlete who was born in one country and represented another. The fact that many Olympic-level athletes have extensive international travel experience may have worked to reduce the possibility of finding cultural differences, but this would be an acceptable Type II error.

force, and control onto his or her back; an athlete pins the opponent to the ground for 25 s; an athlete submits because of the effects of a choke; or an athlete submits because of the effects of an armlock. Photographs of the athletes were taken from the time their match ended until the referee announced the decision, a time period that generally lasted 15 to 30 s. The photographer was told that the focus of the study was on expressions, but was given no information about the specific type of expressions or channel of expression; emotion was not mentioned. At the time of the Olympic Games, the photographer had no formal training in psychology and did not know either the literature related to the study or our hypotheses.

The photographer used a Nikon D2H professional digital camera, which has a high frames-per-second rate (8 frames per second, with 37-ms shutter-time lag) and high resolution (4.1 megapixels effective). The camera was set to use automatic focus and manual exposure using available light; shots were saved as JPEG files. The camera's ISO (International Standards Organization) sensitivity (digital equivalent of film speed) ranged between 400 and 800, giving shutter speeds around 1/500th of a second. A variety of interchangeable Nikkor lenses—including 28- to 70-mm f/2.8, 70- to 200-mm f/2.8, and 300-mm f/2.8 lenses—were used.

Selection of Photographs and Expressions

Approximately 3,000 photographs were taken during each day's competition, resulting in about 21,000 photographs across the 7 days. We examined all photographs taken from the precise moment when a medal match was over until the decision was announced by the referee. This resulted in a preliminary selection of 2,735 photographs. Of the 84 athletes, there were no usable photos for 6.⁴

Expressions were identified in these photographs using two criteria: inclusion of a clear view of at least a profile of the face and contraction of at least one facial muscle. When an expression was identified, we examined the series of photographs from the beginning to the end of that expression and then selected the apex of the expression for subsequent coding.⁵ A face was determined to display a second expression if there was a change in which facial muscles were contracted or if the existing facial muscle contractions increased by at least two intensity levels (generally after a decrease during offset of the prior expression) according to the Facial Action Coding System (FACS; Ekman & Friesen, 1978).

⁴One bronze-medal match did not occur because of injury to an athlete. Another bronze-medal match ended quickly, and the photographer was focused on the other, simultaneously occurring, bronze-medal match. For two gold-medal matches, there was no usable photo for the silver medalist.

⁵In expressions involving multiple facial muscles, the muscles often have different timing dynamics, and thus reach their apex of contraction at slightly different times. We selected for coding and analysis the expressions in which the muscle configurations as a whole appeared to be at their highest intensity.

Expression Coding and Classification of Emotions and Expressive Styles

All expressions were coded using FACS (Ekman & Friesen, 1978). This system identifies each of the facial muscle movements (action units, or AUs) that can occur independently, as well as head and eye positions. Two certified FACS coders (one was author D.M., the other was blind to the hypotheses and goals of the study) coded each of the expressions for any AUs it exhibited. Interrater reliability, calculated by doubling the number of codes on which coders agreed and dividing by the total number of codes used, was .79.

FACS codes were then used to classify the expressions into emotion categories by matching the observed AUs to AU combinations theoretically or empirically related to emotions in previous work (published in Ekman & Friesen, 1975, 1978).⁶ Of the 118 expressions, 4 were not classifiable. The FACS-generated emotion categories included anger, contempt, disgust, fear, sadness, variations of happiness, and blends of different emotions.

We then returned to each athlete's sequence of photos and listed the changes in his or her emotional expressions in chronological order, noting the start and end time for each. This analysis included both the coded emotional expressions and the athlete's neutral faces (i.e., faces in which no expression was identified). Two coders (D.M. and A.O.) agreed on a classification for each athlete's set of expressions. The following categories of expressive style (adapted from Ekman & Friesen, 1969) were used for these classifications:

- *Express*: The athlete initially expressed an emotion and expressed the same emotion at least once more.
- *Deamplify*: The athlete initially expressed an emotion and subsequently expressed it again at lower intensity or together with facial controls. (Facial controls were facial actions involving AU 14, 17, or 24 from FACS. These facial actions give the appearance of "biting one's lip," or trying to control one's feelings.)
- *Neutralize*: The athlete initially expressed an emotion and subsequently displayed no other emotions, displayed facial controls only, or displayed no emotions throughout.⁷

⁶The specific AU combinations on which FACS bases its emotion classifications have been published and are available for public scrutiny. These AU combinations typically consist of components of full-face, prototypical expressions, and have been associated with emotion signaling in a wide range of studies of actual expression production by individuals from all parts of the world, not just Westerners (Matsumoto, Keltner, Shiota, Frank, & O'Sullivan, 2008). This ensures that the facial configurations predicted to be associated with emotion are not just Western prototypes.

⁷Some athletes may have simply had a neutral state after expressing an emotion, and may not have been trying to neutralize an expression per se. From a behavioral perspective, it is easier to interpret facial controls as neutralization attempts than as expressions followed by no expression. (A number of athletes coded in the "neutralize" category exhibited expressions followed by a neutral state.) If an athlete merely displayed nothing, it is impossible to know whether he or she was making a neutralization attempt or experiencing a neutral state, and the results should be interpreted with this caveat.

- *Qualify*: The athlete initially expressed an emotion and then expressed it again at least once blended with other emotions.
- *Mask*: The athlete expressed nothing or an emotion and then displayed a nonenjoyment smile.

Amplification and simulation did not occur in our data set (we assumed that the athletes’ emotional displays were of felt emotions, not simulations).

Ten athletes’ expressions could not be classified because these athletes contributed only one expression and there were no other photos of these individuals to determine what they did subsequently. Thus, the final sample consisted of 68 athletes (84 athletes – 6 with no photos – 10 with only one photo). (The expressive styles were distributed as follows—express: $n = 24$; deamplify: $n = 2$; neutralize: $n = 38$; qualify: $n = 1$; mask: $n = 3$.) Because of small sample sizes, we combined categories, classifying “express” and “deamplify” as “express,” and all other expressions as “mask” (because neutralization can be considered a form of masking).

Cultural Variables

For each country represented by an athlete, we obtained population size, percentage of arable land, national religion, and per capita purchasing power parity (as a proxy for affluence) from *The World Factbook* (Central Intelligence Agency, 2004). Data were available for all countries. We computed a population-density index by dividing the total population by the total amount of arable land.

We also obtained each country’s scores on Hofstede’s (2001) five cultural dimensions: IC, power distance, uncertainty avoidance, masculinity, and time orientation. (Index scores were used.) Country data on the first four dimensions were available for 58 athletes; time-orientation data were available for 48 athletes.

RESULTS

Cultural Differences

Separate regressions on expressive style (express vs. mask) using affluence (per capita purchasing power parity), population density, and the five cultural dimensions as independent variables revealed significant effects of affluence, $R(68) = .29, p < .05$; population density, $R(68) = .27, p < .05$; power distance, $R(48) = .29, p < .05$; and IC, $R(48) = .27, p < .05$. Athletes from countries with more affluence, higher population density, more individualism, and lower power distance expressed their emotions more, whereas athletes from countries with less affluence, lower population density, more collectivism, and higher power distance masked their emotions more (Table 1).

To examine how each of these four variables was associated with expressive style, and to control for the intercorrelations among them, we computed a simultaneous regression, entering affluence, population density, power distance, and IC. The model accounted for 43% of the variance in expressive style,

TABLE 1

Means of the Significant Independent Predictors of Expressive Style Among Athletes Who Expressed Their Emotions and Those Who Masked Their Emotions

Predictor	Expressive style	
	Express	Mask
Affluence (per capita purchasing power parity)	\$28,242.31 (\$12,054.91)	\$20,495.24 (\$12,836.74)
Population density	1,465.35 (1,065.16)	890.89 (1,001.91)
Power distance	48.95 (17.50)	61.29 (17.91)
Individualism versus collectivism	61.25 (19.58)	49.26 (21.49)

Note. Standard deviations are given in parentheses.

$R(48) = .66, p < .001$. Affluence, population density, and IC had significant standardized coefficients (Table 2). Thus, the athletes who came from countries with higher values for these dimensions were more expressive.

Several nonfindings are noteworthy. We used chi-square tests to examine the association between national religion and the expressive styles, but the results were not significant. Also, match outcome and sex of the athlete did not have significant effects. Thus, the findings were not affected by these variables.

Sequential Dynamics

Despite universality in the athletes’ initial emotional reactions (Matsumoto & Willingham, 2006), our analyses revealed cultural differences in the subsequent emotional displays. How quickly did the cultural influences kick in? Fortunately, each of the photographs had a time stamp precise to the second. Thus, we were able to measure the time elapsed from the end of an athlete’s first expression to the beginning of his or her second expression. (Because there was only one photographer who went back and forth between the athletes and competition areas, we did not always have continuous photos for the expressions of a given athlete. For these analyses, we used only those sequences that were shot continuously without breaking away to the other athlete or match, $n = 48$.) Because these data were not normally distributed, we recoded the data using a log-10 transformation.

TABLE 2

Standardized Coefficients for the Significant Predictors of Expressive Style in the Simultaneous Multiple Regression

Variable	β
Population density	-1.59*
Affluence (per capita purchasing power parity)	-2.046*
Individualism versus collectivism	-2.103*

* $p < .001$.

Elapsed time between the initial and subsequent expressions did not differ between athletes who won and those who lost their matches. But, as predicted, elapsed time was shorter for athletes who expressed their emotions ($M = 0.216$, $SD = 0.313$) than for those who masked their emotions ($M = 0.378$, $SD = 0.356$), $t(48) = 1.73$, $p < .05$, $d = 0.46$), a finding consistent with the idea that greater processing time is required for more regulated expressions. In addition, the mode elapsed time was zero, which indicates that many new, culturally influenced expressions began within a second of the initial universal display. Elapsed time was negatively correlated with IC, $r(48) = -.37$, $p < .05$; athletes from individualistic cultures had shorter elapsed times than athletes from collectivistic cultures. (IC was not correlated with match outcomes, so this result did not occur because more athletes from individualistic cultures than from collectivistic countries won their matches.) Elapsed time was not correlated with any of the other cultural variables.

DISCUSSION

Although the athletes' initial emotional expressions were universal, their subsequent expressions were culturally regulated, and reliably associated with population density, affluence, and individualism. As predicted, athletes from individualistic cultures were more expressive than athletes from collectivistic cultures, who were more likely to mask their emotions. Contrary to our prediction, athletes from countries with high population densities were more expressive than athletes from countries with low population densities. One explanation for this unexpected finding is that affluence (per capita purchasing power parity) was highly correlated with population density when analyses controlled for IC, $\beta = .88$, $p < .001$; countries with high population densities were more affluent than countries with low population densities. Thus, individuals from relatively urban, individualistic cultures expressed their emotions more than those from less urban, collectivistic cultures.

These findings demonstrate that, across time, a given individual's emotional expressions in a single context can be both universal and culture-specific. As mentioned earlier, we believe this is true because of the dual neural control of facial expressions, from both the subcortex and the cortical motor strip. We posit that when intense emotions are aroused in situations that require expression regulation, neural impulses from these areas compete to innervate the facial nerve; individuals may display a variety of expressions in a single context as the impulse to express emotion competes with the various learned display rules that regulate expressive behavior. Universal emotional reactions occur first because of the proximity, and direct neural connections, of the subcortical areas controlling expression to the facial nerve. Culturally regulated expressions are subsequent to the initial universal reactions because their impulses emanate from the cortex and require networking of learned display rules. Our data suggest that cultural influences kick in within 1 or 2 s after the onset of an initial universal display.

Expressive styles involving greater modification of the initial reaction required more time for display than expressive styles involving less modification, likely because greater modification requires greater neurocognitive resources. Expressive modes that allow for the continued expression of the initial emotion or only slight alteration of its intensity (deamplification) require less modification of the facial expression, and thus in our study resulted in shorter elapsed times from initial response. Because IC was associated with expressive style, it is no wonder that it was also correlated with elapsed time.

The fact that expressions change across time and are culturally variable subsequent to an initial, immediate, universal emotional reaction explains why people believe that cultural differences in expression are pervasive. When intense emotions are aroused, observers' attention is often drawn to the stimulus event and not the expressive behaviors of the individuals in that event. While observers' attention is focused on the eliciting event, immediate universal reactions may occur but be missed. When attention returns to the individuals, they are already beginning to engage in culturally regulated behavior. Such a process may perpetuate beliefs about the cultural variability of expressive behavior. People tend to believe their experiences, and they believe in the existence of cultural differences in expressive behavior because that is what they often see.

Our study has some limitations, including the lack of corroborating measures of emotion. However, other studies have demonstrated that observers from different cultures reliably judge the expressions that were produced by the athletes in the present study (Matsumoto, Olide, Schug, Willingham, & Callan, in press), and subsequent interviews of these athletes provided self-reports that corroborated the emotions predicted by the facial expressions (Willingham & Matsumoto, 2007). These findings open the door to future research and theory on the temporal dynamics of culturally moderated facial expressions, as well as other types of regulated behaviors, and also provide an initial step toward the identification of the neural networks underlying such regulation. In particular, future studies may examine universality or cultural specificity in the unfolding of different expressions across time and meaning.

REFERENCES

- Central Intelligence Agency. (2004). *The world factbook*. Retrieved October 17, 2008, from <https://www.cia.gov/library/publications/the-world-factbook/>
- Ekman, P. (1972). Universal and cultural differences in facial expression of emotion. In J.R. Cole (Ed.), *Nebraska Symposium on Motivation, 1971* (pp. 207–283). Lincoln: Nebraska University Press.
- Ekman, P., & Friesen, W.V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*, 49–98.
- Ekman, P., & Friesen, W.V. (1975). *Unmasking the face: A guide to recognizing the emotions from facial cues*. Englewood Cliffs, NJ: Prentice Hall.
- Ekman, P., & Friesen, W.V. (1978). *Facial Action Coding System: Investigator's guide*. Palo Alto, CA: Consulting Psychologists Press.

- Georgas, J., van de Vijver, F.J.R., & Berry, J.W. (2004). The ecocultural framework, ecosocial indices, and psychological variables in cross-cultural research. *Journal of Cross-Cultural Psychology*, *35*, 74–86.
- Hofstede, G.H. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). Thousand Oaks, CA: Sage.
- LeDoux, J.E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.
- Matsumoto, D., Keltner, D., Shiota, M.N., Frank, M.G., & O'Sullivan, M. (2008). What's in a face? Facial expressions as signals of discrete emotions. In M. Lewis, J.M. Haviland-Jones, & L.F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 211–234). New York: Guilford Press.
- Matsumoto, D., & Kupperbusch, C. (2001). Idiocentric and allocentric differences in emotional expression and experience. *Asian Journal of Social Psychology*, *4*, 113–131.
- Matsumoto, D., & Lee, M. (1993). Consciousness, volition, and the neuropsychology of facial expressions of emotion. *Consciousness & Cognition: An International Journal*, *2*, 237–254.
- Matsumoto, D., Olide, A., Schug, J., Willingham, B., & Callan, M. (in press). Cross-cultural judgments of spontaneous facial expressions of emotion. *Journal of Nonverbal Behavior*.
- Matsumoto, D., & Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners at the 2004 Athens Olympic Games. *Journal of Personality and Social Psychology*, *91*, 568–581.
- Matsumoto, D., & Yoo, S.H. (2006). Toward a new generation of cross-cultural research. *Perspectives on Psychological Science*, *1*, 234–250.
- Matsumoto, D., Yoo, S.H., Fontaine, J.R.J., Anguas-Wong, A.M., Ariola, M., Ataca, B., et al. (2008). Mapping expressive differences around the world: The relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology*, *39*, 55–74.
- Matsumoto, D., Yoo, S.H., Nakagawa, S., Alexandre, J., Altarriba, J., Anguas-Wong, A.M., et al. (2008). Culture, emotion regulation, and adjustment. *Journal of Personality and Social Psychology*, *94*, 925–937.
- Panksepp, J. (2008). The affective brain and core consciousness: How does neural activity generate emotional feelings? In M. Lewis, J.M. Haviland-Jones, & L.F. Barrett (Eds.), *Handbook of emotions* (3rd ed., pp. 47–67). New York: Guilford Press.
- Rinn, W.E. (1991). Neuropsychology of facial expression. In R. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior* (pp. 3–70). New York: Cambridge University Press.
- Triandis, H.C. (1972). *The analysis of subjective culture*. New York: Wiley.
- Triandis, H.C. (1995). *New directions in social psychology: Individualism and collectivism*. Boulder, CO: Westview Press.
- Waxer, P.H. (1985). Video ethology: Television as a data base for cross-cultural studies in nonverbal displays. *Journal of Nonverbal Behavior*, *9*, 111–120.
- Willingham, B., & Matsumoto, D. (2007). *The thrill of victory and the agony of defeat*. Bristol, England: World of Judo.

(RECEIVED 11/10/08; REVISION ACCEPTED 2/13/09)