

American–Japanese Cultural Differences in Judgements of Expression Intensity and Subjective Experience

David Matsumoto, Fazilet Kasri, and Kristie Kookan
San Francisco State University, USA

A number of studies have reported cultural differences in intensity ratings of facial expressions of emotion. In the previous research, however, observers made only a single intensity rating; thus, it was not clear whether observers rated the external display, or made an inference about the subjective experience of the poser. In this study, we obtained these two intensity ratings separately from American and Japanese observers. Results indicated that Americans perceived greater intensity in display, but Japanese inferred greater intensity of subjective experience. When examined within-culture, Americans rated display more intensely than subjective experience, whereas there was no difference between the two ratings for the Japanese. We discuss these findings in relation to the concept of cultural decoding rules, and outline an agenda for future research that examines the exact nature of these rules, the relationship between decoding, display rules and self-construals, and the role of context in judging emotion.

INTRODUCTION

Research on judgements of facial expressions of emotion has shown high levels of agreement in what emotions are judged in literate (Ekman, Sorenson, & Friesen, 1969; Izard, 1971) and preliterate cultures (Ekman & Friesen, 1971; Ekman et al., 1969). These findings replicate when

Requests for reprints should be sent to Professor David Matsumoto, Department of Psychology, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA, 94132, USA; e-mail: dm@sfsu.edu.

This research was supported in part by a Faculty Affirmative Action Grant and a California State University Award for Research, Scholarship, and Creative Activity awarded to the first author.

We thank Nathan Yrizarry, Sherry Loewinger, Katherine Turner, Cenita Kupperbusch, and Hideko Uchida for their assistance in our general research programme. We also express our gratitude to Paul Ekman for his comments on previous versions of this article.

spontaneous expressions are judged (Ekman, 1972), and when observers are allowed to rate the presence, absence, and intensity of multiple emotions (Ekman et al., 1987). Cultures also agree on the second modal response when judging certain faces, and relative intensity differences within emotions (Ekman et al., 1987; Matsumoto & Ekman, 1989). These studies collectively provide much of the evidence for the universality of facial expressions (see also Ekman's, 1972, cross-cultural study which involves the measurement of expression rather than judgement).

Cultures differ, however, in the absolute level of intensity they attribute to faces (Biehl et al., 1997; Ekman et al., 1987; Matsumoto, 1990). This effect has been found regardless of the race or sex of the poser being judged (Matsumoto, 1990; Matsumoto & Ekman, 1989) and across ethnicities (Matsumoto, 1993). Although agreement about categorical judgements is high, cultures also differ in the exact level of that agreement (Matsumoto, 1992; Russell, 1994; however, see also Ekman, 1994 and Izard, 1994), and those differences are related to stable and meaningful dimensions of cultural variability (Matsumoto, 1989).

Matsumoto and Ekman (1989) have suggested that differences in judgements are based on the existence of cultural decoding rules (cf. Buck, 1984). These are culturally prescribed rules learned early in life that manage the perception and interpretation of other's emotional expressions. Presumably, these rules create tendencies for people of any culture to amplify, deamplify, neutralise, or qualify their judgements, depending on the particular characteristics of their culture and the psychological needs transmitted to members of that culture. For example, Matsumoto (1989) suggested that cultures that encourage and maintain power and status differences among their members, and that encourage more collectivistic orientations, perceive less intensity in expressions of anger, fear, and sadness so as not to disrupt group harmony or status relationships, which would be against the social fabric of these cultures. These types of cultural characteristics may contribute to the consistent findings regarding cultural differences in intensity ratings reported in the previous research cited earlier (Biehl et al., 1997).

One limitation in our knowledge about cultural differences in ratings of intensity, however, was that the methods used in previous studies reporting these differences did not specify the exact nature of the ratings requested, either in studies that used single intensity ratings of emotion (e.g. Biehl et al., 1997) or in those using multiple scalar ratings (e.g. Ekman et al., 1987; Matsumoto & Ekman, 1989). Thus, it was not clear whether observers rated the external display itself, or the presumed, subjective experience underlying the expression. This delineation is crucial, as cultural decoding rules may act to influence judgements of one or both, and examination of possible variability in these differences is necessary to elucidate cultural influences on intensity judgements.

Cultural decoding rules may act on these two ratings in several ways. For example, some cultures may encourage greater attributions of intensity to external display of posers while minimising inferences about actual experience. Other cultures, however, may act in the exact opposite manner, minimising attributions of intensity to external displays relative to inferences about internal experience. Yet other cultures may equate their intensity ratings of display with inferences about feelings. And, these cultural influences can result in various combinations of differences when two or more cultures are compared on either or both types of ratings.

We distinguished these ratings in an American–Japanese comparison. This comparison is compelling as several studies have already documented that Americans rate emotional expressions more intensely (Biehl et al., 1997; Ekman et al., 1987; Matsumoto, 1990; Matsumoto & Ekman, 1989). We do not know, however, whether the differences are specific to ratings of external appearance, or whether they can be generalised to inferred emotional state. In Study 1, observers from both countries were shown facial expressions corresponding to seven emotions. For each expression, observers made a categorical judgement and the two intensity ratings. We speculate here that American–Japanese cultural differences in intensity ratings are specific to judgements of external appearance, and that Japanese judges will attribute more intense emotional experiences than will Americans. This speculation is based on findings of previous studies that have shown that the Japanese deamplify their emotional expressions, at least in certain social contexts (Ekman, 1972; Matsumoto, 1990) relative to their American counterparts, which suggests that they may infer greater felt experiences than what they perceive on the outside. Thus, the first part of this study tested the following hypotheses:

Hypothesis 1: Americans will give significantly higher ratings than the Japanese on external expression, but

Hypothesis 2: Japanese will give significantly higher ratings on subjective experience.

If Hypotheses 1 and 2 are supported, differences between expression and experience ratings within each culture would need to be examined in order to elucidate the nature of cultural influences on these judgement processes. There are at least three mutually exclusive explanations of the findings, each of which has a different implication for our understanding of the nature of cultural decoding rules in each country. First, in the US, expression ratings may be moderately higher than experience ratings (e.g. 8.0 vs. 5.0), but the opposite is true in Japan. Second, in the US, expression ratings may be much higher than experience ratings (e.g. 8.0 vs. 2.0), although there is no difference in Japan (e.g. 5.0 vs. 5.0). Third, there may be no difference between the two ratings in the US (e.g. 5.0 vs. 5.0), whereas in

Japan, experience is rated higher than expression (e.g. 7.0 vs. 3.0, respectively).

Because of the multiple ways by which these differences could occur, we tested the following nondirectional hypothesis:

Hypothesis 3: That there will be within-culture differences between expression and experience ratings, and the nature of these differences will be different for Americans and Japanese.

(We acknowledge here that, if Hypotheses 1 and 2 are supported, Hypothesis 3 by default must be supported in this general version due to the interaction between culture and rating type. We offer this hypothesis solely as a means to explore the exact nature of the differences between the ratings within both cultures.)

STUDY 1

Method

Observers. The observers were 128 Americans (85 females and 43 males) and 80 Japanese (47 females and 33 males). The Americans were university graduates who were born and raised in the US, as were their parents, and whose first language was English. They were recruited from introductory psychology classes at San Francisco State University, and participated in partial fulfilment of class requirements. Forty-six reported their ethnicities as Caucasian; 12 were African Americans; 30 were Asian Americans including individuals of Filipino descent; 17 were Hispanic/Latino Americans; and the remaining 23 were other classifications. The Japanese were all born and raised in Japan, as were their parents, and their first language was Japanese. They were recruited from psychology classes at Osaka University of Education, and participated in fulfilment of class requirements.

Facial Stimuli. Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE) slides were used. This set contains 56 expressions, each posed by a different individual. There are eight examples of seven universal emotions: anger, contempt, disgust, fear, happiness, sadness, and surprise. Within each emotion, there are two examples each of Japanese male and female posers, and of Caucasian male and female posers. Each expression corresponds to the prototypic universal expressions described by Ekman and Friesen (1975) and Ekman and Heider (1988), and as verified by Ekman and Friesen's (1978) Facial Action Coding System (FACS). This stimulus set has reliably produced

cross-cultural agreement in categorical emotion judgements under the methodological parameters described here (Biehl et al., 1997).

Judgement Tasks and Procedures. Observers were tested in small groups, and the procedures were exactly the same in both countries. All protocols were drafted in English and translated into Japanese; accuracy was verified by back translation. Observers were given the following instructions:

You will be shown 56 slides of different faces displaying expressions of emotion. Your task is to (1) determine what emotion is being displayed, (2) indicate the intensity level of that facial expression, and (3) indicate the intensity level of what you think the poser is actually feeling.

The first judgement was categorical, asking the observers to select from a list of seven alternatives the emotion term they thought best described the emotion portrayed. The alternatives were anger, contempt, disgust, fear, happiness, sadness, and surprise. For the next two ratings, observers used an 8-point scale (1–8) labelled A Little (1) to A Lot (8). An example of the use of the rating scale was given, highlighting the fact that the first rating was of the external display of the expression whereas the second was of the poser's internal experience. Observers were shown three expressions that were not used in this study, one at a time, and were asked to use the rating scales. Questions were answered, and the experiment started only after ascertaining that the observers understood their task.

The stimuli were presented one at a time, for 5sec (with a 5sec inter-stimulus blank), in a random order except with the criterion that no more than three stimuli of the same emotion, race or sex appeared in succession. The presentations proceeded uninterrupted until all slides were judged. Completion of the ratings for the last slide marked the end of the experiment.

Results

Preliminary Analyses and Standardisation of Scores. We averaged each participant's ratings across the two slides of each of the four poser types (Caucasian or Japanese males or females) within each of the seven emotions. We then computed a six-way Analysis of Variance (ANOVA) using country (2), observer sex (2), emotion (7), poser race (2), poser sex (2), and rating type (2) as independent variables, the latter four being repeated measures. A significant and large culture main effect (US > Japan) was found [$F(1, 155) = 28.75, P < .001, R^2 = .15$], which replicated

previous US > Japan differences using single intensity rating judgement tasks.

The culture main effect also indicated, however, that subsequent cultural differences found on either scale may be confounded by cultural response sets, which are cultural tendencies to use certain parts of a scale. If they exist, or at least the possibility that they are operating exists, cultural differences on the scales would be inconclusive because the possible cultural response sets confound the comparison. To address this issue, we standardised each observer's 112 ratings (i.e. 56 slides \times 2 ratings) to that observer's country mean and standard deviation across all 112 items. This procedure eliminates overall cultural response bias across all items, but leaves intact within culture differences on all within-subjects factors (i.e. emotion, poser race, poser sex, and rating type), because the standardisation procedure does not alter the relationship among these variables within an individual participant or within that participant's culture. Thus, we could still test for cultural differences on separate scales, and rating differences within cultures, even after scores were standardised (see Matsumoto, 1994, for a fuller discussion of this issue).

Hypotheses 1 and 2. To ensure further that the findings were not confounded by differences in the emotion category perceived in each expression, we analysed the rating data by including only those judges who selected the emotion category intended. We computed a three-way ANOVA separately for each of the 56 expressions, using country (2), judge sex (2), and rating type (2) as the independent variables. We were interested in the two-way interaction between country and rating type, as a significant effect here would indicate that country differences differed as a function of rating. Across the 56 analyses, 43 produced a significant two-way interaction between these factors (average R^2 effect size for the significant interactions = .074, SD = .04). For each, we examined the simple effects of country, using the error term that tested the original two-way interaction. For expression ratings, Americans had a higher mean than the Japanese 40 out of the 43 possible times; 30 of these were statistically significant (average R^2 for the significant effects = .07, SD = 0.05). For subjective experience ratings, Japanese had a higher mean than the Americans 38 out of the 43 possible times; 23 were statistically significant (average R^2 for the significant effects = .07, SD = 0.06). These results provide strong support for both Hypotheses 1 and 2. (A table reporting these findings in detail is available from the authors on request.)

We also tested these cultural differences using the standardised ratings averaged across both expressions of each of the four poser types (i.e. Japanese or Caucasian male or female) within each emotion. We computed a six-way ANOVA using culture, observer sex, emotion, poser

race, poser sex, and rating type as independent variables, the latter four factors being repeated measures. The five-way interaction between culture, emotion, poser race, poser sex, and rating type was significant [$F(6,930) = 6.83, P < .001$]. We then tested the simple effects of culture separately for each emotion, poser race, poser sex, and rating type, using the error term from the overall six-way ANOVA testing the significant five-way interaction (Table 1).

For the expression intensity ratings, 23 of the 28 analyses produced statistically significant results. Of these, 22 indicated that Americans gave significantly higher expression intensity ratings than did the Japanese. For the subjective experience ratings, 23 of the 28 analyses produced statistically significant results. Of these, 19 indicated that Japanese gave significantly higher subjective ratings than did the Americans. These analyses lent further support to Hypotheses 1 and 2.

Averages of 78% and 67% of the American and Japanese samples, respectively, were included in the above analyses. These averages are comparable to categorical emotion judgements reported in other studies. Inclusion of all observers regardless of judged emotion yielded essentially the same results, most likely because the bulk of the participants in these analyses were those that produced the just-cited results. It would have been best to compare participants who selected a nonintended emotion category with those who selected the intended category to examine whether the cultural differences reported above occurred regardless of emotion category judged. This comparison, however, would have yielded unreliable findings because of cell sizes of those not identifying the intended emotion being too small for statistical comparison, and thus was not done.

Hypothesis 3. We tested the simple effects of rating type, separately for each of the four poser types, seven emotions, and two countries. Significant effects were found in all 28 analyses for the Americans, and all indicated that the expression means were significantly greater than the subjective experience means. For the Japanese, only 9 of the 28 analyses were significant, five (one photo type of anger, three of contempt, and one of sadness) indicating that the experience means were greater, the other four indicating the reverse (two disgust, two surprise).

To determine the degree to which these differences were present at the level of the individual, we computed sign tests. A significantly greater proportion of Americans gave higher ratings to the expressions than to inferred experience in all 28 comparisons (average proportions expression > experience = 68%, experience > expression = 18%, ties = 14%, z s range from 2.07 to 8.06, P s range from < .05 to < .0001). For the Japanese,

TABLE 1
American–Japanese Differences in Expression and
Experience Ratings (Study 1)

	<i>US</i>	<i>Japan</i>	<i>F</i>	<i>P</i>
Expression Ratings				
<i>Anger</i>				
JM*	0.59 (1.90)	−0.45 (1.83)	54.08	<.001
JF	0.52 (1.87)	−0.58 (1.68)	61.24	<.001
CM	1.49 (1.78)	1.35 (1.79)	1.03	n.s.
CF	0.55 (1.83)	−0.07 (1.50)	19.52	<.001
<i>Contempt</i>				
JM	−1.33 (2.23)	−1.69 (1.71)	6.48	<.001
JF	−0.87 (2.39)	−0.62 (1.96)	3.28	n.s.
CM	−1.15 (1.97)	−1.74 (1.77)	16.76	<.001
CF	−1.50 (1.96)	−1.39 (1.61)	.64	n.s.
<i>Disgust</i>				
JM	1.57 (1.48)	0.54 (1.87)	52.96	<.001
JF	1.18 (1.84)	1.58 (1.96)	7.91	<.001
CM	1.33 (1.55)	0.60 (1.74)	26.54	<.001
CF	0.93 (1.82)	−0.16 (1.84)	59.70	<.001
<i>Fear</i>				
JM	1.70 (1.66)	1.49 (1.52)	2.29	n.s.
JF	1.29 (1.86)	0.25 (1.73)	54.60	<.001
CM	−0.18 (1.82)	−0.25 (1.97)	.24	n.s.
CF	1.65 (1.53)	0.68 (1.70)	48.09	<.001
<i>Happiness</i>				
JM	2.08 (1.21)	1.22 (1.52)	37.73	<.001
JF	2.33 (1.21)	1.87 (1.51)	10.29	<.001

TABLE 1
American–Japanese Differences (continued)

	<i>US</i>	<i>Japan</i>	<i>F</i>	<i>P</i>
CM	0.41 (1.91)	−0.33 (1.70)	28.26	< .001
CF	1.83 (1.43)	0.88 (1.73)	45.54	< .001
<i>Sadness</i>				
JM	0.13 (1.77)	−1.33 (1.68)	109.46	< .001
JF	0.66 (1.75)	−1.07 (1.80)	150.99	< .001
CM	−0.19 (2.01)	−0.85 (1.69)	21.90	< .001
CF	−0.06 (1.85)	−0.76 (1.76)	25.27	< .001
<i>Surprise</i>				
JM	0.46 (1.72)	−0.53 (2.06)	50.51	< .001
JF	1.74 (1.60)	0.27 (1.74)	109.13	< .001
CM	0.52 (1.83)	−0.13 (1.72)	21.50	< .001
CF	1.22 (1.57)	−0.35 (1.64)	125.66	< .001
Experience Ratings				
<i>Anger</i>				
JM	−1.45 (2.27)	−0.77 (1.78)	23.18	< .001
JF	−1.09 (2.25)	−0.06 (1.75)	54.21	< .001
CM	−0.84 (2.46)	0.97 (1.78)	165.58	< .001
CF	−1.18 (2.38)	0.19 (1.89)	94.45	< .001
<i>Contempt</i>				
JM	−2.08 (2.07)	−1.06 (1.89)	52.29	< .001
JF	−1.48 (2.54)	0.70 (2.04)	238.59	< .001
CM	−2.28 (2.10)	−1.23 (2.00)	55.75	< .001
CF	−1.97 (2.30)	−1.42 (1.60)	15.10	< .001

TABLE 1
American–Japanese Differences (continued)

<i>Disgust</i>				
JM	−0.71 (2.16)	0.04 (1.83)	28.52	<.001
JF	−0.54 (2.13)	0.35 (1.90)	40.37	<.001
CM	−0.61 (2.14)	0.40 (1.97)	50.84	<.001
CF	−0.62 (1.98)	−0.34 (1.84)	3.84	n.s.
<i>Fear</i>				
JM	−0.17 (2.52)	1.74 (1.64)	188.17	<.001
JF	0.07 (2.20)	0.35 (1.88)	3.94	<.05
CM	−1.42 (2.15)	−0.26 (1.98)	68.81	<.001
CF	−0.36 (2.38)	0.83 (1.82)	72.77	<.001
<i>Happiness</i>				
JM	1.64 (1.48)	1.02 (1.64)	19.65	<.001
JF	2.00 (1.53)	1.97 (1.68)	.03	n.s.
CM	−0.37 (2.13)	−0.02 (1.89)	6.27	<.05
CF	1.18 (1.83)	0.80 (1.59)	7.15	<.05
<i>Sadness</i>				
JM	−0.68 (2.30)	−1.07 (1.75)	7.89	<.001
JF	−0.35 (2.20)	−0.46 (2.05)	.56	n.s.
CM	−1.17 (2.50)	−0.56 (1.77)	19.26	<.001
CF	−0.83 (2.43)	−0.07 (1.79)	41.22	<.001
<i>Surprise</i>				
JM	−1.30 (2.18)	−1.03 (2.06)	3.71	n.s.
JF	0.45 (2.15)	0.08 (1.78)	7.05	<.05
CM	−1.17 (2.16)	−0.75 (1.73)	9.37	<.001
CF	−0.43 (2.28)	−0.30 (1.89)	.79	n.s.

* JM, Japanese male; JF, Japanese female; CM, Caucasian male; CF, Caucasian female.

only 8 of the 28 comparisons were significant; of these, four were in one direction, and the other four were reversed. Thus, the Japanese tended to rate expressions and experience at the same intensity levels.

Post-hoc Analyses. Cultural decoding rules may act to produce differences not only in mean tendencies, but also in the relationship between the two rating types. For example, in some cultures the relation between the two ratings may be strong and positive, where attributions of strong displays are associated with inferences of strong emotional experience. In other cultures, the relationship may be strong and negative, where inferences of strong experience are associated with subtle or low intensity displays. In yet other cultures, the two ratings may not be associated at all. This study allowed us to examine these ideas because of the assessment of two intensity ratings.

Thus, we first computed the correlations separately for American and Japanese judges for each of the 56 expressions. There were positive correlations for each expression for both cultures (mean correlations across 56 photos were: $+ .49$, $SD = 0.10$ and $+ .54$, $SD = 0.09$, for Americans and Japanese, respectively). We compared the correlations obtained from both countries by transforming the r s to Fisher z s and comparing the z -scores. Of the 56 comparisons, only nine were statistically significant, seven indicating that the correlations for the Japanese judges were greater than those for the Americans.

The second method computed a correlation between the two ratings within each participant across the 56 expressions. We then treated each participant's r as a dependent variable, and computed a one-way ANOVA using country as the independent variable. There was no difference [mean r for Americans = $.56$, for Japanese = $.60$, $F(1,204) = .90$, n.s.].

Both sets of analyses indicated that the correlation between the two types of ratings was high and positive for both cultures, with no differences between them.

Discussion: Study 1

One possible limitation to Study 1 was the fact that the observers made all three ratings at the same time when viewing an expression. Being asked to rate the same stimulus three times, with two of the scales being similar, may be seen as an invitation to make different ratings on those scales, which may have contributed to the findings. One way to address this potential problem would be to present the stimuli three times, asking observers to make just one rating in each viewing. Study 2 addressed this possibility.

STUDY 2

Method

Observers. These were 50 American (35 females and 15 males, average age = 25.02 years) university undergraduates who were born and raised in the US, as were their parents, and whose first language was English. They were recruited from a psychology course at San Francisco State University, and participated voluntarily. Fourteen reported their ethnicities as Caucasian; three were African Americans; 11 were Asian Americans including individuals of Filipino descent; 10 were Hispanic/Latino Americans; the remaining 12 were other classifications.

Facial Stimuli, Judgement Tasks, and Procedures. The judgement tasks, stimuli, and procedures were exactly the same as in Study 1, with the exception that observers were shown the stimuli three times. On the first pass, observers judged the emotion displayed; on the second pass, observers rated the intensity level of the external appearance of the expression; and on the third pass, they rated the intensity level of what they thought the poser was actually feeling. The stimuli were presented in the same random order on each pass. With 56 stimuli, it is unlikely that observers' ratings on one presentation affected their ratings on another.

Results

We again averaged the two ratings for each of the four poser types within each emotion, and computed a five-way ANOVA on the ratings, using observer sex as a between-subjects factor, and emotion, poser race, poser sex, and rating type as within-subjects factors. The emotion \times poser race \times poser sex by rating type interaction was significant [$F(6,258) = 13.66$, $P < .001$]. We then tested the simple effects of rating type separately for each emotion, poser race, and poser sex (Table 2). Of the 28 comparisons, 22 indicated that the mean for the external display rating was higher than that for the internal experience rating; of these, 15 were statistically significant. Of the 6 comparisons in the opposite direction, only one was statistically significant. These findings essentially replicated the findings for Hypothesis 3 for American observers in Study 1.

As in Study 1, we also computed sign tests between external display and internal experience ratings. Of the 28 comparisons, 25 indicated a greater proportion of observers who gave external display ratings a higher value than internal experience ratings; 17 were statistically significant. These findings also replicate those found in Study 1.

TABLE 2
Means and Results of Simple Effects Comparisons of Rating
Type Separately for Each Emotion, Poser Race, and Poser
Sex (Study 2)

<i>Expression & Poser Types</i>	<i>Means</i>		<i>F</i>	<i>P</i>
	<i>Display</i>	<i>Experience</i>		
JM Anger	5.99	4.90	24.60	< .001
JF Anger	5.51	4.61	17.88	< .001
CM Anger	6.15	4.92	17.85	< .001
CF Anger	5.41	4.92	2.71	n.s.
JM Contempt	3.18	3.02	1.30	n.s.
JF Contempt	3.23	3.43	0.97	n.s.
CM Contempt	3.33	3.13	2.13	n.s.
CF Contempt	3.85	3.67	1.07	n.s.
JM Disgust	5.98	4.84	21.08	< .001
JF Disgust	5.97	4.97	22.07	< .001
CM Disgust	5.94	4.99	19.63	< .001
CF Disgust	5.82	5.00	10.60	< .01
JM Fear	6.41	5.35	16.45	< .001
JF Fear	5.73	5.10	8.43	< .01
CM Fear	5.19	4.53	13.58	< .001
CF Fear	5.89	5.34	6.67	< .01
JM Happiness	6.15	6.26	0.51	n.s.
JF Happiness	6.18	6.18	0.00	n.s.
CM Happiness	5.17	5.26	0.32	n.s.
CF Happiness	6.01	5.99	0.02	n.s.
JM Sadness	4.45	4.23	1.67	n.s.
JF Sadness	4.45	4.81	4.29	< .05
CM Sadness	3.94	3.73	0.92	n.s.
CF Sadness	4.04	4.24	1.34	n.s.
JM Surprise	5.16	4.35	18.38	< .001
JF Surprise	5.58	5.08	6.38	< .01
CM Surprise	5.45	4.32	31.16	< .001
CF Surprise	4.45	3.98	8.30	< .01

JM, Japanese male; JF, Japanese female; CM, Caucasian male;
CF, Caucasian female.

We then computed the product-moment correlation between the two ratings separately for all expressions. Forty-four of the 56 correlations were statistically significant, and all were positive (mean $r = .33$, $SD = 0.15$). These findings replicated those found in Study 1.

Finally, we computed the product-moment correlation between the two ratings within each individual observer. Treating these r s as a dependent variable, we then tested the mean of the r s ($.54$, $SD = 0.29$) against zero using a single sample t -test. The mean r was significantly different than zero and positive, again replicating the results of Study 1 [$t(50) = 13.12$, $P < .001$].

GENERAL DISCUSSION

Americans gave higher ratings than Japanese to external appearance, but Japanese gave higher ratings than Americans on inferences about internal feelings. These findings were obtained because Americans rated expressions higher than experience, whereas Japanese attributed the same levels to expression and experience. There were also strong, positive correlations between the two types of ratings for both Americans and Japanese. Study 2 suggests that these findings could not be attributed to the fact that observers made all their ratings in one viewing of each stimulus.

These findings were hypothesised on the basis of the presumed existence of cultural decoding rules that manage the perception and interpretation of emotion. That Americans gave higher intensity ratings than the Japanese on external display replicates several previous studies that have reported the same finding (e.g. Biehl et al., 1997; Ekman et al., 1987; Matsumoto, 1989, 1990; Matsumoto & Ekman, 1989), and localises the source of this difference to ratings of the external expressions themselves. That Japanese gave higher intensity ratings than the Americans on internal, subjective experience is a new finding that highlights the importance of distinguishing between these two types of ratings. Earlier, we speculated that this difference may occur because the Japanese may be accustomed to compensating their inferences about internal experience, given that they are relatively more likely to alter their external displays than Americans depending on social circumstance. This possible link suggests that a relationship may exist between rules governing expressive behaviour (display rules) and rules governing perception and interpretation (decoding rules), and this possibility should be explored in future research. A number of other studies have demonstrated that emotion judgements are related to emotional expressions in the same participants, lending support to this notion (e.g. Lanzetta & Kleck, 1970; Levy, 1964; Zuckerman, Hall, DeFrank, & Rosenthal, 1976; Zuckerman, Larrance, Hall, DeFrank, & Rosenthal, 1979).

The findings regarding relative differences between the two ratings within both cultures were somewhat surprising. Although writers (e.g. Matsumoto & Ekman, 1989) have interpreted previous American-Japanese cultural differences in intensity ratings as occurring because the Japanese suppress their ratings, the present findings suggest that it is the Americans who exaggerate their ratings of external display, relative to inferences about internal experience, and not the Japanese who suppress. Indeed, there were no differences for the Japanese between the two rating types. Future research needs to explore the reasons why Americans downplay inferences about subjective experience relative to judgements about display intensity. We speculate that this pattern of interpretation suggests a particular way in which Americans regulate emotions. For example,

Americans may learn tendencies because they know that American display rules generally encourage outward displays of emotion that are stronger than what is actually felt. In this case, the cultural decoding rule may appear consistent with the cultural display rule, compensating for the amplification in display. It may also be that American display rules encourage a level of expression congruent to the level of emotion felt, but that it is functionally more adaptive for Americans to infer less intense felt emotion in others. This tendency may arise because inferences of strongly felt emotions in others may interfere with the primacy of self-concerns, given the generally independent construals of self-existent in the individualistic American culture (Markus & Kitayama, 1991). For example, their recognition may result in other-directed behaviours (e.g. helping behaviours) that are discrepant with self-concerns. Which of these is correct can be addressed empirically in future research where self-construals and behavioural tendencies are obtained along with emotion judgements.

At the same time, future research needs to explore why Japanese did not differ in their ratings of expression and experience. Again, we speculate that the fact that the Japanese attributed the same degree of intensity to both external display and internal experience suggests a particular way in which the Japanese regulate emotions. For example, it may be that emotional displays are deemphasised in Japan, which results in lower ratings of external display, and that because of this, inferences about the subjective experiences of others are raised in a compensatory fashion. This mechanism would be related to the relative importance placed on the existence and well-being of others in Japan, given that interdependent self-construals (Markus & Kitayama, 1991) are prevalent in a collectivistic culture. On the other hand, this finding may simply be related to the fact that emotional displays are contextualised based on display rules, so that when they are observed, they correspond to the same degree of felt emotion. Thus, when judges in our experiment view emotional faces, their attribution about the faces could only be that the situation allowed the expressions to occur, and when they do, they correspond to that degree of internal experience. As discussed earlier, these possibilities can and should be addressed empirically.

That the correlations between the two ratings were high and positive in both studies, both within individuals and within expressions, suggest that cultures may be broadly similar with respect to the relation between intensity of experience and intensity of expression. Future studies testing this relation in different cultures can assess the degree to which it is pancultural. If so, it may be another universal aspect of emotion, suggesting a link between the intensity of emotional experience and expression. These correlations also suggest that cultural influences on these ratings involve a "shifting" of the absolute levels of the ratings being made, and

not a switching of the relation between the ratings. Studies examining shifts in ratings utilising expressions of different intensity levels with varying contextual information will elucidate the nature and basis for these shifts. Future studies could also examine the degree of shift in decoding rules across subcultures or ethnic groups within a larger culture, such as the US.

New research should also embark on uncovering the exact nature of the cultural decoding rules that are presumed to underlie the findings obtained here. Indeed, there is now a substantial literature demonstrating that cultures differ in their judgements of emotion, and much of this literature has invoked the concept of cultural decoding rules to account for their findings. Yet, little is known about the nature of these rules. Are they propositional? Volitional? Must they be verbal and explicit? Clearly, the mere existence of cultural differences by itself does not necessarily establish the existence of cultural decoding rules. Future research can elucidate the nature of these rules, and uncover how and why they produce differences in judgements of emotion. Such research also needs to address the relationship between culture and those rules, linking the particular characteristics of different cultures to the exact nature of those rules.

Although Study 2 addressed some possible methodological influences on the data, consideration needs to be given to other methodological parameters within which these studies were conducted, and how these may have influenced the results. The lack of contextual information may play an important role in producing these findings, and future judgement research needs to build context into research designs. Japanese culture is more contextualised than the American culture (Hall, 1969), and as such, Japanese judgements of expression or subjective experience may differ if contexts were manipulated. For that matter, American culture is not context-free, and such effects may be obtained with the Americans as well. A point related to context involves the degree to which the observers may have thought that the expressions were posed or not, and how this may have affected their judgements of either scale. Another related point involves the question of whether the findings would replicate with less intense emotional expressions than those portrayed in the JACFEE, or with partial and blended expressions. Future studies should incorporate these types of judgements as well, and involve posed and spontaneously expressed stimuli of different intensities.

Manuscript received 14 March 1997

Revised manuscript received 10 August 1998

REFERENCES

- Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Jacfee reliability and national Differences. *Journal of Nonverbal Behavior, 21*, 3-21.
- Buck, R. (1984). *The communication of emotion*. New York: Guilford Press.
- Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. In J. Cole (Ed.), *Nebraska Symposium on Motivation*. Lincoln, NE: University of Nebraska Press.
- Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's Mistaken Critique. *Psychological Bulletin, 115*, 268-287.
- Ekman, P., & Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124-129.
- Ekman, P., & Friesen, W.V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.
- Ekman, P., & Friesen, W.V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., Friesen, W.V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W.A., Pitcairn, T., Ricci-Bitti, P.E., Scherer, K., Tomita, M., & Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*, 712-717.
- Ekman, P., & Heider, K.G. (1988). The universality of a contempt expression: A replication. *Motivation and Emotion, 12*, 303-308.
- Ekman, P., Sorenson, E.R., & Friesen, W.V. (1969). Pan-cultural elements in facial displays of emotion. *Science, 164*, 86-94.
- Hall, E.T. (1969). *The hidden dimension*. Garden City, NY: Doubleday.
- Izard, C. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.
- Izard, C. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*, 288-299.
- Lanzetta, J.T., & Kleck, R.E. (1970). Encoding and decoding of nonverbal affect in humans. *Journal of Personality and Social Psychology, 16*, 12-19.
- Levy, P.K. (1964). The ability to express and perceive vocal communications of feeling. In J.R. Davitz (Ed.), *The communication of emotional meaning* (pp. 43-55). New York: McGraw-Hill.
- Markus, H.R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.
- Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20*, 92-105.
- Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14*, 195-214.
- Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology, 23*, 72-84.
- Matsumoto, D. (1993). Ethnic differences in affect intensity, emotion judgments, display rule attitudes, and self-reported emotional expression in an American sample. *Motivation and Emotion, 17*, 107-123.
- Matsumoto, D. (1994). *Cultural influences on research methods and statistics*. Pacific Grove, CA: Brooks/Cole.
- Matsumoto, D., & Ekman, P. (1989). American-Japanese cultural differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13*, 143-157.
- Russell, J.A. (1994). Is there universal recognition of emotion from facial expression? A review of cross-cultural studies. *Psychological Bulletin, 115*, 102-141.

- Zuckerman, M., Hall, J.A., DeFrank, R.S., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. *Journal of Personality and Social Psychology*, *34*, 966-977.
- Zuckerman, M., Larrance, D.T., Hall, J.A., DeFrank, R.S., & Rosenthal, R. (1979). Posed and spontaneous communication of emotion via facial and vocal cues. *Journal of Personality*, *47*, 712-733.

Copyright of Cognition & Emotion is the property of Psychology Press (T&F) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.