A NEW TEST TO MEASURE EMOTION RECOGNITION ABILITY: MATSUMOTO AND EKMAN'S JAPANESE AND CAUCASIAN BRIEF AFFECT RECOGNITION TEST (JACBART)

David Matsumoto, Jeff LeRoux, Carinda Wilson-Cohn, Jake Raroque, Kristie Kooken, Paul Ekman, Nathan Yrizarry, Sherry Loewinger, Hideko Uchida, Albert Yee, Lisa Amo, and Angeline Goh

ABSTRACT: In this article, we report the development of a new test designed to measure individual differences in emotion recognition ability (ERA), five studies examining the reliability and validity of the scores produced using this test, and the first evidence for a correlation between ERA measured by a standardized test and personality. Utilizing Matsumoto and Ekman's (1988) Japanese and Caucasian Facial Expressions of Emotion (JACFEE) and Neutral Faces (JACNeuF), we call this measure the Japanese and Caucasian Brief Affect Recognition Test (JACBART). The JACBART improves on previous measures of ERA by (1) using expressions that have substantial validity and reliability data associated with them, (2) including posers of two visibly different races (3) balanced across seven universal emotions (4) with equal distribution of poser race and sex across emotions (5) in a format that eliminates afterimages associated with fast exposures. Scores derived using the JACBART are reliable, and three studies demonstrated a correlation between ERA and the personality constructs of Openness and Conscientiousness, while one study reports a correlation with Extraversion and Neuroticism.

Research on judgments of emotion from facial expressions has a long and important history in psychology, and has contributed greatly to the literature concerning the universality of emotion, and to knowledge con-

David Matsumoto, Jeff LeRoux, Carinda Wilson-Cohn, Jake Raroque, and Kristie Kooken, San Francisco State University; Paul Ekman, University of California, San Francisco; Nathan Yrizarry, Sherry Loewinger, Hideko Uchida, Albert Yee, Lisa Amo, and Angeline Goh, San Francisco State University.

We thank Chu Kim, Sunita Paul, and Rebecca Ray for their assistance in our general research program.

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

Journal of Nonverbal Behavior 24(3), Fall 2000

cerning differences between gender, ethnicity, culture, and psychiatric status. Studies examining the relationship between individual differences in judgments of emotion (hereon referred to as Emotion Recognition Ability—ERA) and personality also have a considerable history, but is checkered with inconsistent findings. On one hand, ERA has been correlated with emotional expression (e.g., Lanzetta & Kleck, 1970; Levy, 1964; Zuckerman, Hall, DeFrank, & Rosenthal, 1976; Zuckerman, Larrance, Hall, DeFrank, & Rosenthal, 1979), self-monitoring (Mill, 1984; Mufson & Nowicki, 1991; Riggio & Friedman, 1982); social desirability (Cunningham, 1977); depression, control, aggression, and gregariousness (Toner & Gates, 1985); and social style, mental ability, achievement, and psychological mindedness (LeRoux, 1987). On the other hand, Cunningham (1977) failed to replicate a relationship between self-monitoring and ERA, and Zuckerman et al. (1979) found a relationship for women but not men. Buck, Savin, Miller, and Caul (1972) found a relationship between ERA and extraversion, as did Zuckerman et al. (1979). Cunningham (1977), however, did not replicate these findings, and instead found a relationship with neuroticism.

Theoretically, it is not unreasonable to consider that ERA should be related to stable personality traits. Individuals who are better at judging emotions in others should have greater degrees of interpersonal consciousness or concern; they should be more in tune with their environment, and with others. As an important component of our nonverbal communication system, such skills would be necessary for successful adaptation and manipulation of the environment, ensuring the stability and integrity of the self.

Because ERA is an important part of our daily lives, it is easy to consider how it should be related to various personality constructs, such as those specified in the five factor model. Extraversion, for example, is associated with stimulation seeking from others and the environment. As such, extraverts should be more willing to take in data concerning the emotions of others, being more interpersonally conscious of others in the environment. Individuals who score high on neuroticism, however, tend to be emotionally avoidant; because they are prone to experience negative emotions, they should have a tendency to avoid the recognition and awareness of others' emotions. The personality construct of openness is similar to extraversion in the sense that open individuals tend to be curious and interested in stimulation; they should be more attendant to the emotions of others. Conscientiousness is related to cooperation with and attending to others; conscientious individuals are more thorough, reliable, and efficient. They should be better at recognizing emotions because they are more attentive to details, and are better able to participate in such emotion judgment tasks.

Why have previous attempts to establish a relationship between personality and ERA been awash with contradictory findings? One possible reason is the stimuli used in previous studies, which were different in each study and thus not equivalent across the studies nor, as Bruner and Tagiuri (1954) suggested, did they cover a representative spectrum of emotional expressions (LeRoux, 1987). Another possible reason is the fact that, with only one exception (LeRoux, 1987, but these data are not published), many previous studies used measures specifically generated in each study rather than accepted, standardized tests. This distinction is important (O'Sullivan, 1982), because there is no guarantee that accuracy judgments were made against a valid standard.

If a standardized test were available, data could be generated using a valid standard, and the same test can be used across studies. At the very least, inconsistencies in the nature of the stimuli could be ruled out as a possible moderator of the contradictions.

Previous Tests of ERA

There has been a number of such tests developed in the past, each assessing some aspect of ERA (and its close relative, nonverbal decoding skills,)[1] but each with its own limitations (see review by O'Sullivan, 1982). Some focus on nonverbal behaviors, such as the Profile of Nonverbal Sensitivity (PONS: Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979), the Social Skills Inventory (SSI: Riggio, 1986), the Social Interpretations Test (Archer & Akert, 1977), and the Diagnostic Analysis of Nonverbal Accuracy Scale (DANVA: Nowicki & Duke, 1994). But, these do not focus on the recognition of discrete emotional states. Other tests focus more closely on emotion, such as the Communication and Reception of Affect Test (CARAT: Buck, 1976), the Test of Emotion Styles (TES: Allen & Hamsher, 1974), the Understanding our Feelings test (Elmore, 1985), the Feldstein Affect Judgment Test (Wolitzky, 1973), the Affective Communication Test (Friedman, Prince, Riggio, & DiMatteo, 1980), and the Contextual and Affective Sensitivity test (CAST: Trimboli & Walker, 1993). But, these are also questionable because of the lack of validity of the expressions used to portray emotion, the ability to produce specific scores on discrete emotions, or the lack of balance within the test to portray encoder characteristics (e.g., sex, race) equally.

The use of facial expressions of emotion that are universally recognized would address one concern. The data associated with expressions of anger, contempt, disgust, fear, happiness, sadness, and surprise provide suf-

ficient evidence of their external validity to portray accurately and reliably these discrete emotional states.

In fact, some studies have used these expressions as measures (e.g., Matsumoto, 1989, 1992). But, while they address some concerns, one artifact of their use is the high agreement level in judgments, which precludes the measurement of individual differences.

There are at least three ways to address the issue of high agreement levels: (1) reduce image size, (2) distort temporal and/or spatial resolution, or (3) increase presentation speed. Ekman, Brattesani, O'Sullivan, and Friesen (1979) explored the first method, using two cameras to videotape nurses during "honest" and "dishonest" interviews. One camera provided the "small face" condition in which the image size was one-fifth the area of a typical human face. The other camera provided the "large face" condition in which the image size was twice the area. Image size did not affect judgments about the nurses' affective states. Ekman et al. (1979) concluded "facial actions provide consistent information despite considerable size reduction" (p. 61).

Wallbott (1992) examined the second method, using a series of videotaped sequences developed by Scherer (1986) that depict 14 emotional states, and distorted either spatial resolution (pixel resolution—the number of points or squares constituting a video frame) or temporal resolution (refreshment rate—the number of frames transmitted per second). Although recognition rates decreased as distortion increased, most recognition rates still remained above chance levels. The stimuli used, however, did not meet independent criteria for validity.

The third method is to present the stimuli at such fast speeds that judgment accuracy is compromised. Ekman and Friesen's (1974) Brief Affect Recognition Task (BART) was created in this manner. It involves very brief (under 1/5s) presentations of facial stimuli, based on Ekman and Friesen's (1969) observation of micro-momentary expressions that occur almost outside of conscious awareness, and has been used to assess individual ERA (Ekman & Friesen, 1974; Mufson & Nowicki, 1991). One problem, however, is that facial physiognomy and poser sex are not balanced across emotions; another is the production of afterimages that affect judgments.

Matsumoto and Ekman's (1988) Japanese and Caucasian Facial Expressions of Emotion (JACFEE) addresses the limitation of the expressions used in Ekman and Friesen's (1974) BART, and improves on them in several ways. First, it includes equal numbers of posers of two visibly different ethnic groups, and of males and females within each group, for each of the seven universal emotions. Second, the faces were scored using Ekman and

Friesen's (1978) Facial Action Coding System (FACS; reliability = .91) to verify that the same expressions were shown across posers within each emotion, and that these are associated with universal emotions (Ekman & Friesen, 1975, 1986). Third, observers in multiple countries and cultures agree in their emotion judgments of the JACFEE expressions (Biehl et al., 1997). There is, therefore, ample support for the validity and reliability of these expressions. The next issue is how to alter their presentation to produce reliable individual difference scores.

The Development of the JACBART

We used Matsumoto and Ekman's (1988) JACFEE and Neutral Faces (JACNeuF) (consisting of neutral poses by the JACFEE posers) to develop a new test of ERA which we call the Japanese and Caucasian Brief Affect Recognition Test (JACBART). Items were created by embedding a JACFEE expression in the middle of a 1s presentation of that poser's JACNeuF expression on videotape. This format eliminated after-images of the target JACFEE expression. Items were placed in a random order, with the condition that the same emotion was not presented consecutively. There was a 3s inter-stimulus interval, with an orienting tone accompanied by a presentation number shown 1s prior to the item. This format was repeated for all 56 items.

Overview of the Studies Reported Here

We report five studies that explore the reliability and validity of the JACBART. Within the realm of reliability, two issues need to be considered—internal and temporal reliability. With 56 items measuring different aspects of expressions—emotion, poser race, and poser sex—and multiple items representing each aspect, items measuring each characteristic of the expressions must be internally consistent with each other, and consistent across time.

There are multiple concerns about validity. Face validity is assessed by the overall appearance of the test, and the use of the JACFEE and JACNeuF amply addresses this concern. Content validity concerns the lexical, logical, and methodological definitions of the construct (O'Sullivan, 1982), and is addressed by the structure of the JACBART, its rating scales and instructions. Two forms of construct validity—convergent and divergent—are demonstrated empirically. Convergent validity refers to the ability of

the JACBART to correlate with measures of the same or similar construct, or by intercorrelations among the different JACBART scale scores. Divergent validity requires that the test demonstrate that it is not perfectly correlated with an already existing test, and that it assesses an aspect of the construct that already existing tests do not assess.

Predictive validity refers to the ability of a test to accurately predict other constructs, and there are two types. Concurrent validity refers to the ability of the JACBART to predict scores on a different construct when the scores are gathered at the same time. Future predictive validity refers to its ability to predict scores on a different construct measured at a future time. A different version of predictive validity is incremental validity, which refers to its ability to predict a different construct above and beyond what is already predicted by other similar tests. In the studies reported below, predictive validity was assessed by examining correlations between the JACBART and widely used personality measures.

A final consideration is the need to demonstrate that the ERA scores generated by the JACBART are specific to the nature of emotion judgment, and not to general abilities related to taking such tests, including visual acuity, motivation, and the like.

Study 1 first examined what exposure duration produces the most internally reliable scores. We produced three versions, each differing solely in the length of exposure of the JACFEE expression, with one presenting the target at 1/15s, another at 2/15s, and a third at $1/5s^2$ (Versions 1, 2, and 3, respectively). These times were selected on the basis of previous work with the BART that suggested that they reduce the agreement levels for each expression for maximum item discrimination. The same random order was used on each tape.

Study 1 also addressed the concurrent validity of the JACBART with a five-factor personality scale. In Studies 1, 2, and 5, we used scales that operationalized the five-factor model of personality, considering that one limitation of previous research was the lack of consistent use of a standardized personality test. Study 2 examined the internal reliability of the JACBART using a simplified response format, and its concurrent validity with the same five-factor personality test used in Study 1. Study 3 tested its temporal reliability in a test-retest procedure. Study 4 examined its concurrent validity with a different personality scale that operationalized two of the five factors. Study 5 examined concurrent validity obtained in Studies 1, 2, and 4 with a different personality measure; it also examined the degree to which such evidence is confounded by individual differences in visual acuity not related to emotion judgment. All studies examined the internal reliability of the JACBART ERA scores, as well as one type of its

convergent validity via the intercorrelations among the JACBART scale scores.

Study 1

Method

Stimuli. JACBART Versions 1, 2, and 3 were used as stimuli in this study.

Judgment task: Multi-scalar ratings. For each item, participants rated the presence or absence of seven emotion terms—anger, contempt, disgust, fear, happiness, sadness, and surprise—by rating its intensity using a 9-point (0–8) scale labeled: Not at All (0), A Little (1), Moderate (4), and A Lot (8). Scores for each item for each subject were computed by correlating the 7 ratings with the means of the same seven ratings for that expression from the American data (n = 114) reported in Matsumoto (1986) and Matsumoto and Ekman (1989). Sixteen scores were then computed for each subject by averaging correlations within the following expression types: anger, contempt, disgust, fear, happiness, sadness, and surprise (8 expressions each); Caucasian and Japanese (28 expressions each); males and females (28 expressions each); Caucasian and Japanese males and females (14 expressions each); and a total score (all 56 expressions).

Personality measure. The Big Five Inventory-54 (BFI; John, 1989) was the measure of personality. Participants indicated whether each of 54 statements is true of themselves. Responses were averaged across items loading onto five scales: Extroversion, Neuroticism, Openness, Conscientiousness, and Agreeableness.

Participants and procedures. A total of 363 individuals participated in this study. All were university undergraduates participating in partial fulfillment of class requirements, and participated in small groups. Upon arrival to the laboratory, they completed the BFI, and then were randomly assigned to one of three groups that viewed JACBART Versions 1, 2, or 3. Group 1 included 103 participants (42 males, 61 females; mean age = 21.5; 40% Caucasian, 48% Asian, 12% other ethnicities); Group 2 included 114 participants (37 males, 77 females; mean age = 22.4, 31% Caucasian, 45% Asian, 24% others); and Group 3 included 146 participants (47 males, 99 females; mean age = 21.5, 28% Caucasian, 43% Asian, 29% others).[3]

The instructions were the same for all groups. Participants were told that they would be seeing a videotape containing facial expressions of emotion, and that they would rate each for the degree of anger, contempt, disgust, fear, happiness, sadness, and surprise displayed in the expression. They were provided with a sheet of definitions of these emotion words taken from a standard dictionary. They were told to not focus on the neutral face that started and ended each item, but on the target expression imbedded within the neutral. They were then given two examples of completed rating sheets, ensuring that they knew how to use them. Once any questions were answered and all participants understood their task, the videotape was started, and did not stop until the presentation of the last item. The experiment ended with completion of the ratings for the last item.

Results

Reliability analyses and descriptive statistics. Cronbach alphas computed for all sixteen scores (Table 1) were high and acceptable for all three tapes, with an increasing trend across tapes. Alphas were highest for the total recognition score. These findings provided ample support for the internal reliability of the ERA scores computed. Their means (Table 2) indicated that the degree of agreement was considerably lower than that usually obtained with judgments of still photographs. Variances also suggested considerable individual differences.

Concurrent analyses. Pearson correlations were computed between the ERA scores and the BFI scales, separately for each of the three video speeds, and for all three speeds combined.[4] These results were comparable to each other, and for parsimony we report the results only for all three speeds combined.[5] Total ERA was significantly correlated with Openness, $r(363) = .21$, $p < .001$; 14 of its remaining 15 correlations were also significant (range = .13 to .24; see Table 3). Total ERA was also significantly correlated with Conscientiousness, $r(363) = .11$, $p < .05$; 8 of its remaining 15 correlations were also significant (see Table 3). Thus, there was support for the concurrent validity of ERA and Openness and Conscientiousness.

The JACBART scores in general correlated well with each other (Table 4). The only emotion that did not correlate consistently well with the others was happiness. This may have been due to a ceiling effect, as evidenced by its means (Table 2). On the whole, however, the data indicated considerable support for the internal convergent validity of the test.

TABLE 1

Results of Reliability Analysis for JACBART Scores from All Studies

	Study 1 (Multi-scalar ratings)			Study 2 (Forced-choice)	Study 3, session 1 (Forced choice)	Study 3, session 2 (Forced choice)	Study 4 (Forced choice)	Study 5 (Forced choice)
Expression Type	1/15 s (n = 103)	2/15 s (n = 114)	1/5 s (n = 146)	1/5 s (n = 89)	1/5 s (n = 56)	1/5 s (n = 56)	2/15 s (n = 27)	1/5 s (n = 44)
Anger	.54	.75	.77	.71	.61	.66	.67	.72
Contempt	.66	.82	.81	.80	.79	.87	.45	.82
Disgust	.62	.75	.78	.65	.69	.67	.48	.58
Fear	.81	.78	.85	.68	.65	.68	.55	.69
Happiness	.64	.69	.79	.54	.76	.81	.72	.46
Sadness	.63	.71	.79	.69	.72	.73	.56	.48
Surprise	.78	.73	.81	.40	.63	.76	.71	.51
Caucasian	.76	.77	.84	.75	.80	.81	.83	.72
Japanese	.76	.76	.85	.59	.75	.77	.77	.78
Males	.78	.77	.84	.70	.73	.77	.75	.70
Females	.72	.77	.85	.66	.81	.80	.82	.81
Caucasian Males	.66	.62	.71	.57	.60	.64	.72	.43
Caucasian Females	.54	.60	.70	.58	.67	.68	.66	.63
Japanese Males	.55	.60	.70	.44	.54	.53	.52	.55
Japanese Females	.58	.59	.76	.32	.67	.66	.73	.71
Total	.86	.87	.92	.82	.87	.89	.90	.87

TABLE 2

Means and Standard Deviations for the Emotion Recognition Scores for all Studies

	Study 1 (Multi-Scalar Ratings)			Study 2 (Forced-Choice)	Study 3, Session 1 (Forced Choice)	Study 3, Session 2 (Forced Choice)	Study 4 (Forced Choice)	Study 5 (Forced Choice)
Expression Type	1/15 s (n = 103)	2/15 s (n = 114)	1/5 s (n = 146)	1/5 s (n = 89)	1/5 s (n = 56)	1/5 s (n = 56)	2/15 s (n = 27)	1/5 s (n = 44)
Anger	.30 (.21)	.48 (.26)	.60 (.23)	.64 (.27)	.67 (.15)	.73 (.12)	.30 (.15)	.68 (.26)
Contempt	.23 (.22)	.34 (.27)	.36 (.26)	.39 (.31)	.50 (.09)	.56 (.05)	.20 (.06)	.53 (.33)
Disgust	.47 (.23)	.59 (.23)	.62 (.23)	.64 (.25)	.70 (.11)	.74 (.08)	.29 (.17)	.69 (.23)
Fear	.34 (.30)	.57 (.25)	.62 (.27)	.58 (.26)	.60 (.16)	.65 (.09)	.34 (.16)	.59 (.27)
Happiness	.73 (.22)	.86 (.18)	.88 (.19)	.96 (.09)	.91 (.05)	.93 (.05)	.68 (.17)	.96 (.09)
Sadness	.24 (.24)	.40 (.28)	.59 (.29)	.64 (.25)	.70 (.18)	.77 (.12)	.29 (.16)	.74 (.20)

Surprise	.67	.83	.85	.93	.87	.89	.67	.90
	(.26)	(.16)	(.17)	(.11)	(.12)	(.06)	(.12)	(.14)
Caucasian	.44	.59	.65	.70	.75	.77	.41	.76
	(.16)	(.14)	(.16)	(.15)	(.16)	(.13)	(.23)	(.14)
Japanese	.42	.58	.64	.66	.70	.75	.36	.70
	(.16)	(.14)	(.17)	(.12)	(.19)	(.18)	(.25)	(.16)
Males	.44	.57	.63	.68	.72	.76	.40	.74
	(.16)	(.15)	(.17)	(.14)	(.18)	(.14)	(.25)	(.14)
Females	.41	.59	.66	.69	.72	.75	.38	.71
	(.15)	(.14)	(.16)	(.13)	(.19)	(.14)	(.23)	(.17)
Caucasian Males	.48	.59	.65	.71	.75	.78	.40	.78
	(.19)	(.16)	(.16)	(.16)	(.16)	(.12)	(.20)	(.14)
Caucasian Females	.40	.58	.66	.70	.74	.76	.41	.73
	(.16)	(.16)	(.16)	(.16)	(.17)	(.13)	(.25)	(.17)
Japanese Males	.41	.55	.61	.65	.68	.74	.40	.71
	(.17)	(.16)	(.18)	(.15)	(.20)	(.17)	(.28)	(.16)
Japanese Females	.43	.60	.67	.67	.69	.74	.33	.69
	(.17)	(.15)	(.18)	(.13)	(.21)	(.16)	(.63)	(.19)
Total	.43	.58	.65	.68	.72	.76	.38	.73
	(.15)	(.14)	(.16)	(.13)	(.18)	(.14)	(.24)	(.14)

TABLE 3

Significant Correlations Between ERA Scores and Personality Dimensions, Studies 1, 2, 4, and 5

Expression Type	Openness (Study 1)	Conscientiousness (Study 1)	Openness (Study 2)	Conscientiousness (Study 2)	Extraversion-Introversion (Study 4)	Neuroticism (Study 4)	Openness BFI (Study 5)	Conscientiousness BFI (Study 5)	Openness NeoPI-R (Study 5)	Conscientiousness NeoPI-R (Study 5)
Anger	.14**	.06	.13	.01	.37*	−.51**	.20	.17	.27+	.12
Contempt	.13**	.10*	.16+	.08	−.08	−.00	.12	−.09	−.07	−.20
Disgust	.19***	.11*	.38***	.06	.61***	−.35*	.38*	.41**	.50***	.39*
Fear	.19***	.06	.26**	.15+	.34*	−.08	.17	.36*	.23	.25
Happiness	.03	.05	−.06	.00	.60***	−.39*	.06	−.00	−.11	.01
Sadness	.13**	.05	.21*	.14	.29	−.51**	.17	.45**	.29+	.40**
Surprise	.18***	.08+	.30*	.12	.33*	−.36*	.30+	.38*	.38*	.21
Caucasian	.20***	.11*	.37***	.12	.52**	−.40*	.28+	.37*	.30+	.22
Japanese	.20***	.10*	.34**	.17	.28	−.42*	.30*	.31*	.34*	.23
Males	.24***	.12*	.39***	.12	.46**	−.45**	.27+	.38*	.38*	.23
Females	.17**	.09+	.32**	.16+	.48**	−.46**	.30*	.29+	.27+	.23
Caucasian Males	.22***	.09*	.35***	.10	.49**	−.28	.23	.38*	.27+	.16
Caucasian Females	.17**	.10*	.34**	.12	.41*	−.48**	.28+	.30+	.26+	.24
Japanese Males	.23***	.13**	.35***	.11	.11	−.39*	.26+	.32*	.40**	.24
Japanese Females	.15**	.06	.22*	.17+	.39*	−.37*	.28+	.24	.23	.19
Total	.21***	.11*	.38***	.15+	.47**	−.45**	.31*	.35*	.33*	.24

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.
All tests in Study 1 two-tailed; in subsequent studies one-tailed.

TABLE 4

Intercorrelations Among the JACBART Scores

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Anger	.12	.25**	.34***	.12	.30**	.26**	.52***	.50***	.54***	.49***	.50***	.43***	.50***	.44***	.54***
(1)	.22**	.33***	.39***	.33***	.16*	.53***	.63***	.72***	.69***	.65***	.61***	.54***	.66***	.68***	.71***
	.18*	.37***	.47***	.30***	.24**	.39***	.60***	.61***	.63***	.57***	.61***	.52***	.59***	.55***	.62***
	.29**	.48***	.08	.20*	.14	.20*	.63***	.61***	.63***	.61***	.59***	.56***	.54***	.51***	.67***
	.26*	.11	.13	.27*	.11	.33*	.54***	.43***	.46***	.51***	.45***	.53***	.32**	.41***	.54***
	.23*	.14	.26*	.23*	.15	.35*	.52***	.55***	.55***	.52***	.47***	.46***	.55***	.47***	.58***
	.46	.72**	.59*	.64**	.67**	.60*	.84***	.75***	.82***	.77***	.85***	.69**	.67**	.71**	.82***
	.38**	.35**	.46***	.11	.34*	.34*	.71***	.63***	.63***	.68***	.58***	.69***	.56***	.58***	.70***
	.23	.32	.33	.25	.22	.36	.60	.60	.61	.58	.57	.53	.55	.54	.63
Con-	—	.36***	.28**	.06	.01	.31**	.48***	.44***	.47***	.46***	.44***	.43***	.44***	.39***	.49***
tempt		.41***	.21*	.11	.04	.24**	.56***	.52***	.53***	.54***	.51***	.51***	.48***	.48***	.56***
(2)		.36***	.30***	.19*	.10	.25**	.52***	.49***	.52***	.49***	.48***	.51***	.52***	.41***	.52***
		.22*	.04	.16	.04	.04	.52***	.52***	.58***	.46***	.54***	.42***	.49***	.39***	.56***
		.25*	.05	.02	.23*	.32**	.53***	.47***	.52***	.49***	.44***	.54***	.48***	.36**	.53***
		.22	.07	.03	.14	.12	.51***	.45***	.48***	.50***	.44***	.51***	.47***	.41***	.49***
		.26	.46	.34	.28	.42	.55*	.53*	.40	.63**	.42	.58*	.48	.55*	.51
		.26*	.35**	.02	.25	.34*	.67***	.64***	.68***	.61***	.68***	.55***	.56***	.59***	.68***
		.32	.21	.11	.10	.23	.53	.50	.53	.50	.49	.49	.49	.43	.54
Disgust		—	.45***	.12	.20*	.53***	.64***	.64***	.65***	.64***	.61***	.54***	.60***	.60***	.68***
(3)			.37***	.18*	.15	.51***	.70***	.64***	.67***	.67***	.65***	.64***	.58***	.61***	.70***
			.51***	.34***	.34***	.50***	.69***	.72***	.68***	.73***	.62***	.68***	.66***	.70***	.72***
			.15	.08	.30**	.19*	.63***	.64***	.65***	.62***	.57***	.59***	.61***	.48***	.68***
			.34**	.44***	.52***	.50***	.70***	.66***	.60***	.70***	.59***	.67***	.51***	.63***	.72***
			.39***	.49***	.45***	.42***	.70***	.63***	.68***	.65***	.64***	.65***	.61***	.54***	.70***
			.50	.63**	.37	.49*	.70**	.49	.76***	.62**	.68**	.57*	.51*	.51*	.69**
			.35**	.11	.41**	.45***	.62***	.64***	.64***	.60***	.55***	.57***	.62***	.55***	.66***
			.39	.25	.31	.45	.67	.66	.66	.67	.61	.62	.60	.60	.70

TABLE 4 (*Continued*)

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Fear (4)			—	.32***	.22*	.67***	.75***	.77***	.76***	.78***	.65***	.70***	.76***	.68***	.80***
				.21*	.14	.47***	.63***	.65***	.64***	.64***	.51***	.64***	.66***	.54***	.67***
				.49***	.43***	.54***	.75***	.80***	.77***	.78***	.71***	.72***	.76***	.76***	.80***
				.05	.28**	.06	.46***	.43***	.44***	.45***	.44***	.40***	.34***	.40***	.48***
				.36**	.44***	.37**	.49***	.64***	.55***	.57***	.51***	.40**	.47***	.64***	.61***
				.38**	.57***	.42***	.61***	.68***	.61***	.69***	.60***	.55***	.54***	.72***	.67***
				.45	.62*	.52*	.80***	.73**	.82***	.65**	.71**	.74***	.82***	.47	.79***
				.14	.63***	.38**	.69***	.74***	.70***	.70***	.56***	.69***	.71***	.62***	.75***
				.30	.35	.44	.65	.68	.66	.67	.59	.61	.63	.62	.70
Happi-				—	.03	.34***	.44***	.40***	.47***	.38***	.48***	.29**	.38***	.37***	.45***
ness (5)					.04	.29**	.46***	.38***	.45***	.39***	.53***	.31***	.29**	.41***	.44***
					.35***	.70***	.64***	.66***	.66***	.64***	.66***	.56***	.59***	.65***	.67***
					.14	.29**	.28**	.37***	.30**	.34***	.24*	.28**	.30**	.32***	.34***
					.33**	.59***	.61***	.54***	.56***	.60***	.55***	.57***	.40***	.54***	.62***
					.47***	.59***	.60***	.58***	.59***	.61***	.62***	.51***	.49***	.60***	.62***
					.64**	.71***	.90***	.80***	.87***	.85***	.79***	.78***	.67**	.71***	.90***
					.16	.28*	.05	.19	.05	.18	.08	.02	.01	.30*	.13
					.21	.46	.48	.48	.48	.47	.50	.39	.39	.48	.50
Sadness					—	.25**	.38***	.49***	.41***	.47***	.35***	.33***	.42***	.50***	.46***
(6)						.00	.39***	.40***	.34***	.45***	.26**	.46***	.36***	.38***	.41***
						.49***	.66***	.60***	.61***	.65***	.62***	.63***	.54***	.60***	.65***
						.23*	.56***	.47***	.48***	.57***	.46***	.58***	.41***	.40***	.56***
						.44***	.63***	.66***	.67***	.61***	.64***	.53***	.56***	.60***	.69***
						.49***	.68***	.67***	.73***	.65***	.68***	.60***	.67***	.58***	.70***
						.64**	.77***	.83***	.68**	.85***	.71***	.70**	.61*	.83***	.80***
						.47***	.64***	.72***	.63***	.70***	.52***	.63***	.62***	.67***	.71***
						.32	.55	.55	.53	.58	.49	.53	.49	.52	.58

Surprise	—	.78***	.75***	.76***	.78***	.70***	.70***	.72***	.68***	.81***
(7)		.62***	.68***	.69***	.59***	.63***	.50***	.65***	.60***	.68***
		.76***	.76***	.75***	.76***	.74***	.71***	.69***	.74***	.78***
		.35***	.33***	.28**	.41***	.20*	.44***	.30**	.27**	.37***
		.71***	.71***	.70***	.69***	.67***	.62***	.59***	.65***	.75***
		.67***	.65***	.65***	.70***	.60***	.66***	.62***	.61***	.71***
		.83***	.84***	.76***	.86***	.72***	.79***	.69**	.80***	.84***
		.65***	.63***	.57***	.68***	.60***	.58***	.44***	.68***	.67***
		.66	.66	.65	.67	.61	.61	.60	.61	.69
Cauca-		—	.78***	.91***	.88***	.91***	.88***	.77***	.69***	.94***
sian (8)			.84***	.91***	.92***	.91***	.91***	.77***	.79***	.96***
			.90***	.94***	.95***	.95***	.95***	.84***	.86***	.97***
			.72***	.88***	.87***	.91***	.92***	.66***	.58***	.94***
			.77***	.87***	.90***	.92***	.93***	.61***	.73***	.94***
			.80***	.92***	.9l***	.93***	.94***	.77***	.72***	.95***
			.90***	.97***	.93***	.94***	.91***	.83***	.79***	.98***
			.85***	.89***	.91***	.89***	.93***	.74***	.78***	.95***
			.82	.91	.91	.92	.92	.76	.75	.95
Japanese			—	.89***	.91***	.73***	.66***	.94***	.94***	.94***
(9)				.92***	.91***	.76***	.77***	.93***	.92***	.96***
				.95***	.94***	.87***	.84***	.95***	.94***	.98***
				.85***	.85***	.65***	.68***	.88***	.83***	.91***
				.87***	.87***	.69***	.72***	.86***	.89***	.94***
				.90***	.92***	.75***	.74***	.93***	.94***	.94***
				.87***	.97***	.80***	.88***	.84***	.91***	.97***
				.90***	.92***	.75***	.79***	.88***	.92**	.97***
				.90	.91	.76	.75	.92	.91	.95
Males				—	.82***	.94***	.68***	.92***	.76***	.96***
(10)					.82***	.92***	.75***	.92***	.77***	.96***
					.89***	.95***	.84***	.95***	.85***	.97***
					.73***	.90***	.72***	.89***	.55***	.93***

TABLE 4 (*Continued*)

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
										.74***	.89***	.72***	.88***	.66***	.92***
										.85***	.93***	.80***	.92***	.76***	.96***
										.86***	.94***	.90***	.91***	.65**	.96***
										.78***	.89***	.74***	.92***	.72***	.93***
										.82	.92	.76	.92	.74	.95
Females										—	.72***	.88***	.80***	.90***	.95***
(11)											.74***	.93***	.76***	.92***	.95***
											.86***	.95***	.83***	.95***	.97***
											.69***	.91***	.61***	.85***	.93***
											.73***	.92***	.58***	.93***	.94***
											.79***	.91***	.80***	.91***	.97***
											.85***	.92***	.74**	.93***	.97***
											.72***	.92***	.70***	.94***	.96***
											.76	.92	.75	.92	.95
Cauca-											—	.61***	.71***	.66***	.87***
sian												.66***	.69***	.72***	.87***
Males												.81***	.81***	.83***	.93***
(12)												.68***	.60***	.51***	.86***
												.71***	.57***	.64***	.86***
												.76***	.72***	.69***	.89***
												.72**	.72**	.73***	.93***
												.64***	.65***	.69***	.84***
												.70	.70	.69	.88
Cauca-												—	.66***	.58***	.81***
sian													.71***	.71***	.88***
Females													.79***	.80***	.92***
(13)													.60***	.55***	.87***
													.54***	.70***	.89***

							.72***	.67***	.89***
							.86***	.72***	.94***
							.69***	.73***	.89***
							.69	.68	.88
Japanese							—	.76***	.90***
Males								.70***	.88***
(14)								.79***	.92***
								.47***	.81***
								.53***	.77***
								.74***	.90***
								.55*	.86***
								.62***	.85***
								.68	.87
Japanese								—	.87***
Females									.89***
(15)									.93***
									.75***
									.56***
									.87***
									.84***
									.89***
									.85

$p < .05$, $^{**}p < .01$, $^{***}p < .001$.
1st entry—Study 1, Version 1.
2nd entry—Study 1, Version 2.
3rd entry—Study 1, Version 3.
4th entry—Study 2.
5th entry—Study 3, Session 1.
6th entry—Study 3, Session 2.
7th entry—Study 4.
8th entry—Study 5.
9th entry—Weighted average of the eight entries.

Additional analyses: Videotape differences. We averaged each subject's score across both expressions of the same poser type (i.e., Caucasian and Japanese males and females) for each emotion, and computed a five-way Analysis of Variance (ANOVA) using participant sex and videotape as between-subject factors, and emotion, poser race, and poser sex as within-subject factors. The main effect for videotape was significant, $F(2, 307) = 45.38$, $p < .001$, and follow-up comparisons indicated that recognition scores for Version 2 were significantly higher than for Version 1, but significantly lower than for Version 3, $F(1, 215) = 63.65$, $p < .001$; and $F(1, 258) = 12.36$, $p < .01$, respectively. The videotape by emotion interaction was also significant, $F(12, 1842) = 3.63$, $p < .001$. Analytic comparisons indicated that Version 2 produced higher recognition scores than Version 1 for all emotions; Version 3, however, produced significantly higher recognition scores than Version 2 on only anger and sadness.[6]

A significant poser sex by videotape interaction, $F(2, 307) = 11.26$, $p < .001$, led to analytic comparisons that showed that female expressions had higher recognition scores than males on Versions 1 and 3, $F(1, 102) = 8.83$, $p < .01$; and $F(1, 145) = 22.34$, $p < .01$, but there were no differences on Version 2.

Additional analyses: Emotion differences. A significant emotion main effect, $F(6, 1842) = 217.23$, $p < .001$, was analyzed by ranking the emotions according to their mean accuracy and testing differences between adjacent pairs, combining non-significant emotions, which produced the following results: Happiness > Surprise > Disgust > Fear = Anger = Sadness > Contempt.

Discussion

All versions produced alpha coefficients for all scores well in the acceptable range; the alphas for Version 3 were generally the highest. The means indicated that Version 3 produced more desirable item discrimination, and the variances were indicative of acceptable individual differences. There were, however, some differences in recognition scores as a function of emotion and poser sex. Also, the recognition scores for happiness and surprise were surprisingly high. These differences suggest that a different presentation speed may be necessary for some emotions (e.g., faster speeds for happiness) or some poser characteristics (e.g., male or female faces). This, however, poses its own problems. If happy faces were shown at a faster speed than other faces, they may literally be outside of conscious awareness. There would also be a practice effect, as items presented at

slower speeds will appear even slower, which would bias the recognition scores. The altering of one or two emotions or poser characteristics may unintentionally produce an association between items and response alternatives in judgment tasks, introducing yet another bias. Because of these limitations, we conclude that standardized item presentation speeds are optimal, and that Versions 2 and 3 were optimal in assessing ERA.

That ERA was correlated with the Openness and Conscientiousness provided some evidence for the concurrent validity of the JACBART. People who score high on Openness are generally original, introspective, aesthetically sensitive, intellectually curious, attentive to inner feelings; they value intellectual matters, and have an active imagination. Thus, they are more attentive and receptive to their environment and the people around them, and will be more in tune with reading emotions of others. People who score high on Conscientiousness are generally thorough, reliable, and perseverant. They may attend to details better than others, including details concerning facial expressions, which would explain the obtained correlations. Alternatively, the finding may reflect differences in effort exerted during the experimental task. That these findings were obtained across the different ERA scores suggests that these correlations are not emotion-specific, and instead may be part of a general ability to recognize emotions.

The intercorrelations provided evidence for the convergent validity of the JACBART scores. And, the differences among the emotions were generally the same as found in judgment studies using full face presentations of still photographs of these same expressions, despite the substantial change in the nature of the judgment task. There are several possibilities that explain these findings. First, the physical features associated with the expressions may make judgments of some easier than others. Hager and Ekman (1979), for example, found that happy expressions were easiest to judge at further distances than were the other emotions. The findings also suggest possible neural circuitry that allows for easier discrimination of some emotions than others. Given that happiness and surprise were the only non-negative emotions, the findings may also reflect discrimination along a positive-negative dimension, which may be related to Darwin's (1872) principle of antithesis.

One limitation of Study 1 was the multi-scalar ratings, which require considerable effort by the observers. In addition, the findings need to be replicated. Thus, the goal of Study 2 was to replicate the findings from Study 1: internal reliability of the recognition scores, concurrent validity with the BFI, convergent validity via intercorrelations, and emotion differences in recognition scores, while at the same time extending the findings of Study 1 by using a simpler judgment task.

Study 2

Method

Stimuli and personality measure. JACBART Version 3 and the BFI-54 were used in this study.

Judgment task: Forced-choice judgments. For each item, participants were given a list of seven emotion words—anger, contempt, disgust, fear, happiness, sadness, and surprise—and chose one that best represented the emotion portrayed. The nominal judgments were converted into dichotomous accuracy scores by recoding them to "1" if it was the emotion term intended for that expression, and "0" for all other terms selected. The same 16 scores described in Study 1 were computed by averaging across expressions.

Participants and procedures. Participants were 89 university undergraduates participating in partial fulfillment of class requirements (67 female, 22 male, mean age = 22.1). All procedures for data collection were exactly the same as in Study 1.

Results

Reliability analyses and descriptive statistics. Kuder-Richardson 20s (Table 1) were generally lower compared to the alphas reported in Study 1, but they were acceptable. In particular, the total emotion recognition score was still high. Descriptive statistics (Table 2) indicated that the emotions were recognized well above chance (1/7), and in the case of happiness and surprise, were almost comparable to judgments of still photographs. With the exception of these two emotions, there was considerable individual variation in the degree of emotion recognition, which was desired.

Concurrent analyses. Total ERA was again positively correlated with Openness, $r(89) = .38$, $p < .001$; 12 of the remaining 15 correlations with this dimension were also significant. Total ERA was marginally correlated with conscientiousness, $r(89) = .15$, $p < .10$. Total recognition was not correlated with any other BFI scale (Table 3). The intercorrelations among the JACBART scores (Table 4) again showed considerable evidence for convergence among the scores.

Emotion differences. We summed the recoded dichotomous accuracy scores (0–1) within each poser type (i.e., Caucasian and Japanese males

and females) separately for each emotion and computed a four-factor ANOVA, using judge sex as a between-subject factor, and emotion, poser race, and poser sex as within-subject factors. A significant emotion main effect, $F(6, 522) = 57.34$, $p < .001$, was analyzed by ranking the emotions according to their mean accuracy, and testing differences between adjacent pairs, which produced the following results: Happiness > Surprise > Disgust = Sadness = Anger = Fear > Contempt.

Discussion

The results of Study 2 confirmed the reliability of the JACBART and its ability to produce meaningful individual difference scores of ERA. The lower reliability estimates than Study 1 were most likely due to the decreased discrimination in the response task. They may also be due to differences in the two reliability statistics. With the exception of happiness and surprise (which were the same exceptions in Study 1), the descriptive statistics were well within the desirable range for item discrimination.

Study 2 also replicated the findings of Study 1 with regard to convergence with the two personality scales, and the intercorrelations among the JACBART scores. Given the internal reliability of the JACBART to produce meaningful individual difference scores of ERA using a single choice response format, Study 3 examined its temporal reliability using the same format.

Study 3

Method

Participants were 56 university undergraduates participating in partial fulfillment of class requirements (42 female, 14 male, mean age = 23.4). All procedures for data collection were exactly the same as in Study 2, with the exception that participants were tested twice, with the testing sessions separated by 3 to 4 weeks. The scoring procedures were exactly the same as those in Study 2, with 16 ERA scores produced for both sessions.

Results

Alphas were computed for all sixteen scores from both sessions (Table 1), and were again generally high and acceptable; they were again highest for the total score. The intercorrelations replicate the findings of Studies 1 and 2 (Table 4). Pearson correlations between the 16 ERA scores for ses-

sions 1 and 2 (Table 5) were positive, high, and statistically significant, providing convincing evidence for the temporal reliability of the scores.

One curious and unexpected trend in the data was that all sixteen scores in Session 2 were higher than in Session 1 (Table 2). We tested each pair of scores using a paired-sample t-test. The ts for fear, sadness, males, Japanese, Caucasian males, Japanese males, Japanese females, and total scores were significant, $t(54) = 1.97$, $p < .06$; $t(54) = 2.94$, $p < .01$; $t(51) = 3.01$, $p < .01$; $t(52) = 3.17$, $p < .01$; $t(53) = 1.91$, $p < .06$; $t(52) = 2.96$, $p < .01$; $t(54) = 2.27$, $p < .05$; and $t(50) = 3.06$, $p < .01$, respectively. These results suggested the existence of a practice effect for the test.

TABLE 5

Test-Retest Correlations for the 16 ERA Scores Generated from the JACBART, Study 3

Expression Type	Test-Retest Correlation
Anger	.44***
Contempt	.61***
Disgust	.58***
Fear	.62***
Happiness	.66***
Sadness	.72***
Surprise	.53***
Caucasian	.75***
Japanese	.67***
Males	.73***
Females	.70***
Caucasian Males	.66***
Caucasian Females	.68***
Japanese Males	.62***
Japanese Females	.57***
Total	.78***

***$p < .001$.

Discussion

The findings from Study 3 again replicated the internal reliability and convergent validity of the JACBART, and demonstrated its test-retest reliability.

While Studies 1 and 2 established the correlation between ERA as measured by the JACBART and Openness and Conscientiousness, the correlations between ERA and extraversion and neuroticism were not significant. In order to examine whether these non-findings were limited to the personality test used, Study 4 examined the relationship between ERA and extraversion and neuroticism measured by a different scale.

Study 4

Method

Personality measure. The Eysenck Personality Inventory (EPI) was used as a measure of personality. It contains 57 items that assess extraversion-introversion and neuroticism (24 items each), as well as a Lie scale (9 items). The test was administered and scored according to the standard procedures (Eysenck & Eysenck, 1968).

Participants, stimuli, and procedures. The participants were 27 university undergraduates, participating as volunteers in a course activity (16 females, 11 males). Participants completed the EPI prior to viewing the stimuli, and were tested in small groups. JACBART Version 2 was used in this study. All procedures concerning the collection of the judgment data using the JACBART were exactly the same as those used in Study 2 with forced-choice response formats. Sixteen scores were again generated.

Results

Alphas were again generally high and acceptable, and again highest for the total score (Table 1). Product moment correlations were computed between Extraversion-introversion and neuroticism and the sixteen ERA scores (Table 3). Extraversion-introversion was positively correlated with total ERA, as well as a number of other ERA scores. Neuroticism was significantly negative correlated with total ERA, as well as other ERA scores. The two personality scales were not correlated with each other, $r(25) = .009$, ns, and thus could not account for the correlations obtained with ERA. The

intercorrelations among the JACBART scores (Table 4) again replicated the same results from the previous studies.

Discussion

Study 4 again replicated the internal reliability and convergent validity of the JACBART, and demonstrated its concurrent validity with a different personality scale. The fact that ERA was positively correlated with extraversion suggests that individuals who are more sociable and outgoing are more in tune with the emotions of others, and can better read and interpret them. At the same time, the negative correlation between ERA and neuroticism suggests that those individuals who are more emotionally labile do worse at interpreting others' emotions. This may occur because they are more in tune with their own emotional states, and cannot attend to the emotions of others as well as those who are not as emotionally labile.

It is curious that these correlations were obtained with the Eysenck scale but not the BFI. One possibility is that the correlations may be limited to the specific scale used to measure personality. To examine this possibility, Study 5 included another personality scale that measured all five personality factors. Study 5 also addressed the possible confounding of the JACBART scores with the general ability to perceive visual stimuli presented at such high speeds. It may be the case, for example, that the correlations obtained so far are indicative not necessarily of the ability to judge emotion, but instead of one's visual acuity. This concern was addressed by the incorporation of a measure of visual acuity with presentation parameters comparable to the JACBART.

Study 5

Method

Personality measures. In addition to the BFI-54, we also used the Revised Neo Personality Inventory (NeoPI-R; Costa & McCrae, 1992), which is a 240 item measure of the five major dimensions of personality and some of the more important traits or facets that define each domain. The measure has 30 facet scales which, when summed, generate the 5 domain (global) scales. Participants rate their responses to each item using a 5-point Likert scale labeled from *strongly agree (0)* to *strongly disagree (4).* For both scales, scores on the big five personality constructs were scored using standard procedures.

ERA and visual acuity. JACBART Version 3 was again used to assess ERA, with forced-choice response formats; sixteen ERA scores were generated. In addition, we used Ekman's Facial Identification Test (FIT), which was designed by Paul Ekman a number of years earlier in order to control for the ability to recognize non-emotional facial expressions. It includes 18 items in which a neutral expression is presented for 1s, and a target expression of the same poser is imbedded within the 1s neutral presentation. Like the JACBART Version 3, the target expression was presented for 1/5s. Unlike the JACBART, however, the FIT target expressions were of one of three non-emotional faces: eyes open, mouth open; eyes closed, mouth closed; eyes closed, mouth open. Observers are presented with these three response alternatives for each item, and select the expression they believed was portrayed. Responses are scored as correct or incorrect, and the total number of correct responses is the observer's score (range = 0–18). The FIT was almost exclusively used as a control measure in pilot research in Ekman's laboratory; no published data are available concerning its development.

Participants and procedures. The participants were 44 university undergraduates, participating as volunteers in a course activity (32 females, 12 males). Participants were tested in small groups, and upon arrival at the laboratory, viewed either the JACBART or FIT in a randomly chosen order. After completion of both judgment tasks, they then completed the BFI and NeoPI-R, also in a randomly chosen order.

Results

The reliability data, descriptive statistics, and intercorrelations again supported the internal reliability and convergent validity of the JACBART scores (Tables 1, 2, and 4). Partial correlations were computed between the JACBART scores and the big five scales, separately for the BFI and NeoPI-R, partialling the individual's FIT scores (Table 3). Total ERA was again correlated with Openness on both scales, despite the fact that individual differences in visual acuity were controlled for. In addition, total ERA was correlated with BFI Conscientiousness; it was positive and of moderate value for NeoPI-R Conscientiousness as well. These findings replicate and extend the findings from Studies 1 and 2, providing stronger support for the convergent validity of the JACBART. Total ERA was also correlated with BFI Agreeableness; this correlation was not significant, however, on the NeoPI-R, nor was it found in the previous studies. We opt, therefore, not to interpret this finding as reliable. ERA was not significantly correlated with either

extraversion or neuroticism, suggesting that the findings obtained in Study 4 may be specific to the EPI.

General Discussion

All studies provided strong evidence for the internal reliability of the JACBART, as well as its convergent validity through its intercorrelations. Study 3 provided strong evidence for the temporal reliability of the scores, while Studies 1, 2, and 5 provided strong evidence for its concurrent validity with Openness, and moderate evidence for its concurrent validity with Conscientiousness. Study 4 also provided preliminary support for its concurrent validity with extraversion and neuroticism, but specific to one scale. Study 5 demonstrated that the validity coefficients between JACBART scores and the personality measures were not confounded by individual differences in visual acuity.

These are the first findings in support of a valid and reliable measure of ERA. These data, combined with the external validity associated with the JACFEE expressions used in the JACBART from previous judgment studies and FACS coding, and with the balanced poser race and sex design of the JACFEE, allows the JACBART to improve on limitations of other existing tests. We recommend the use of JACBART Version 3 with multi-scalar ratings, as these produce the best internal reliability statistics. But, these ratings are cumbersome, and should they be unwieldy, the forced-choice judgment task is much more user-friendly. And, there is no sacrifice in terms of reliability or validity associated with its use.

To our knowledge, the findings reported are also the first published evidence of a correlation between a psychometrically sound measure of ERA and standard personality tests across multiple studies. That ERA was significantly correlated with these traits in three studies involving two different measures suggests clearly that the ability to recognize emotions in others is a reliable correlate of these personality dimensions. As mentioned earlier, individuals who score high on Openness are interested and curious, and are receptive to external stimuli. Apparently, these stimuli include characteristics related to other people's expressions of feelings and emotions. Likewise, individuals who score high on Conscientiousness are reliable, efficient, and attentive to detail. The findings from these studies suggest that one of the details they attend to is facial expressions of emotion.

To be sure, there was some range in the size of the correlations across studies. We interpret the variance in the size of the correlations to be asso-

ciated with sampling error. For instance, the correlations between total ERA and Openness in Studies 2 and 5 were .38 and .30, respectively (BFI only). The same correlation in Study 1 was .21. Studies 2 and 5 were associated with sample sizes of 89 and 44, while the sample size in Study 1 was 363. Clearly, the larger sample sizes introduce greater error into the data set, which reduces the absolute size of the correlations computed. This same trend was observed with the Conscientiousness correlations, and with correlations using other ERA scores.

The pattern of correlations between ERA and Openness and Conscientiousness provide some conflicting views of the emotion-specificity of the correlations. On one hand, in Study 1, the correlations with Openness were significant for all emotions except happiness, suggesting non-specificity. Also, in Study 2, these correlations were significant for disgust, fear, sadness, and surprise, and was marginally significant for contempt, again suggesting some degree of non-specificity. In Study 5, however, this correlation was only significant with disgust. Future studies, therefore, will need to explore the possible emotion-specificity of these correlations more fully. These studies will not be without considerable theoretical import. Non-specificity, on one hand, would suggest the existence of a general ERA, much like an intelligence "g" score. Such a concept would have major implications for future models of emotion and personality. If such a construct existed, subsequent studies can further explore its possibilities, and the implications it has for development, other personality and psychological correlates, social interactions, and its biological substrates. Some research using the JACBART, for instance, has already documented its predictive ability to detect lies (Frank & Ekman, 1997). Specificity, on the other hand, would implicate a special role for specific emotions in either interpersonal relationships or personality construction. In either case, the potential ramifications are interesting and provocative.

We had also suggested in the introduction that ERA was related to extraversion and neuroticism. The results from Studies 1, 2, and 5, however, indicated that there was no correlation between these constructs. That the correlation was obtained only in Study 4, which used the EPI, suggests that there may be something about these dimensions measured by the EPI and *not* measured by the BFI-54 or NeoPI-R that may be related to ERA. Or, the findings from Study 4 may be due to sampling error. Clearly, this finding needs to be replicated. If replicated, then conceptual work needs to occur to flush out some of the theoretical differences between the extraversion and neuroticism measured in the EPI as opposed to the BFI or NeoPI-R.

The five studies reported here do not address all of the reliability and

validity concerns of the JACBART. Future studies will need to examine the predictive validity of the JACBART with other measures of personality, and personality constructs other than those measured in the studies reported here. Such research will need to occur within a more refined theoretical framework of the relationship between ERA and personality, which was premature here. Also, future studies will need to address the relationship between JACBART and other currently existing measures of ERA and decoding ability. One issue, of course, concerns the degree to which JACBART provides predictive ability above and beyond other currently existing measures, regardless of the specificity of emotion recognition that is assessed. A final methodological issue concerns the addition of faces of people of various ethnic and racial backgrounds, not just the two presented by the JACBART.

Future researchers will also need to deal with concerns about the ecological validity of the JACBART. Given that the JACBART focuses on the judgments of facial expressions of emotion in a laboratory setting, the degree to which ERA scores derived from such an administration are variable in multiple contexts—with other channels of communication and other contextual cues available—is an empirical question that needs to be addressed. It may be that the JACBART captures much of the variance in full context communication; or, it may be that the JACBART captures only a portion of such variance. A related issue concerns the relationship between the JACBART scores and real-life behaviors, and the actual emotion recognition schemes and abilities that are used in everyday life.

The studies reported here, however, do provide considerable evidence for the ability of the JACBART to reliably and validly measure individual differences in ERA. The availability of this test suggests a whole host of new and exciting research opportunities investigating its nature and scope, including the neural circuitry associated with its function and the contexts and factors that may influence it. Future research, for instance, may examine characteristics of the posers—such as race, sex, hair style and color, facial physiognomy and morphology, skin color or tone, and the like—that may influence ERA, as well as characteristics of the observers—such as emotional or mood state, culture, sex, facility with making emotion judgments, and the like. These possibilities also open the door to studies that investigate the possibility of the relationship between ERA of specific emotions with specific personality, mood, and psychopathologies (e.g., ERA of sadness with depressed individuals, of anger in hostile individuals, etc.). The possibility of detecting an ERA g factor is an intriguing and interesting one that deserves considerable attention. The ability of facial expression modeling via computer graphics affords researchers with the tools neces-

sary to alter systematically and objectively characteristics of faces that will be useful for future studies. That ERA is an important component of many personality constructs is a concept that has been suggested for years; now, researchers can have a valid tool with which to pursue those ideas in a scientifically sound fashion.

Notes

1. Many previous studies examining individual differences in decoding ability have generated such measures by asking participants to decode their own or other's spontaneous expressions obtained as part of the research protocol. We chose not to include these adhoc methods in this review, electing instead to focus on formal attempts to develop objective measurement techniques of the construct.
2. These times are somewhat affected by technological parameters of the videotape. Because videotape creates 30 image fields per second, the fastest exposure duration we could accomplish would have been to lay a JACFEE expression on a single field, resulting in a 1/30 s presentation. Our pilot work, however, had indicated that this presentation time was extremely fast, and judges reported that they had absolutely no confidence in their judgments. At two fields per second, that is 1/15 s, judges report being able to see (barely) the expressions. Thus, with this as a baseline speed, we decided to create two more stimulus tapes at equal increments, that is, 2/15 s and 3/15 = 1/5 s durations.
3. All reliability and validity analyses reported below were also conducted separately for Caucasians and Asians, as these were the two ethnicities that had adequate sample sizes for reliable analyses, and for males and females. All results replicated the results for the entire group, and are available upon request. Thus, no further mention of ethnicity or gender will be made in the main text.
4. We decided to conduct an aggregated analysis for all three speeds based on the comparability of the alphas reported in Table 1.
5. A detailed report of all analyses, including separate analyses for males and females, can be obtained from the first author.
6. A detailed table of these findings is available from the first author.

References

Allen, J. R., & Hamsher, J. H. (1974). The development and validation of a test of emotional styles. *Journal of Consulting and Clinical Psychology, 42,* 663–668.

Archer, D., & Akert, R. (1977). Words and everything else: Verbal and nonverbal cues in social interaction. *Journal of Personality and Social Psychology, 35,* 443–449.

Biehl, M., Matsumoto, D., Ekman, P., Hearn, V., Heider, K., Kudoh, T., & Ton, V. (1997). Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. *Journal of Nonverbal Behavior, 21* (1), 3–22.

Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology, Vol 2.* (pp.634–654). Reading, MA: Addison-Wesley.

Buck, R. W. (1976). A test of nonverbal receiving ability: Preliminary studies. *Human Communication Research, 2,* 162–171.

Buck, R. W., Savin, V. J., Miller, R. E., & Caul, W. F. (1972). Communication of affect through facial expressions in humans. *Journal of Personality and Social Psychology, 23* (3), 362–371.

Costa, P. T., & McCrae, R. R. Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*, 5–13.

Cunningham, M. R. (1977). Personality and the structure of the nonverbal communication of emotion. *Journal of Personality, 45,* 564–584.

Darwin, C. (1872). *The expression of emotions in man and animals.* New York: Philosophical Library.

Ekman, P., Brattesani, K. A., O'Sullivan, M., & Friesen, W. V. (1979). *Journal of Nonverbal Behavior, 4* (1), 57–61.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32* (1), 88–106.

Ekman, P., & Friesen, W. V. (1974). Nonverbal behavior and psychopathology. In R. J. Friedman & M. Katz (Eds.), *The psychology of depression: Contemporary theory and research* (pp. 3–31). Washington, DC: Winston & Sons.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face.* 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. (1986). A new pan-cultural expression of emotion. *Motivation and Emotion, 10,* 159– 168.

Elmore, B. C. (1985). Emotionally handicapped comprehension of nonverbal communication. *Journal of Holistic Medicine, 7,* 194–201.

Eysenck, S. B., & Eysenck, H. J. (1968). The measurement of psychoticism: A study of factor stability and reliability. *British Journal of Social and Clinical Psychology, 7* (4), 586–294.

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

Friedman, H. S., Prince, L. M., Riggio, R. E., DiMatteo, M. R. (1980). Understanding and assessing nonverbal expressiveness: The Affective Communication Test. *Journal of Personality and Social Psychology, 39* (2), 333–351.

Hager, J. C., & Ekman, P. (1979). Long distance transmission of facial affect signals. *Ethology and Sociobiology, 1*, 77–82.

John, O. (1989). *The BFI-54.* Unpublished test, Institute of Personality and Social Research, Department of Psychology, University of California, Berkeley.

Lanzetta, J. T., & Kleck, R. E. (1970). Encoding and decoding of nonverbal affect in humans. *Journal of Personality and Social Psychology, 16,* 12–19.

LeRoux, J. A. (1987). *Personality and the encoding and decoding of facial expressions of emotion.* Unpublished doctoral dissertation, University of California, Berkeley.

Levy, P. K. (1964). *The ability to express and perceive vocal communications of feeling.* In J. R. Davitz (Ed.), The communication of emotional meaning. New York: McGraw-Hill.

Matsumoto, D. (1986). *Cross-cultural communication of emotion.* Unpublished doctoral dissertation, University of California, Berkeley.

Matsumoto, D. (1989). Cultural influences on the perception of emotion. *Journal of Cross-Cultural Psychology, 20,* 92–105.

Matsumoto, D. (1992). American-Japanese cultural differences in the recognition of universal facial expressions. *Journal of Cross-Cultural Psychology, 23,* 72–84.

Matsumoto, D., & Ekman, P. (1988). *Japanese and Caucasian facial expressions of emotion (JACFEE)* [Slides]. San Francisco, CA: Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University.

Matsumoto, D., & Ekman, P. (1989). American-Japanese differences in intensity ratings of facial expressions of emotion. *Motivation and Emotion, 13* (2), 143–157.

Mill, J. (1984). High and low self-monitoring individuals: Their decoding skills and empathic expression. *Journal of Personality, 52,* 372–388.

Mufson, L., & Nowicki, S. Jr. (1991). Factors affecting the accuracy of facial affect recognition. *The Journal of Social Psychology, 131,* 815–822.

Nowicki, S. Jr., & Duke, M. P. (1994). Individual differences in the nonverbal communication

of affect: The diagnostic analysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior, 18,* 9–35.

O'Sullivan, M. (1982). Measuring the ability to recognize facial expressions of emotion. In P. Ekman (Ed.), *Emotion in the human face* (pp.281–314). New York, NY: Cambridge University Press.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51,* 649–660.

Riggio, R. E., & Friedman, H. S. (1982). The interrelationships of self-monitoring factors, personality traits, and nonverbal social skills. *Journal of Nonverbal Behavior, 7,* 33–45.

Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal cues: The PONS test.* Baltimore, MD: Johns Hopkins University Press.

Scherer, K. (1986). Vocal affect expression: Review and a model for future research. *Psychological Bulletin, 99,* 143–165.

Toner, H. L., & Gates, G. R. (1985). Emotional traits and recognition of facial expression of emotion. *Journal of Nonverbal Behavior, 9,* 48–66.

Trimboli, A., & Walker, M. (1993). The CAST test of nonverbal sensitivity. *Journal of Language and Social Psychology, 12,* 49–65.

Wallbott, H. G. (1992). Effects of distortion of spatial and temporal resolution of video stimuli on emotion attributions. *Journal of Nonverbal Behavior, 16,* 5–20.

Wolitzky, D. (1973). Cognitive control and person perception. *Perceptual and Motor Skills, 36,* 619–623.

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.