Motiv Emot
DOI 10.1007/s11031-011-9212-2

ORIGINAL PAPER

Evidence for training the ability to read microexpressions of emotion

David Matsumoto · Hyi Sung Hwang

Abstract Microexpressions are extremely quick facial expressions of emotion that appear on the face for less than ½ a s. To date, no study has demonstrated that the ability to read them can be trained. We present two studies that do so, as well as evidence for the retention of the training effects. In Study 1 department store employees were randomly assigned to a training or comparison group. The training group had significantly higher scores than the comparison group in microexpression reading accuracy at the end of the training; 2 weeks later the training group had better third-party ratings of social and communicative skills on the job. Study 2 demonstrated that individuals trained in reading microexpressions retained their ability to read them better than a comparison group tested 2–3 weeks after initial training. These results indicated that the ability to read microexpressions can be trained and are retained.

Keywords Emotion · Facial expressions · Microexpressions · Training · Nonverbal behavior

Introduction

The ability to read nonverbal behavior is related to a number of important social skills (Ambady et al. 1995; Costanzo and Archer 1989; DePaulo and Rosenthal 1979; Nowicki and Duke 1994). Accuracy in recognizing emotions in particular has been related to personality traits (Matsumoto et al. 2000), sociocultural adjustment and mental health (Carton et al. 1999; Yoo et al. 2006), negotiation effectiveness (Elfenbein et al. 2007), and workplace performance (Rosenthal et al. 1979). A meta-analysis of studies examining the relationship between emotion recognition and workplace performance reported an average effect of $r = .20$ (Elfenbein et al. 2007).

Because of the importance of the ability to read emotions accurately, researchers as early as 80 years ago reported that it was possible to improve emotion recognition (Guilford 1929), a finding that has been corroborated recently as well (Elfenbein 2006; McKenzie et al. 2000). An important development in this literature has been the reported training efficacy in individuals with Asperger's syndrome (Barnhill et al. 2002), autism (Bolte et al. 2006; Solomon et al. 2004), schizophrenia (Frommann et al. 2003; Silver et al. 2004; Wolwer et al. 2005), mental retardation (McAlpine et al. 1991, 1992a, b; Stewart and Singh 1995), and acquired brain injury (Guercio et al. 2004).

This literature has focused on the recognition of macroexpressions. These are expressions that are produced when emotions occur and there is no reason for them to be modified or concealed. They typically last between 0.5 and 4 s on the face and involve the entire face (Ekman 2003). Macroexpressions differ from microexpressions, which occur much more quickly, less than 0.5 s and sometimes as fast as 67 ms.

The idea that microexpressions exist has its roots in Darwin's (1872) inhibition hypothesis that suggested that facial actions that cannot be controlled voluntarily may be produced involuntarily even if the individual is trying to control his or her expressions.[1] Research on the neuroanatomical bases of

D. Matsumoto (✉) · H. S. Hwang
Department of Psychology, San Francisco State University and Humintell, LLC, 1600 Holloway Avenue, San Francisco, CA 94132, USA
e-mail: dm@sfsu.edu

[1] This idea may actually have earlier roots in the work of Duchenne (1862/1990), who demonstrated a difference in specific facial muscles between smiles of true enjoyment and non-enjoyment smiles. Darwin's inhibition hypothesis was different in that it suggested that suppressed emotional expressions could leak out involuntarily.

Journal : Large 11031 | Dispatch : 9-3-2011 | Pages : 11
Article No. : 9212 | LE | TYPESET
MS Code : MOEM591 | CP | DISK

emotional expressions suggests how this occurs. There are two distinct neural pathways that mediate facial expressions, each originating in a different area of the brain (Rinn 1984). The pyramidal tract drives voluntary facial actions and originates in the cortical motor strip, whereas the extrapyramidal tract drives involuntary emotional expressions and originates in the subcortical areas of the brain. When individuals are in intensely emotional situations but need to control their expressions, they likely activate both systems, which engage in a neural "tug of war" over control of the face, allowing for the quick, fleeting leakage of microexpressions.

The existence of microexpressions was verified almost a century after Darwin by Haggard and Isaacs (1966) while scanning films of psychotherapy sessions.[2] Later Ekman and Friesen (1974) demonstrated that microexpressions occurred in their frame by frame analysis of interviews with depressed inpatients. Most recently Porter and ten Brinke (2008) demonstrated that microexpressions occurred when individuals attempted to be deceitful about their emotional expressions.[3]

Microexpressions are likely signs of concealed or deceptive emotions.[4] As such the ability to detect them may be important for individuals such as health care workers, psychotherapists, law enforcement officers, or anyone whose profession requires face-to-face interpersonal skills. This could benefit the development of rapport, trust, and collegiality, in making credibility assessments, evaluating truthfulness and detecting deception, and provide the basis for better cooperation, negotiation, or sales. But because they are so quick, observers may miss them when they occur; to wit they were only detected initially when films were analyzed in frame by frame detail (Ekman and Friesen 1974; Haggard and Isaacs 1966). Thus it is questionable as to whether or not the ability to detect microexpressions can be trained.

One limitation of this area has been the lack of validated tools to measure and train the ability to read microexpressions. Based on their discovery Haggard and Isaacs (1966) developed a test of microexpression detection ability that was used in two published reports (Garwood et al. 1970; Taylor et al. 1969). Later Ekman and Friesen developed the Brief Affect Recognition Test (BART; Ekman and Friesen 1974), in which facial expressions were presented very rapidly to observers tachistoscopically. One of the problems with this approach was that presented images remained on the retina longer than intended, obscuring the assessment. Matsumoto et al. (2000) corrected this problem by imbedding emotional expressions within a 1 s presentation of the same expressor's neutral expression, producing a forward and backward mask that eliminated the after effects of the image. They provided evidence for the internal and temporal reliability of the scores generated, and for convergent and concurrent validity. Scores were also not influenced by individual differences in visual acuity for rapidly presented faces. The emotion recognition task used in the studies reported here is based on this methodology.

We present two studies that examine whether microexpression recognition training can improve individuals' ability to read microexpressions of emotion (Studies 1 and 2) and whether this improved ability is sustained across time (Study 2). In Study 1 a sample of working adults from a non-western culture was randomly assigned to either a microexpression training group or a comparison group. Later, colleagues evaluated the trainees on scales measuring communication skills. We hypothesized that training in the ability to read microexpressions of emotion will produce improvements in the ability to read them immediately at the end of training and greater subsequent third-party ratings.

Study 2 examined whether or not the training benefit documented in Study 1 was retained several weeks after initial training in a different sample of working adults in the US. We hypothesized that accuracy scores of participants of a training workshop would be higher than those of a comparison group, even 2–3 weeks after the workshop.

Study 1

Method

Participants

The participants were 81 employees at a major retail store near Seoul, South Korea (32 males, 48 females, 1 declined to answer; mean age = 30.30 years, SD = 8.12). All were born and raised in South Korea and spoke Korean as their first and primary language. Thirty-nine had high school

[2] Relatedly, Condon and Ogston (1967) published a microanalysis of interactional behavior. While not focused on facial expressions they employed a micro-level, frame-by-frame analysis of nonverbal behavior in a psychotherapeutic interview of a mother, father, and child. They suggested that nonverbal behavior can be analyzed with such precision and that such micro-level nonverbal behaviors signal something meaningful about the mental states of the encoders.

[3] To our knowledge Porter and ten Brinke's (2008) study was the first empirical study published in a peer-reviewed journal that demonstrated the existence of microexpressions, as previous evidence for microexpressions were limited to book chapters and books. Many peer-reviewed articles on expression in deceptive situations do exist (e.g., Ekman et al. 1981; Frank and Ekman 1993) but these focus on the form of the expressions—i.e., the presence or absence of specific facial muscles, ala Duchenne's (1862/1990) work mentioned earlier—and not on their speed of occurrence (microexpressions).

[4] This claim is consistent with the theoretical and empirical work on microexpressions to date. It may be that microexpressions are also signs of rapidly processed but unconcealed emotional states. Future research will need to investigate this and other possibilities.

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

degrees, 29 had either two-year junior college or uncompleted four-year university educations, 9 had university undergraduate degrees, and the rest declined to comment. Twenty were married, 35 were single, and the rest declined to comment. Fifty-four percent were in sales; the remainder was in administration or staff positions. The use of a non-American sample ensured that any benefits would not be specific to western cultures, and the use of a non-university sample ensured that findings were not specific to students.

Microexpression training and judgment tasks

We used the Microexpression Recognition Training Tool (MiX) to assess microexpression recognition accuracy and to deliver the training and practice. It included 84 expressions of male and female expressers of six ethnic groups (Caucasian, African, Asian, Hispanic, Middle Eastern, and South/Southeast Asian). All expressions were selected on the basis of meeting criteria for signaling emotions (Ekman and Friesen 1975, 1978) and verified by two coders using the Facial Action Coding System (FACS; Ekman and Friesen 1978; reliability = .86), ensuring that the expressions were valid analogs of the expressions that actually occur when emotions are aroused (Ekman and Friesen 1975; Matsumoto et al. 2008). Two studies obtaining observer judgments of these expressions confirmed that they were reliably judged as portraying the emotions intended (Matsumoto and Hwang 2011). The expressions were presented for 67 ms sandwiched in between two 1 s presentations of the same expresser's neutral face.

The MiX was comprised of five sections: (1) a 14-item pretest that had two examples of anger, contempt, disgust, fear, enjoyment, sadness and surprise; (2) an instructional section involving rolling morphs and audio in which each of the expressions was introduced and described and where pairs of expressions most commonly confused with each other were compared; (3) a 42-item practice section featuring more examples of microexpressions, with immediate feedback for judgments and the ability to replay and freeze frame items; (4) a review section, which reviewed the material presented in instruction with different expressers; and (5) a 28-item post-test that featured expressions not used in any of the previous sections. Expressions were balanced for expresser ethnicity and sex to the extent possible across the five sections. Because of time constraints only half of the 42 item practice section and only two items from the review were used.

For the pre- and post-tests, participants completed a fixed-choice judgment in which they selected a single emotion label from a list provided that corresponded to the emotion displayed. The response alternatives included anger, contempt, disgust, fear, happiness, sadness, surprise, and neutral, along with a 'none' option to help prevent any artifactual agreement (Frank and Stennett 2001). Accuracy scores were computed by coding each participant's selections as 1 when they matched the emotion intended and 0 when they did not. Accuracy scores were then averaged to compute pre- and post-test scores separately for each emotion and then across emotions for a total score. Cronbach alphas were low for most of the pre-test scores (range .10–.83), likely because there were only two items per score; results below should be interpreted with that caveat. Reliabilities for the post-test scores were higher than pre-test and with only one exception in the acceptable range (range .37–.91). Reliabilities for the pre- and post-test total scores for the entire group were .62 and .87, respectively.

Design and procedures

All participants attended a two-h workshop conducted by the authors, who had not met any of the trainees prior to the event. Prior to their arrival at the workshop, 2/3rds of the trainees were randomly assigned to the training group, the remainder to the comparison group. (Trainees were unaware of what group they were assigned to, and the cooperating department store had requested more of their employees actually receive the training). Upon arrival at the event, which was held in a small theater, participants received a packet containing a brief demographic questionnaire and the answer sheets for the microexpression training. They completed the demographic questionnaire.

The session was conducted in English by the first author and translated into Korean by the second author. All participants were given a brief introduction to the training and then the MiX pre-test. Then they were told that there were two versions of the training; one group (the training group) would stay in the theater while the other (the comparison group) would go to an adjacent room. After the comparison group left, the training group then remained and had a brief lecture about emotions and facial expressions, followed by the instruction, practice, and review sections of MiX. In the other room, which was configured exactly the same way as the first, the comparison group watched a one-h video of the first author giving a lecture on emotions and facial expressions covering essentially the same content (dubbed in Korean), and introducing the facial expressions of emotion in still photos. They were also given a Korean translation of the first chapter from the book *Emotions Revealed* (Ekman 2003). The comparison group did not receive specific training on reading microexpressions.

After the 1 h training the comparison group returned to the theater and joined the training group. Neither group was instructed about the difference between the two groups. All participants then completed the MiX post-test and were given a brief overview of how to use the ability to recognize emotions in others in their daily lives and on the

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

job. The experimental design was a 2 (training vs. comparison) by 2 (pre- and post-test) mixed design.

Third-party outcome measures

Approximately 2 weeks after the training event, outcome data were obtained from third party colleagues (some peers and some higher in status) who were blind to the nature of the training, conditions, and assignments, and completed a set of evaluations about the participants.[5] The measures were originally designed by a training company in Japan in consultation with the first author, and piloted in a study of Japanese company employees (N = 35) investigating emotional competence in the workplace. It consisted of two sections, each of which produced two scores (thus four scores total). The first section presented eight vignettes describing real-life job situations (e.g., "A serious problem has occurred in a project that you and your colleague are responsible for. At this rate, the project will not be finished by the deadline, and both of you are worried. You have started to discuss the situation with your colleague."). Each vignette was then followed by two items, the first asking about the degree to which the participant would realize the gravity of the situation or the emotions of the third-party, and the second asking about whether the participant would take appropriate action to alleviate the situation or intervene with the third party's emotions. Each of the ratings was made using four-point scales. The ratings for both sets of questions were averaged to produce two composite scores labeled Emotional Sensitivity and Using Emotions, respectively (αs = .92 and .94).

The second section consisted of 20 items selected from an original pool of 75 items assessing a variety of work related competencies and behaviors of the participant. The 20 items consisted of the 10 highest loading items on two factors based on an exploratory factor analysis of the original item pool in the Japanese study. The first factor was labeled Social Skills and the second Communication Skills (see "Appendix" for the scales). All items were rated using 4-point scales, labeled 1, Strongly Disagree; 2, Disagree; 3, Agree; 4, Strongly Agee. Ratings for both scales were averaged (αs = .95 and .95).

Third-party ratings for each participant were obtained from at least one person, and in some cases more than one, in which case ratings were averaged to produce a single set of ratings for each participant.[6] All protocols were translated into Korean and accuracy was verified using back-translation procedures. There were no problems encountered in the translation.

Results and discussion

We computed a mixed, two-way ANOVA on the MiX Total scores using Condition (training vs. comparison) and Time (pre- vs. post-test) as factors. The interaction was significant, $F(1, 72) = 10.90$, $p < .01$, MSE = .195, $\eta_p^2 = .13$.[7] Simple effects of Time indicated that, as predicted, the training group increased in their MiX scores from pre- (M = .47, SD = .16) to post-test (M = .65, SD = .20), $F(1, 48) = 33.05$, $p < .01$, MSE = .025, $\eta_p^2 = .41$. There was no change, however, for the comparison group from pre- (M = .49, SD = .24) to post-test (M = .50, SD = .23), $F(1, 24) = .02$, $p = .89$, MSE = .022.

Separate analyses testing each of the seven emotions indicated that the training group's scores significantly improved on anger, $t(48) = 4.44$, $p < .01$; contempt, $t(48) = 4.32$, $p < .01$; disgust, $t(48) = 2.61$, $p < .01$; fear, $t(48) = 5.64$, $p < .01$; and sadness, $t(48) = 3.05$, $p < .01$. Scores on happiness, $t(48) = 1.11$, $p = .27$, and surprise, $t(48) = 1.00$, $p = .32$, were also in the same direction but not significant (Table 1). The non-findings on happiness and surprise were likely due to ceiling effects; the pre-test scores were already quite high (.82 and .77, respectively). None of the emotion scores for the comparison group improved (which is a bit surprising, given that some improvement may be expected from the practice in the task alone).

Training efficacy on third-party outcome measures

We computed univariate ANOVAs on each of the outcome measures using Condition as the independent variable. There were no significant differences between the groups on Emotional Sensitivity (M = 2.90, SD = .66 and M = 2.72, SD = .61, for training and comparison groups, respectively) or Using Emotions (M = 2.87, SD = .66 and M = 2.70, SD = .58, respectively. The training group had significantly higher scores, however, than the comparison group on Social Skills (M = 2.88, SD = .66 and M = 2.38, SD = .59, respectively), $F(1, 68) = 8.98$, $p = .01$, MSE = .411, $\eta_p^2 = .117$; and Communication Skills (M = 2.88, SD = .63 and M = 2.51, SD = .55, respectively), $F(1, 68) =$

[5] Even though the third-party raters were initially blind to condition of the individuals, it was very possible that they might not have been blind to condition at the time of the ratings. Participants may have communicated with each other and their raters about their training/conditions.

[6] Unfortunately, the number of participants who had more than one rater was too small to conduct any meaningful analyses.

[7] The analyses reported here utilized only those participants with complete data ($N = 72$).

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

Table 1 MiX Means and SD for each emotion, both studies

	Study 1				Study 2	
	Training group		Comparison group		Training group	Comparison group
	Pre-test	Post-test	Pre-test	Post-test		
Anger	.40 (.36)	.60 (.31)	.43 (.46)	.49 (.36)	.89 (.22)	.56 (.08)
Contempt	.40 (.36)	.70 (.33)	.47 (.40)	.51 (.36)	.88 (.22)	.36 (.34)
Disgust	.37 (.34)	.55 (.33)	.33 (.36)	.33 (.33)	.68 (.28)	.48 (.28)
Fear	.19 (.35)	.53 (.37)	.28 (.37)	.33 (.34)	.71 (.29)	.43 (.29)
Joy	.82 (.31)	.86 (.21)	.83 (.36)	.85 (.29)	.94 (.15)	.64 (.36)
Sadness	.41 (.39)	.58 (.33)	.48 (.37)	.38 (.38)	.91 (.16)	.53 (.37)
Surprise	.77 (.31)	.83 (.27)	.78 (.34)	.79 (.32)	.87 (.17)	.70 (.34)

5.77, $p = .02$, MSE = .371, $\eta_p^2 = .08$).[8] The findings, therefore, provided some support for the notion that training produced better third-party ratings.

Do MiX scores mediate the training effect on third-party outcome measures?

Condition differences on two of the outcome measures and five of the MiX pre-post emotion scores established two of the criteria required for mediation analysis, namely that the independent variable (condition) had effects on both the dependent variable (outcomes) and the mediator (emotion recognition scores) (Baron and Kenny 1986). To establish a third condition of mediation, that changes in MiX emotion scores were related to the outcome variables after controlling for condition differences, we computed a hierarchical regression on Social Skills and Communication Skills entering Condition on the first step and difference scores for anger, contempt, disgust, fear, and sadness on the second, using stepwise inclusion criteria (p value of F required for entry = .05; for removal = .10). For Social Skills, differences in fear ($\beta = .340$, $p = .01$) and contempt ($\beta = .291$, $p = .04$) were entered into the equation, final $R(62) = .511$, $p < .01$. For Communication skills, differences in fear ($\beta = .307$, $p = .02$) was entered into the equation, final $R(62) = .406$, $p < .01$. Thus changes in these emotions were significantly related to the two outcome variables that differed between the groups, meeting the third criterion for mediation.

To examine whether differences in recognition of fear or contempt mediated the training effects, we computed Sobel tests. For Social Skills, the Sobel tests for fear was significant, Sobel = 1.96, $p = .05$; while contempt was marginally significant, Sobel = 1.65, $p = .10$. For Communication Skills, the Sobel test for fear was again significant, Sobel = 1.96, $p = .05$. At the same time, the regression coefficients for condition in the final equation of the regressions were still significant. Thus changes in two emotion scores as a result of MiX training partially mediated the training effect on Social Skills and Communication Skills.

There was a possibility that individuals in the training group came to the experiment with better ability to read microexpressions and better outcome ratings to begin with. But separate analyses testing differences between the two groups on each of the seven emotion scores and MiX Total score at pre-test generated no significant findings, and correlations between the MiX pre-test scores and the four outcome ratings were all non-significant. Correlations between gender, travel experience, income, and experience working collaboratively and the outcome ratings were significant; but when these demographic variables were included as covariates in the ANOVA, the condition effect still survived.

Thus MiX training produced the desired improvement in the ability to see microexpressions, and these improvements were associated with better third-party ratings at a subsequent time. One reason why changes in fear (and contempt) recognition mediated the training effect on third-party ratings may have been because the improvements on these emotions were the largest of all the emotions, and because of the intercorrelations among the emotions. That is because changes in fear and contempt recognition were the best discriminator of the differences between the training and comparison groups they entered the regressions; because changes among the seven emotions were intercorrelated, no other emotions were entered.

[8] It may be the case that these findings occurred because of Hawthorne or halo effects. On one hand, that positive effects were found on some scales but not others may argue against Hawthorne or halo in confounding the findings because it would be difficult to argue that such effects occurred differentially on some scales but not others. And the two non-significant effects did not occur because of low reliabilities among the dependent variables. On the other hand, all four variables showed the same group effect in the same direction, and it could be argued that the effect just happened to be larger for two variables, pushing them past statistical significance. The results should be interpreted with this caveat.

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

One limitation of Study 1 was the fact that there was no follow up test of the retention of the training benefit. That is, because the post-test was taken immediately at the end of the training session, there is no reason to believe that the training benefit sustained itself thereafter and contributed to the third-party ratings. Study 2 examined whether the training benefit was retained a few weeks after initial training. No such retention data exist in the literature.

Study 2

Method

Participants

Participants in the training group were 25 practicing trial consultants (11 male, 14 female, mean age = 46.31 years) who participated in a workshop conducted by the first author on reading facial expressions of emotion. The workshop was part of pre-convention activities for a society of trial consultants. The participants were mainly psychologists or attorneys with an average of 15.69 years in the profession.

A comparison group that received no training also participated (13 males, 17 females, mean age = 49.83). They were recruited using a snowball method through acquaintances of the authors, targeting the average age of the training group. Consequently, there were no age differences between the two groups, $t(41) = .99$, $p = .33$.

Training and retention assessments

The training group received training in reading microexpressions using the same version of MiX used in Study 1, using the same procedures. Later both they and the comparison group completed a different version of a microexpression training tool. This tool was administered online and presented an entirely new set of expressions produced by different expressors. There were four examples of each of the seven emotions, resulting in a 28-item test. Each of the target emotional expressions was FACS coded by both authors, ensuring that each expression included the target facial movements associated with the prototypic emotional expressions. Each target expression was presented for 200 ms, sandwiched in between two 1 s neutral expressions of the same expressor. Despite the fact that the initial training with MiX presented expressions at 67 ms, we chose a 200 ms presentation time for several reasons. First, it was still below the threshold of most normal expressions that are not concealed (500 ms; see Ekman 1993). Second, in our experience most microexpressions do not occur as fast as 67 ms; thus a 200 ms exposure is more likely to be ecologically valid. And third, the only study to actually compare different exposure times presented in the same format reported that a 200 ms exposure is optimal for producing individual differences in judgment accuracy rates (Matsumoto et al. 2000).

Procedures and tasks

Participants in the training group were recruited from the workshop described above, which was essentially the same as that delivered in Study 1, but geared to the work content of the target audience. During the workshop, the participants completed the MiX pre-test, received instruction on the appearance of each of the seven emotions, practiced reading them in microexpression format, and completed the MiX post-test, just as participants in Study 1 did.

At the end of the workshop trainees were asked to participate in a follow-up study examining retention. All participants consented and were asked to provide their email addresses. Two weeks later a notice was sent to them by email with a link to the online post-test tool. Thirteen of the 25 participants completed the second trial within a week of that notice, for a 52% response rate between 2 and 3 weeks after the initial training (the time period corresponding to the same time period in Study 1 when the third-party ratings were assessed).

The same link was provided to the comparison group via the acquaintances of the authors. When judging the emotions all participants used a fixed-choice judgment task with the alternatives anger, contempt, disgust, fear, happiness, sadness, surprise, neutral, and other, clicking one of these buttons adjacent to the presentation area of the face. Accuracy scores were computed the same way as in Study 1; across both groups, Cronbach's αs were .68, .82, .44, .62, .79, .76, .67, and .92 for anger, contempt, disgust, fear, happiness, sadness, surprise, and Total scores, respectively. We also captured the participants' reaction times to each item, defined as the time elapsed from the end of the presentation of the item until a response (click) was made.

Results

Training efficacy manipulation check

We averaged the accuracy scores across all expressions of the same emotion and conducted a two-way, repeated measures ANOVA using Time (2: pre v post) and Emotion (7) on the training group's MiX scores from the workshop. The main effect of Time was significant, $F(1, 22) = 32.30$, $p < .01$, MSE = .078, $\eta_p^2 = .60$, indicating that scores improved from pre (M = .45, SD = .13) to post (M = .63, SD = .15). The Time by Emotion interaction was also significant, $F(1, 22) = 50.28$, $p < .01$, MSE = .071,

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

$\eta_p^2 = .70$. Participants increased their scores on anger, $F(1, 24) = 17.13$, $p < .01$, MSE = .044, $\eta_p^2 = .42$; contempt, $F(1, 23) = 54.75$, $p < .01$, MSE = .064, $\eta_p^2 = .70$; disgust, $F(1, 24) = 5.19$, $p = .04$, MSE = .131, $\eta_p^2 = .18$; and fear, $F(1, 23) = 30.71$, $p < .01$, MSE = .069, $\eta_p^2 = .57$. The means for sadness were also in the direction of improvement, but not significant, probably due to sampling error of either the expressions used or individual sensitivities to recognition of sadness. The non-findings for happiness and surprise were similar to that of Study 1. These results largely replicated those from Study 1 and ensured that training was effective.[9]

Retention of microexpression reading ability

We averaged the accuracy scores across all four examples of each of the emotions in the retention test and computed a Condition (training vs. comparison) by Emotion (7) two-way mixed ANOVA. The Condition main effect was significant, $F(1, 41) = 17.90$, $p < .01$, MSE = 5.472, $\eta_p^2 = .30$, indicating that the training group had significantly higher accuracy scores (M = .84, SD = .12) than the comparison group (M = .53, SD = .25). The interaction between Condition and Emotion was also significant, $F(6, 246) = 2.55$, $p = .04$, MSE = .818, $\eta_p^2 = .06$. Simple effects analyses indicated that the training group had significantly or marginally significantly higher scores on all seven emotions than the comparison group, $F(1,41) = 10.81$, $p < .01$, MSE = 1.429, $\eta_p^2 = .21$; $F(1,41) = 24.02$, $p < .01$, MSE = 1.673, $\eta_p^2 = .37$; $F(1,41) = 4.54$, $p = .04$, MSE = 1.255, $\eta_p^2 = .10$; $F(1,41) = 8.60$, $p < .01$, MSE = 1.306, $\eta_p^2 = .17$; $F(1,41) = 8.45$, $p < .01$, MSE = 1.553, $\eta_p^2 = .17$; $F(1,41) = 12.24$, $p < .01$, MSE = 1.702, $\eta_p^2 = .23$; $F(1,41) = 2.71$, $p < .10$, MSE = 1.464, $\eta_p^2 = .06$, for anger, contempt, disgust, fear, happiness, sadness, and surprise, respectively (see Table 1 for descriptives). Thus training produced a retention of the ability to recognize all seven emotions.

We also compared the retention rates of the training group against norm data for expressions presented at the same speed. Recall that a previous report had compared judgment accuracy rates for three speeds of presentation, including 200 ms (Matsumoto et al. 2000). We computed a weighted average of the mean accuracy rates for 200 ms across the studies in that report (69.16%) and computed a single sample t-test comparing the mean accuracy rate of the training group in this study to that norm value. The training group had significantly higher scores than the norm comparison group, $t(12) = 4.53$, $p < .01$.

Response times

We averaged the response times across the four examples of each of the seven emotions and computed a Condition (training vs. comparison) by Emotion (7) two-way mixed ANOVA. The Condition main effect was significant, $F(1, 40) = 9.32$, $p < .01$, MSE = 3.617×10^{-7}, $\eta_p 2 = .19$, indicating that the training group had faster response times (M = 3.25, SD = 1.03) than the comparison group (M = 5.56, SD = 2.62). The interaction was not significant.

Important non-findings

It was possible that the trial consultants who participated in the retention test differed from those who elected not to. We compared those who participated in the retention test with those who did not on gender and age but there were no differences on either, $\chi^2(26) = .004$, p = .95; and $t(24) = .82$, p = .42, respectively. We also compared them on their original MiX pre- and post-test scores during the original training, but there were no differences between the groups on either test, $t(23) = .85$, $p = .40$; and $t(25) = .49$, $p = .63$, respectively.

Discussion

The training group had significantly higher retention scores than both a comparison group and norm data, indicating that the ability to read microexpressions was retained 2–3 weeks after initial training, which itself was initially effective (replicating the training benefit reported in Study 1). The training group was not only more accurate but also quicker in their responses. Thus the improved ability to read microexpressions was retained after training.

To be sure, there were limitations to Study 2. Perhaps the most important concerned the nature of the participants of the training group who participated in the retention assessment. Although the response rate was not 100%, it is entirely possible that this was a self-selecting group of individuals who tried to use the skills they had learned in the workshop and thus were more conscious of reading facial expressions and were simply better at it or more motivated than those who did not. On one hand this group's scores may inflate the difference between training and comparison, as we did not (could not) include participants who didn't want to participate and who may have had lower scores. On the other hand, this group may be precisely the group to compare, as in any training there will be some who resonate with the training and carry it forward

[9] All analyses reported in this section and the next were recomputed using only participants with complete data; all findings were replicated. Detailed report of the analyses is available from the authors.

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

and those who will not. Future studies will need to control for this possible self-selection bias in participation.

While the training and comparison groups were matched in gender and age, they likely differed on a number of other important characteristics such as in their professional backgrounds and motivation to learn how to read nonverbal behavior. Although the non-significant differences between the trial consultants who participated in the retention test and those who did not mitigated these concerns somewhat, it is still possible that the differences between the training and comparison groups on other characteristics were so large to question whether or not they are appropriate to compare in the first place. Such questions would also be germane to the comparison of the training group with the normative values as well. The results reported above, therefore, should be interpreted with these caveats and future studies should address this issue to ensure greater comparability of groups.

General discussion

Training improved the ability to read microexpressions, this ability was retained a few weeks after initial training, and the improvements were associated with better third-party ratings of socio-communicative skills. The training effects were obtained in working adults with a wide range of ages and educational backgrounds. Study 1 participants were not Americans; they were South Koreans, and in that study training was done in English translated to Korean.

There were many limitations to the studies. One had to do with the nature of the expressions used in the training, which were all full-face, relatively high intensity expressions. The only peer-reviewed evidence for the existence of microexpressions to date (Porter and ten Brinke 2008) indicated that many microexpressions were shown in either the upper or lower halves of the face only. This suggests that the expressions of anger, fear, sadness and surprise that we used, which were displayed in both the upper and lower parts of the face, may not reflect how these emotions are actually produced in real-life microexpressions, thus limiting our findings' generalizability. Future research will need to examine whether or not microexpressions of partial versions of these expressions can be trained to be recognized. Our findings on contempt, disgust, and happiness are likely unaffected by this issue as the defining action of these emotions occurs only in one part of the face anyway. (It is interesting to note that the training effects for contempt and disgust were on par with those of anger, fear, sadness, and surprise.)

Another limitation of our work was the format of the training. MiX presents a target emotional expression imbedded within presentations of the same expressor's neutral expression. In real life, however, microexpressions may not always occur as changes from a neutral baseline back to a neutral. They can occur from a neutral baseline and end in another emotion (or series of emotions), or they can start from a different emotion. Future research should create analogs of these different ways in which microexpressions actually occur in real life and examine whether the training benefits differ for different presentation formats.

A major limitation of Study 1 concerned the nature of the third-party outcome ratings, and in particular the lack of data demonstrating their ecological validity. Future studies surely need to examine the validity of those measures and to utilize other measures of outcomes. Even if the measures we used were considered valid, they were limited in other ways. Kirkpatrick (1998) described four levels of training evaluation criteria: reactions, learning, behaviors, and results. The ratings obtained in Study 1 are reactions (participants' improved microexpression scores may be indicative of learning). Clearly the data we obtained do not speak to actual on-the-job behaviors or to objective performance measures. These kinds of data should be collected in the future.

Nevertheless, these data are the first to demonstrate that microexpression recognition ability can be trained, and lead to a number of interesting possibilities for future research. For example they open the door to new studies examining the boundaries of the training effect, e.g., the optimal length of practice, the balance of emotions across items, the key characteristic of the training that produces a benefit, or the retention and recidivism rates across longer time periods. It would be interesting to examine whether training benefits differ depending on how non-morphological cues affect the images such as facial hair, glasses, or head or face covers. Future research needs to identify the active ingredient that contributes to the training effect; is it the introduction of each of the expressions, the presentation of the expressions in micro format, or the ability to practice viewing them? The training benefits may also differ depending on the signal clarity of the expressions used in the training. The expressions used in the MiX are high-intensity, full-face, high signal clarity expressions. Spontaneous expressions, however, are usually not as clear, especially when people are talking. The inclusion of spontaneous expressions may further test the boundary of the training effect. The fact that the comparison group in Study 1 was merely given information about emotions and faces and shown still photo examples of the seven emotions suggests that the training benefits were not simply due to learning about emotion and expression; but it does open the door to the interesting possibility that macroexpression and microexpression training may have different benefits, which future studies may explore. The ceiling effects for happiness and surprise in both studies suggest that these

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

emotions need to be trained with different parameters—such as faster presentation speeds—in order to produce enough variation to allow for training effects to occur. Additionally the characteristics of some of the expressions in the training may be a limitation that needs to be addressed in the future.

One potentially important avenue for future research is the examination of benefits of microexpression recognition training to deception detection. Several studies have already documented that the ability to recognize microexpressions is correlated with the ability to detect deception (Frank and Ekman 1997; Warren et al. 2009). And as mentioned in the Introduction theoretical work on the existence of microexpressions is based on its relationship to deception (Darwin 1872; Ekman 1985). Future research should consider the interesting possibility that training in microexpression recognition can improve deception detection abilities, or conduct interviews or interrogations.

Future research also needs to examine if microexpression training (or any other kind of training in reading nonverbal behavior) produces real-world benefits much more comprehensively such as in sales, business negotiations, health care situations, or cross-cultural adaptation. All of these are at least theoretically possible because microexpressions are signs of concealed emotions and people who can recognize them should be in a better position to address an expresser's true emotions. Thus the improved ability to see microexpressions should aid groups to work collaboratively; colleagues to build better relationships, interpersonal trust, and rapport; bosses to communicate intent; and subordinates to read and interpret intent. The ability to read microexpressions gives people one more important tool in their interpersonal toolkit, providing the basis for better cooperation, negotiation, and sales. All of these are potentially fruitful lines of future research.

To be sure improved ability to read microexpressions, or any nonverbal behavior for that matter, is just the first step. What one does with the information is an important second step in the process of interaction. Being overly sensitive to nonverbal behaviors such as microexpressions and other forms of nonverbal leakage can be detrimental to interpersonal outcomes as well, as discussed in literature on eavesdropping (Blanck et al. 1981; Elfenbein and Ambady 2002; Rosenthal and DePaulo 1979). Individuals who call out other's emotions indiscriminately can be considered intrusive, rude, or overbearing. Dealing effectively with emotion information of others is also likely to be a very crucial part of the skill set one must have to interact effectively with others. Knowing when and how to intervene, to adapt one's behaviors and communication styles, or engage the support and help of others, are all tactical skills that must be brought into play once emotions are read.

Acknowledgments Portions of this report were prepared with the support of research grant W91WAW-08-C-0024 from the Army Research Institute, and FA9550-09-1-0281 from the Air Force Office of Scientific Research to the first author. We thank Gil-Won Hwang and the Sopooong Department Store in South Korea for their assistance in the project.

Appendix

Questionnaire used by colleagues for third-party ratings in study 1

How much do you agree or disagree with the following statement about the behavior of the target individual? Please indicate your rate of agreement and mark your answer on your answer sheet

1 Strongly Disagree.
2 Disagree.
3 Agree.
4 Strongly Agree.

Social skills

1. Breakthrough current situation and promote new ideas even when experiencing gridlock or resistance.
2. Often attracting and leading people.
3. Always pursuing challenging goals and making the effort towards accomplishing those goals.
4. Set high self-motivation standard to accomplish tasks.
5. Promote change by engaging others when many are not willing to change
6. Good at getting people involved.
7. Recruiting people to initiate action even when others had given up.
8. Possess human network in a wide range of departments.
9. Indicate clear vision even in a confusing situation.
10. Proactively detect risks and lead others even when problems are not exposed at the surface.

Communication skills

1. Understand the flow of conversation from the positioning of the person.
2. Take time to listen to others without interrupting.
3. Accept others' opinions if they are right although having his/her own opinions.
4. Never behave in a manner that may provide unpleasant feelings to others regardless of how difficult the situation becomes.

Journal : Large 11031 Dispatch : 9-3-2011 Pages : 11
Article No. : 9212 LE TYPESET
MS Code : MOEM591 CP DISK

5. Handle severe criticism by others without expressing anger.
6. Listen to others even when busy and pressured.
7. Listen to others, placing him/herself in the other person's position.
8. Act with respect towards others although being in a position of authority.
9. Listen to others with positive gestures.
10. Does not hold on to the past success.

NB: These items were randomized in a single questionnaire.

References

Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology, 69*(3), 518–529.

Barnhill, G. P., Cook, K. T., Tebbenkamp, K., & Myles, B. S. (2002). The effectiveness of social skills intervention targetting nonverbal communication for adolescents with Asperger syndrome and related pervasive developmental delays. *Focus on Autism and Other Developmental Disabilities, 17*(2), 112–118.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1882.

Blanck, P. D., Rosenthal, R., Snodgrass, S. E., DePaulo, B. M., & Zuckerman, M. (1981). Sex differences in eavesdropping on nonverbal cues: Developmental changes. *Journal of Personality and Social Psychology, 41*(2), 391–396.

Bolte, S., Hubl, D., Feineis-Matthews, S., Prvulovic, D., Dierks, T., & Poustka, F. (2006). Facial affect recognition training in autism: Can we animate the fusiform gyrus? *Behavioral Neuroscience, 120*(1), 211–216.

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior, 23*, 91–100.

Condon, W. S., & Ogston, W. D. (1967). A segmentation of behavior. *Journal of Psychiatric Research, 5*, 221–235.

Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The Interpersonal Perception Task. *Journal of Nonverbal Behavior, 13*(4), 225–245.

Darwin, C. (1872). *The expression of emotion in man and animals*. New York: Oxford University Press.

DePaulo, B. M., & Rosenthal, R. (1979). The structure of nonverbal decoding skills. *Journal of Personality, 47*(3), 506–517.

Duchenne de Boulogne, G. B. (1862/1990). *The mechanism of human facial expression*. New York: Cambridge University Press.

Ekman, P. (1985). *Telling lies: clues to deceit in the marketplace, politics, and marriage* (1st ed. ed.). New York: Norton.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist, 48*(4), 384–392.

Ekman, P. (2003). *Emotions revealed* (2nd ed. ed.). New York: Times Books.

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, D. C.: Winston and Sons.

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing the emotions from facial cues*. Englewood Cliffs, NJ: Prentice Hall.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: Investigator's guide*. Palo Alto, Calif.: Consulting Psychologists Press.

Ekman, P., Hager, J., & Friesen, W. V. (1981). The symmetry of emotional and deliberate facial actions. *Psychophysiology, 18*, 101–106.

Elfenbein, H. A. (2006). Learning in emotion judgments: Training and the cross-cultural understanding of facial expressions. *Journal of Nonverbal Behavior, 30*(1), 21–36.

Elfenbein, H. A., & Ambady, N. (2002). Predicting workplace outcomes from the ability to eavesdrop on feelings. *Journal of Applied Psychology, 87*(5), 963–971.

Elfenbein, H. A., Foo, M. D., White, J., Tan, H. H., & Aik, V. C. (2007). Reading your counterpart: The benefit of emotion recognition accuracy for effectiveness in negotiation. *Journal of Nonverbal Behavior, 31*, 205–223.

Frank, M. G., & Ekman, P. (1993). Not all smiles are created equal: The differences between enjoyment and nonenjoyment smiles. Special Issue: Current issues in psychological humor research. *Humor: International Journal of Humor Research, 6*(1), 9–26.

Frank, M. G., & 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.

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80*(1), 75–85.

Frommann, N., Streit, M., & Wolwer, W. (2003). Remediation of facial affect recognition impairments in patients with schizophrenia: A new training program. *Psychiatry Research, 117*(3), 281–284.

Garwood, R., Guiora, A. Z., & Kalter, N. (1970). Manifest anxiety and perception of microomentary expression. *Comprehensive Psychiatry, 11*(6), 576–580.

Guercio, J. M., Podolska-Schroeder, H., & Rehfeldt, R. A. (2004). Using stimulus equivalence technology to teach emotion recognition to adults with acquired brain injury. *Brain Injury, 18*(6), 593–601.

Guilford, J. P. (1929). An experiment in learning to read facial expression. *Journal of Abnormal and Social Psychology, 24*(2), 191–202.

Haggard, E. A., & Isaacs, K. S. (1966). Micro-momentary facial expressions as indicators of ego mechanisms in psychotherapy. In L. A. Gottschalk & A. H. Auerbach (Eds.), *Methods of Research in Psychotherapy* (pp. 154–165). New York: Appleton-Century-Crofts.

Kirkpatrick, D. L. (1998). Evaluating training programs the four levels Available from http://www.columbia.edu/cgi-bin/cul/resolve?clio4245080.

Matsumoto, D., & Hwang, H. S. (2011). Judgment of Profile Views of Universal Facial Expressions of Emotion. *Manuscript currently submitted for publication*.

Matsumoto, D., Keltner, D., Shiota, M. N., Frank, M. G., & O'Sullivan, M. (2008). What's in a face? Facial expressions as signals of discrete emotions. In M. Lewis, J. M. Haviland, & L. Feldman Barrett (Eds.), *Handbook of emotions* (pp. 211–234). New York: Guilford Press.

Matsumoto, D., LeRoux, J. A., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., et al. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24*(3), 179–209.

Journal : Large 11031 | Dispatch : 9-3-2011 | Pages : 11
Article No. : 9212 | ☐ LE | ☐ TYPESET
MS Code : MOEM591 | ☑ CP | ☑ DISK

McAlpine, C., Kendall, K. A., & Singh, N. N. (1991). Recognition of facial expressions of emtoion by persons with mental retardation. *American Journal of Mental Retardation, 96*(1), 29–36.

McAlpine, C., Singh, N. N., Ellis, C. R., & Kendall, K. A. (1992a). Enhancing the ability of adults with mental retardation to recognize facial expressions of emotion. *Behavior Modification, 16*(4), 559–573.

McAlpine, C., Singh, N. N., Kendall, K. A., & Ellis, C. R. (1992b). Recognition of facial expressions of emotion by persons with mental retardation: A matched comparison study. *Behavior Modification, 16*(4), 543–558.

McKenzie, K., Matheson, E., McKaskie, K., Hamilton, L., & Murray, G. C. (2000). Impact of group training on emotion recognition in individuals with a learning disability. *British Journal of Learning Disabilities, 28*(4), 143–147.

Nowicki, S. J., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The diagnostic anlaysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior, 19*, 9–35.

Porter, S., & ten Brinke, L. (2008). Reading between the lies: Identifying concealed and falsified emotions in universal facial expressions. *Psychological Science, 19*(5), 508–514.

Rinn, W. E. (1984). The neuropsychology of facial expression: A review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin, 95*, 52–77.

Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology, 37*(2), 273–285.

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

Silver, H., Goodman, C., Knoll, G., & Isakov, V. (2004). Brief emotion training improves recognition of facial emotions in chronic schizophrenia: A pilot study. *Psychiatry Research, 128*(2), 147–154.

Solomon, M., Goodlin-Jones, B. L., & Anders, T. F. (2004). A social adjustment enhancement intervention for high functioning autism, Asperger's Syndrome, and Pervasive Development Disorder. *Journal of Autism and Developmental Disorders, 34*(6), 649–668.

Stewart, C. A., & Singh, N. N. (1995). Enhancing the recognition and production of facial expressions of emotion by children with mental retardation. *Research in Developmental Disabilities, 16*(5), 365–382.

Taylor, L. L., Guiora, A. Z., Catford, J. C., & Lane, H. L. (1969). The role of personality variables in second language behavior. *Comprehensive Psychiatry, 10*(6), 463–474.

Warren, G., Schertler, E., & Bull, P. (2009). Detecting deception from emotional and unemotional cues. *Journal of Nonverbal Behavior, 33*, 59–69.

Wolwer, W., Frommann, N., Halfmann, S., Piaszek, A., Streit, M., & Gaebel, W. (2005). Remediation of impairments in facial affect recognition in schizophrenia: Efficacy and specificity of a new training program. *Schizophrenia Research, 80*(2–3), 295–303.

Yoo, S. H., Matsumoto, D., & LeRoux, J. A. (2006). Emotion regulation, emotion recognition, and intercultural adjustment. *International Journal of Intercultural Relations, 30*(3), 345–363.

Journal : Large 11031 | Dispatch : 9-3-2011 | Pages : 11
Article No. : 9212 | ☐ LE | ☐ TYPESET
MS Code : MOEM591 | ☑ CP | ☑ DISK