



## ERES ELECTRONIC COURSE RESERVES DOCUMENT

Copyright citation for JOURNAL ARTICLES -- All rights reserved

Must be attached to each chapter or pages on reserve

Article Title:	Measuring The Emotion - Perception Component OF Emotional Intelligence		
Author:	Glenn Geher & Kristin A. Renstrom		
Publication title:	Advances in Psychology Research		
Volume:	27		
Year/Date:	2004		
Number of pages:	27		
Professor's Name:	Glenn Geher		
Course Title and Number:	GLENN'S PUBLISHED WORKS		

PLEASE DO NOT USE PENCIL OR STAPLES

*Chapter 9*

## **MEASURING THE EMOTION-PERCEPTION COMPONENT OF EMOTIONAL INTELLIGENCE**

*Glenn Geher and Kristin L. Renstrom*  
State University of New York at New Paltz

### **ABSTRACT**

This chapter summarizes our research pertaining to issues surrounding the emotion-perception component of emotional intelligence. In general, our research was designed to develop and assess the Emotional Accuracy Research Scale (EARS; Mayer & Geher, 1996), a performance measure of emotion-perception ability. The EARS served as the template for the "Stories" subsection of the Multifactor Emotional Intelligence Scale (Mayer, Caruso, & Salovey, 1999). Targets of the EARS were adult individuals who both reported their emotions and described different scenarios that potentially gave rise to those emotions. The EARS is comprised of several scenarios pertaining to these actual targets in addition to forced-choice questions asking participants to choose which of two emotions they believed the target felt more strongly (for each pair of forced-choice items, targets necessarily endorsed feeling one item more strongly than the other). This measurement paradigm allows for an assessment of two indices of emotional intelligence: a consensus score, pertaining to the degree to which observers' judgments tend to match targets' endorsements (weighted by the proportion of subjects choosing the same alternative as the participant), and a target-accuracy score, pertaining to the degree to which observers' judgments matched targets' actual emotion endorsements. Study 1, which included 321 participants, demonstrated both convergent and discriminant validity for the EARS; EARS scores were somewhat positively correlated with extant self-report indices of empathy and were only negligibly related to indices of other, more cognitive indices of intelligence. In a second study, 40 individually-run undergraduate students first completed the EARS and five self-report measures of trait empathy. Participants then engaged in a relatively realistic emotion-judgment task; they were asked to assess the emotions of six target individuals whose discussions of emotional issues were included in a videotape. Compared with their scores on the trait-empathy measures, participants' scores on the EARS were generally more predictive of accurately detecting the emotions of targets, thus providing evidence of the EARS' predictive validity as a measure of emotional intelligence. Measurement issues concerning emotional intelligence and the implications of the conceptual distinction between consensus versus target agreement are discussed.

## INTRODUCTION

Since 1990, academics and laypeople have expressed a great deal of interest in the idea of Emotional Intelligence (EI; Salovey & Mayer, 1990). While this construct has generally been defined as a set of abilities pertaining to emotional understanding in self and other (Mayer, Salovey, & Caruso, 2000a), specific models used to implement research in this area have employed a variety of conceptualizations of EI.

The different models of EI can be understood as representing either a set of mental abilities or as a blend of mental abilities and relevant dispositional attributes. Mayer et al. (2000a) propose a model that is comprised strictly of mental abilities. These authors argue that keeping EI theoretically and empirically distinct from potentially relevant personality traits allows for EI to exist as a traditional form of intelligence (comprised of a set of abilities) while, concurrently, allowing measures of EI to demonstrate discriminant validity. In other words, this mental abilities conceptualization allows for the creation of measures that are likely to be empirically distinct from such potentially relevant personality traits as empathy or openness to experience.

On the other hand, Bar-On (1997) and Goleman (1995) offer conceptualizations that Mayer et al. (2000a) consider "mixed" models of EI. The defining feature of a mixed model, compared to a mental-abilities model, is that mixed models necessarily incorporate both mental-ability constructs in addition to relevant dispositional qualities that are conceptually unrelated to mental abilities. For instance, while Bar-On's (1997) model includes *problem solving* as an important feature of EI, this author also includes *happiness*. Clearly, this model includes both abilities and non-ability-relevant dispositions.

While mixed models of EI have demonstrated both reliability and validity (Bar-On, 1997), the current work is designed to elaborate specifically on the mental abilities conceptualization of EI. The mental ability framework allows for the use of performance, as opposed to self-report measures. Thus, compared with mixed model approaches, the mental ability approach is less likely to yield scores that are affected by impression management and other self-reporting-relevant biases.

Mayer, Salovey, and Caruso's (2000a) mental abilities model of EI is comprised of four major areas. These areas include (a) the perception and expression of emotion, (b) the assimilation of emotion into cognitions, (c) an understanding of emotions, and (d) emotional regulation. Mayer and his colleagues (Mayer et al., 2000a; Mayer, Salovey, Caruso, & Sitarenios, in press) have developed performance-based measures of EI based on this conceptualization. Further, their research using such measures has demonstrated that EI meets traditional criteria of a legitimate and discernable form of intelligence (Mayer et al., 2000). Specifically, this research demonstrated that EI, as conceptualized in terms of the four aforementioned classes of abilities, (a) can be operationally defined as a set of distinct but related abilities, (b) correlates moderately with other indices of intelligence, and (c) develops with age.

## THE EMOTION-PERCEPTION COMPONENT OF EMOTIONAL INTELLIGENCE

While multiple models of EI exist, all such models seem to include a component pertaining to effective perception of emotions in others. For instance, Mayer et al. (2000a) include the identification of emotions in other people as a core element of EI. Similarly, Bar-On (1997) includes *empathy* while Goleman (1995) includes *empathic awareness*. Further, a great deal of research on emotion prior to the establishment of EI as a unique construct, focused on people's ability to detect emotion in others (e.g., Ekman & Friesen, 1971; Rosenthal, Hall, DiMateo, Rogers, & Archer, 1979).

In fact, many have argued that the primary function of emotions is to provide relevant information to others in social contexts. This point of view reflects an evolutionary perspective which suggests that emotions serve the purpose of increasing inclusive fitness by leading to such outcomes as alerting kin to potential dangers (e.g., concern), scaring off potential aggressors (e.g., anger), and indicating to a conspecific that some behavior was desirable and that it would be in that individual's interest to perform similar acts in the future (e.g., pleasure). Such an evolutionary perspective, first delineated in detail by Darwin (1872) and more recently described by Cosmides and Tooby (2000), clearly underscores the importance of the ability to accurately detect emotions in others. Individuals in our ancestral past who were able to accurately read the emotions of conspecifics were more likely to engage in behaviors that would have increased their inclusive fitness. Thus, from an evolutionary standpoint, the ability to accurately perceive emotions in others is of considerable importance and, perhaps, speaks to the ultimate cause of the evolution of emotions in the first place.

### Measuring Emotion-Perception Ability

The current research is primarily concerned with measuring the emotion-perception component of EI. Much of the past work in this area has focused on empathy, a construct that clearly relates to the ability to accurately perceive others' emotions. In past work, different researchers have defined "empathy" differently (Wiske, 1986). Some researchers, such as Batson and Coke (1983), have defined empathy as sharing feelings with others and responding compassionately to others' feelings. Self-report measures of empathy have generally been designed to tap this particular component of empathy (e.g., Mehrabian & Epstein, 1972). Other empathy researchers have focused on empathy as the ability to know what another person is feeling (e.g., Levenson & Ruef, 1992). This research has focused on a component of empathy called "empathic accuracy" (Ickes, Stinson, Bissonette, & Garcia, 1990) as well as "perspective taking" (Batson & Coke, 1983).

Emotional intelligence has been defined in relatively broad terms that encompass several components of empathy. The current work is primarily concerned with the component of EI relating explicitly to accurate perceptions of others' emotions. Different kinds of EI measures have previously been employed to predict the ability to accurately perceive others' emotions.

## Existing Measures of Emotional Intelligence

Existing measures of emotional intelligence may be divided into (a) self-report measures and (b) performance measures.

### *Self-report Measures*

Some self-report measures of the emotion-perception component of EI have attempted to tap this construct by asking participants to rate themselves on this ability (e.g., Batson & Coke, 1983). Such measures rely on participants' accurate self-reports as a basis for whether they are empathic people who can accurately interpret emotional stimuli.

Schutte, Malouff, Hall, Haggerty, Cooper, Golden, and Dornheim (1998) also developed a self-report measure of EI. Their scale requires people to describe how well they are able to "hear" the emotions of others. This scale was found to have high reliability. Also, these researchers demonstrated discriminant validity for this measure by showing that scores on their scale were negligibly related to indices of intellectual intelligence. More recently, Mayer, Salovey, and Caruso (1999) have made available an on-line empathy measure (as part of the larger Emotional IQ Test). This measure is also comprised of a series of self-reported items (e.g., "The suffering of others deeply disturbs me.") whereby subjects rate their perceptions of their own empathy.

While several self-report measures of emotional intelligence have demonstrated their usefulness in many domains, such measures have not generally been found to predict actual performance on realistic emotion-detection tasks (Levenson & Ruef, 1992). In other words, reporting that one is capable of accurately knowing another's emotions does not necessarily translate into the actual ability to do so. Also, self-report measures of emotional intelligence have been found to be significantly related to the social desirability response set (e.g., in one study (Geher, 1994), scores on the Crowne & Marlowe (1960) social desirability scale were positively and significantly correlated with scores on the fantasy, empathic concern, and perspective taking subscales of the Davis (1983) empathy scale).

### *Performance Measures*

Several researchers have employed actual performance measures to tap constructs related to the empathic component of EI. One such measure may be the Profile of Nonverbal Sensitivity (PONS; Rosenthal, Hall, DiMateo, Rogers, & Archer, 1979) which assesses nonverbal sensitivity to emotional content by measuring one's ability to know how others feel based on an interpretation of nonverbal behavior. This measure employed videotapes depicting actors displaying different intensities of several emotions (e.g., anger). In some tapes, the faces of the actors could be seen, while in others, actors' faces were partially covered; thereby forcing subjects to tap some other nonverbal cue (e.g., body language). Subjects completing the PONS are given the task of assessing qualitative and quantitative aspects of actors' emotions. High scores indicated a tendency to accurately determine the targets' intended emotions; an ability that is very similar to the emotion-perception component of Mayer et al.'s (1999) conceptualization of EI.

A similar measure has been developed by Ickes and his colleagues (1990). In this research, participants were asked to write down the emotion they thought a target felt. Then, blind judges rated whether what these participants wrote was similar to corresponding sentences written by the targets themselves. Similarly, Levenson and Ruef (1992) had

participants watch videotapes of others and then assess how the targets of those videotapes were feeling using an affect-rating dial. Ratings were made on a negative-to-positive-affect continuum. Empathic accuracy was measured by how closely the dial ratings of a subject matched the dial ratings of a target. Again, the measure of empathy was how well one can judge the feelings of others, independent of self-perception and response bias.

Recently, Mayer and his colleagues (e.g., Mayer et al., 1999; Mayer et al. (in press)) have developed performance measures of EI that are comprised of a variety of rating tasks. The Multifactorial Emotional Intelligence Scale (MEIS) includes items tapping four constructs thought to comprise EI (emotion perception, assimilation of emotion into cognitions, understanding emotion, and managing emotion). This measure includes items that tap a total of 12 ability-related tasks.

Regarding the emotion-perception construct (or, in their terms, the emotion-perception *branch*), participants are asked to rate faces, music samples, designs, and emotionally laden stories in terms of the perceived presence of a variety of emotions. For instance, for the *faces* task, view several faces and rate the degree to which each face represents each of six basic emotions (e.g., happiness). Generally, performance is gauged by the degree to which participants' ratings match the ratings of other participants. These researchers found that the MEIS had adequate reliability and, further, that this measure predicted a variety of relevant outcome measures such as self-reported empathy and parental warmth.

The *Mayer Salovey Caruso Emotional Intelligence Test (V2.0; MSCEIT)* is similar to the MEIS in its structure. This measure is somewhat streamlined in that it contains eight (as opposed to 12) tasks. The same four constructs thought to underlie EI are tapped by this new measure. Mayer, Salovey, Caruso, and Sitarenios (in press) found that the MSCEIT may be appropriately conceptualized in terms of a one, two, or four factor solution. The one-factor solution proposes a general EI score, while the two-factor solution divides the scales into *experiential* and *strategic* components, while the four-factor solution breaks the items into the four *a priori* branches that these authors believe underlie EI.

### The Emotional Accuracy Research Scale (EARS)

While self-report measures of emotional intelligence may be limited in terms of their validity as predictors of emotional detection ability, they do have some advantages over laboratory measures of empathic accuracy. For instance, self-report measures are typically much easier to administer than laboratory measures. In addition, they are less dependent on a specific apparatus, such as video tape recorders, monitors, and computers.

Like the MEIS and the MSCEIT, the EARS (Mayer & Geher, 1996) was designed to afford the benefits of both self-report and laboratory measures of EI. In fact, the EARS was the intellectual predecessor of the *stories* task of the MEIS. The EARS consists of eight emotionally laden vignettes followed by pairs of mood items. Both the vignettes and the mood items were obtained from actual target individuals. Participants who complete the EARS are asked to read the vignettes and then choose mood items that they believe the target of the vignette reported feeling. For each participant completing the EARS, two emotion-perception scores are obtained. The *consensus agreement* score corresponds to the degree to which participants completing the EARS choose items that are most commonly chosen (weighted by the proportion of all participants who chose each item). The *target agreement*

score corresponds to the degree to which participants are able to accurately choose items that the targets of the vignettes endorsed themselves.

The EARS has the practical benefit of being easy to administer. Further, it seems to tap the actual (as opposed to self-reported) ability to accurately detect others' emotions. In the first study on the EARS, Mayer and Geher (1996) found the EARS to have demonstrated convergent validity (scores on the EARS were positively related to scores on Davis' (1983) measure of trait empathy) and discriminant validity (scores on the EARS were found to be negligibly related to SAT scores). Similar discriminant validation data have been described for the MEIS (Mayer et al., 1999). Mayer et al. describe a study in which they administered the MEIS and a modified version of the Army Alpha test of intelligence (Yerkes, 1921) to address the issue of discriminant validity. Similar to the results of Mayer and Geher (1996), these researchers found a moderate positive relationship between these measures of cognitive intelligence and EI, thus demonstrating some degree of discriminant validation for measures of the emotional perception component of emotional intelligence.

### **The Predictive Validity of Measures of Emotional Intelligence**

The primary study described here was designed to elaborate on past research on the EARS by assessing its predictive validity. While more recent measures of EI have demonstrated validity in several regards, such measures have not been examined vis a vis the ability to accurately perceive emotions in real-life actors in a dynamic, life-like setting. For instance, while the MEIS has been found to be related to empathy (Mayer et al., 1999), empathy in this context was operationally defined in terms of participants' self-reports of empathy. These researchers found the EI, as measured by the MEIS was moderately and positively correlated with self-reported empathy ( $r=.33$ ). This finding suggests that people who scored higher on the MEIS tended to agree with items such as "If someone is upset, I get upset too."

While it is important to demonstrate that indices of EI correlate with existing measures of related constructs, it seems that it would be most informative to demonstrate that such measures predict the ability to perceive others' emotions in a real-world, behavioral setting. The study described here was designed to assess whether scores on the EARS were predictive of accurate emotion detection in such a real-life situation. Additionally, the predictive power of the EARS was juxtaposed with the predictive power of two self-report measures of empathy to assess whether scores on performance measures such as the EARS do, indeed, better predict this ability.

## **METHOD**

A pre-selection procedure was designed to select participants who scored either extremely high or low on the EARS. This extreme-groups design was utilized as participants in the main part of this study were run individually in one-hour sessions. Participants were run individually to ensure that each participant had the opportunity to focus an optimal amount of attention on the stimuli that were presented. To assess the differential predictive value of self-report empathy measures versus the EARS on a performance-based laboratory

measure of emotional perception, participants first completed the EARS and two different self-report measures of empathy. Participants then engaged in a laboratory emotion-perception task.

## Participants

One hundred twenty four undergraduate students (52 males, 72 females) enrolled in an introductory psychology course at a large state university in New England were pretested on the EARS. Of these 124 potential participants, the 10 highest scoring males, 10 highest scoring females, 10 lowest scoring males, and 10 lowest scoring females ( $N=40$ ) were asked to engage in a subsequent laboratory test of emotional detection. For their participation, students received partial credit towards fulfillment of a requirement for their introductory psychology class.

## Materials

### *Preselection Materials*

For the preselection process, participants completed the EARS (Mayer & Geher, 1996) and two self-report empathy measures (Mehrabian & Epstein, 1972; Davis, 1983). In completing the EARS, participants first read eight emotionally laden vignettes written by other students. Then, for each vignette, participants were presented with a series of forced choice items that included a pair of mood-related terms. Participants completing the EARS were given the task of determining which item in each pair they believe the target reported feeling more strongly. Two scores were computed: (a) a *target* score corresponding to the degree to which participants were able to accurately choose items that the actual targets reported and (b) a *consensus* score corresponding to the degree to which participants' responses matched other participants' responses (weighted, for each item, by the proportion of participants who chose that item).

For a more detailed description of the EARS, including the items used in the EARS and scoring procedures, see Appendix A).

### *Laboratory Measures of Emotion Perception*

Participants who served as judges of others' emotions were selected for the second phase of the study completed several empathic accuracy tasks in a laboratory setting. This laboratory measure was conceptually similar to the EARS. However, instead of asking participants to read stories about targets and then rate their emotions (as in the EARS), in completing this laboratory measure, participants watched videotapes of targets who discussed emotionally laden topics. The video tape consisted of six graduate students (three males and three females), who talked about three situations that were affecting their moods at the time they were being taped. After describing each situation, the targets wrote down one adjective which they felt best captured their moods at that time. Finally, the targets completed the Present Reactions Scale (Mayer, Salovey, Gomberg-Kaufman, & Blainey, 1991); a mood scale that targets of the EARS also completed. In completing this scale, targets were asked to rate their moods along several broad dimensions.



### ***Scoring the Laboratory Measure***

After participants (i.e., emotion judges) observed the targets in a video format, four scores were calculated for the laboratory empathy measure for each participant; two scores corresponding to multiple choice items, and two corresponding to forced-choice items.

### ***Multiple Choice Items***

First, recall, that for each of three emotionally laden situations, targets were asked to write one word that best captured their moods. These emotionally laden words were juxtaposed with two similar words to create multiple choice items. For instance, if a target wrote "frustrated" after discussing an emotionally laden situation, "frustrated" was juxtaposed with "angry" and "depressed" to create a multiple choice item. As three such situations existed for each of six targets, a total of 18 multiple-choice items was created.

For each participant who watched the targets in the videos, two multiple choice scores were computed. *Target agreement* scores were computed by summing items for which participants' answers agreed with targets' answers. To compute *consensus agreement* scores, data regarding each multiple choice item from all participants were analyzed to determine the modal score for each item. Then consensus agreement scores were computed by summing instances in which participants chose this modal, or consensual, score.

### ***Forced Choice Items***

Forced choice items were computed with the same method used to compute target agreement and consensus agreement scores for the EARS. Specifically, for each target, 12 pairs of mood items from that target's Present Reactions Scale scores were juxtaposed. Pairings were based on the criterion that one item in each pair was rated by the target as describing his or her mood more accurately. Participants completing this measure were asked to make 72 (12 pairs of items \* six targets) judgments by determining which item in each pair the target reported feeling more strongly.

As with the multiple choice items, two scores were obtained for the forced choice measure. *Target agreement* scores were computed for each participant by summing instances in which that participant chose the alternative in each pair that the target reported feeling more strongly. *Consensus agreement* scores were computed by summing instances in which participants agreed with the item within each pair that was chosen by a majority of other participants.

### ***Procedure***

Participants who fit the criteria for inclusion in the laboratory phase of this study were run individually. After informed consent was obtained, participants were given an answer sheet that included all the multiple choice and forced choice items from the laboratory measures. Participants were asked to watch the videotapes and to try to best assess the emotions of the targets. After each of the 18 emotionally laden segments of the videotapes was presented, the tape was paused for 30 seconds. During this time, participants were asked to complete the multiple choice item corresponding to the appropriate segment of video. During the six segments corresponding to the completion of a particular target (i.e., after each target's third emotionally laden segment), the video was paused for three minutes. During this time, participants responded to the 12 pairs of mood items for a given target by indicating the item within each pair they believed the target reported feeling most strongly across all three

emotionally laden situations. Participants repeated this process for all six target subjects in the videotape. One whole session lasted approximately 55 minutes. When participants were finished, they were thanked and thoroughly debriefed.

## RESULTS

The primary hypothesis of the present study is that performance-based measures of emotion perception should be stronger predictors of behavioral indices of emotional intelligence than standard self-report measures of trait-empathy. Consequently, the analyses performed assessed how well two scores on a performance measure of emotion perception, the EARS (corresponding to both consensus and target agreement), predicted scores on a behavioral measure. Further, for comparison purposes, analyses addressed how well the five scores on the self-report measures of empathy predicted scores on these same dependent variables.

Analyses were performed in the following order: First, to assess the internal reliability of the different measures employed, a Cronbach alpha was computed for each measure. Next, intercorrelations among the predictor variables were determined to assess the degree of shared variance between these variables. Also, to address the primary hypothesis (that scores on the performance-based test of emotion perception should be more strongly associated with behavioral measures of emotion perception than self-report empathy measures), zero-order correlations between all predictor variables and all dependent variables were computed. Additionally, four multiple regressions were performed to assess how well each variable independently predicted empathy in the laboratory.

### Conceptual Issues Surrounding Scoring

In this research, the EARS and the laboratory empathy measures yield scores pertaining to Target and Consensus criteria. As is indicated in the Method section, these criteria offer conceptually different indices of emotional intelligence. Target scores generally speak to whether individuals tend to agree with targets' self-reported emotions while consensus scores assess whether individuals agree with others' judgments regarding the targets' emotions. While target scoring of this sort has been employed in previous research (e.g., Ickes et al., 1990), such a scoring procedure has potential problems. Specifically, targets who produce "correct answers" for such measures may employ response biases and/or they may not accurately know themselves. Consensus scoring procedures, on the other hand, are based on aggregated judgments and, as such, tend to be more reliable. It is important to note that these criteria are mathematically and conceptually different when considering this research.

### Instrument Reliabilities

Cronbach's alpha coefficients were calculated to determine the internal reliabilities of the EARS and of the laboratory measures of empathy. Previous research on the EARS reported relatively low alpha coefficients. Specifically, Cronbach's  $\alpha$  for target agreement was .24

while Cronbach's  $\alpha$  for consensus agreement was .53 (Mayer & Geher, 1996). In the current research, an item analysis was conducted for both EARS scores to improve these reliabilities. After deleting items that reduced alpha for both target and consensus agreement scores, alphas increased considerably to .75 and .80 respectively. A similar item analysis was conducted for the four laboratory measures. After deleting items that reduced internal reliability, the four laboratory measures achieved reasonable levels of reliability. For the multiple choice items, alpha for target agreement was .52 and alpha for consensus agreement was .55. For the forced choice items, alpha for target agreement was .86 and alpha for consensus agreement was .77. While the reliabilities for the multiple choice items are somewhat low, they are considerably higher than the initial reliabilities for the EARS.

It should be noted that the preselection procedure in the current research entailed creating extreme groups based on EARS scores. Overall, this selection process was chosen to optimize the EARS' predictive validity. However, this process may have affected instrument reliabilities of the measures employed. As such, the reliability estimates of the laboratory measures may be somewhat inflated.

## Descriptive Statistics

The means and standard deviations for all predictor and criterion variables are reported in Table 1. Recall that the EARS yields two separate scores differentially representing consensus and target agreement. Similarly, the two kinds of items on the laboratory empathy measures (multiple-choice and forced-choice pairs) each yield consensus and target agreement scores. For each of these three measures (the EARS and two laboratory empathy measures) the mean for consensus scoring was greater than the mean for target scoring. For the EARS, the mean for target scoring (i.e., the average number of items for which the subject agreed with the target as to which alternative the target felt more strongly) was 33.45 ( $SD=6.93$ ), whereas the mean for consensus scoring (i.e., the average number of items for which the subject agreed with the group consensus as to which alternative the target felt more strongly) was 40.30 ( $SD=6.63$ ). A dependent measures t-test revealed that this difference was statistically significant ( $t(39)=4.56$ ,  $p<.01$ ). Similarly, for the two kinds of items on the laboratory empathy measure, the mean for consensus scoring was found to be significantly greater than the mean for target scoring (See Table 1).

**Table 1.** Means and Standard Deviations of the Predictor and Criterion Variables

	M	SD
<b>Predictor variables</b>		
	33.45 <sub>a</sub>	6.93
	40.30 <sub>b</sub>	6.63
	120.33	11.63
	29.53	3.67
	25.85	5.00
	20.23	5.44
	26.95	4.18
<b>Criterion variables</b>		
	9.70 <sub>a</sub>	1.38
	11.03 <sub>b</sub>	1.69
	26.48 <sub>a</sub>	7.60
	35.08 <sub>b</sub>	5.39

<sup>a</sup> (Mehrabian & Epstein, 1972).

<sup>†</sup> (Davis, 1983).

<sub>ab</sub> significant differences between consensus and target agreement scores for a given measure.

### Intercorrelations among the Predictor Variables

The predictor variables in the present study included two scores for the EARS (separate consensus and target agreement scores), one global measure of trait empathy (Mehrabian & Epstein, 1972), and four self-report measures of more specific kinds of empathy from Davis' (1983) empathy scale including the self-reported ability to take the perspective of others, the tendency to be empathically concerned about others, the inclination to engage in fantasy, and the propensity to be overly concerned with one's own personal distress. In general, both target and consensus scores on the EARS were found to be uncorrelated with scores on the five trait measures of empathy (See Table 2). Interestingly, the target and consensus scores on the EARS were found to be uncorrelated with each other as well ( $r(40) = .02$ , *ns*). In other words, the degree to which participants agreed with the targets on the EARS was found to be unrelated to the degree to which participants agreed with the group consensus. This result was also found in previous research on the EARS and will be addressed in the Discussion. Several of the self-report measures of empathy were found to be significantly related to each other in theoretically meaningful ways. For instance, scores on the empathic concern scale were positively related to scores on the general empathy scale ( $r(40) = .55$ ,  $p < .01$ ).

Table 2. Intercorrelations among Predictor Variables

	EARS		Empathy variables			
	Target score	Consensus score	General empathy <sup>†</sup>	Empathic concern <sup>§</sup>	Personal distress <sup>§</sup>	Perspective taking <sup>§</sup>
<i>EARS</i>						
Target						
Consensus	.02					
<i>Empathy</i>						
General empathy	-.21	-.02				
Empathic concern	-.12	-.11	.55 <sup>†</sup>			
Fantasy	-.07	.02	.39 <sup>*</sup>	.28		
Personal distress	-.54 <sup>†</sup>	.05	.30	-.03	-.02	
Perspective taking	.18	-.06	.24	.40 <sup>†</sup>	.22	-.18

<sup>\*</sup>  $p < .05$

<sup>†</sup>  $p < .01$

<sup>‡</sup> (Mehrabian & Epstein, 1972).

<sup>§</sup> (Davis, 1983).

### Zero-order Correlations between the Predictor and Dependent Variables

To address the primary hypothesis that the EARS should better predict behavioral indices of emotional intelligence than self-report empathy measures, zero-order correlations were computed between the predictor variables and four criterion variables. The criterion variables involved in the present analyses include scores from the three-item multiple choice section of the laboratory empathy measure scoring for both target and consensus agreement as well as scores from the forced-choice pairs section of the laboratory empathy measure scoring for both target and consensus agreement.

The zero-order correlations were computed to provide a general sense of the relationship between the different predictor measures of emotional intelligence and the behavioral measures of emotional intelligence. It is noteworthy that none of the 20 correlations between the trait measures of empathy and the four laboratory empathy scores were significant (See Table 3). These correlations ranged from  $r(40) = -.28$  to  $.28$ .

The pattern of correlations between scores on the EARS and the criterion variables revealed generally positive relationships between the EARS and the behavioral indices of emotional intelligence. Of eight correlations computed between scores on the EARS and scores on the criterion variables, three were positive and significant (See Table 3). Specifically, these correlations included the correlation between consensus scores on the EARS and target scores on the three-item multiple-choice section of the laboratory empathy measure ( $r(40) = .37$ ,  $p < .01$ ), the correlation between consensus scores on the EARS and consensus scores on the three-item multiple-choice section of the laboratory empathy measure ( $r(40) = .36$ ,  $p < .01$ ), and the correlation between the target scores on the EARS and the target scores on the forced-choice pairs section of the laboratory empathy measure ( $r(40) = .43$ ,  $p < .01$ ). Of the other five correlations computed between scores on the EARS and scores on the laboratory empathy measure, four were non-significant while one significant negative correlation existed between consensus scores on the EARS and target scores on the forced choice section of the laboratory empathy measure ( $r = -.41$ ,  $p < .01$ ).

**Table 3.** Zero-order Correlations between the Predictor and Criterion Variables<sup>1</sup>

	Laboratory Empathy Measures			
	Multiple Choice Items		Forced Choice Items	
	Target Score	Consensus Score	Target Score	Consensus Score
<i>Predictor variables:</i>				
EARS				
	-.04	.05	.43 <sup>†</sup>	-.09
	.37 <sup>†</sup>	.36 <sup>†</sup>	-.41 <sup>†</sup>	.15
Empathy				
	.23	.25	-.06	.15
	-.22	-.28	.13	-.07
	.07	.06	-.03	.28
	.07	.18	-.21	.12
	-.19	-.01	.05	-.03

<sup>1</sup>Note that target and consensus scoring procedures are conceptually, mathematically, and empirically distinct.

\*  $p < .05$

†  $p < .01$

This negative correlation is inconsistent with predictions regarding the EARS relationship with the laboratory empathy measures. This troubling correlation may derive from several sources. This correlation is between a target index of the laboratory measure and a consensus index of the EARS. It may be that consensus measures and target measures are very different conceptually, and, accordingly, relationships between such measures may not always be positive in nature. Further, this negative correlation may derive from Type I error. In any case, while this finding may be meaningful, it is inconsistent with the basic pattern of results. In general, scores on the EARS were positively correlated with scores on the laboratory empathy measure while scores on the trait measures of empathy were uncorrelated with scores on the laboratory empathy measure, supporting the hypothesis that a performance-based measure of emotional intelligence should be more strongly associated with behavioral measures of emotional intelligence than self-report measures.

### Multiple Regression Analyses

To address the question of how well scores on the predictor variables differentially predict scores on the criterion (laboratory empathy measure) variables, four separate multiple regression analyses were conducted. In each of these analyses one of the four laboratory empathy measure variables was used as the criterion variable while the two EARS scores (consensus and target agreement) and the five self-reported empathy scores served as predictor variables. To treat each predictor variable equally, a standard multiple regression was performed (i.e., all variables were entered on the same step). Thus, in each analysis, each predictor variable was assessed in terms of how well it independently accounted for variance in the dependent variable taking shared variance with other variables into account.

In one of these four multiple regression analyses, no combination of predictor variables accounted for a significant amount of variance in the dependent variable. This particular

analysis employed the consensus-agreement forced-choice score from the laboratory empathy measure as a criterion variable.

In each of the other three regression analyses, a significant amount of variance in the criterion variable was accounted for by some combination of predictor variables. The first of these analyses used target-agreement forced-choice scores as the criterion variable. Of the seven predictor variables in the present analysis, only two contributed significantly to the prediction of scores on the dependent variable. Consistent with the primary hypothesis of this study, these variables were the consensus and target-agreement EARS variables. None of the self-report empathy variables contributed significantly to predicting scores on the dependent variable. The most significant predictor (in terms of the semi-squared partial correlations ( $sr^2$ ), which represents the amount of unique variability in the criterion variable predicted by a specific predictor variable) was the target-agreement EARS variable ( $sr^2=.18$ ) followed by the consensus-agreement EARS variable ( $sr^2=.16$ ). Altogether, 40% (27% adjusted) of the variability in this dependent variable was accounted for by scores on the target and consensus-agreement scores of the EARS.

In the remaining two analyses, scores from the consensus EARS scale, the Davis empathic concern scale, and the Mehrabian and Epstein trait-empathy scale all contributed significantly in accounting for variability in the laboratory empathy measure multiple-choice scale scoring for *both* target and consensus agreement.

**Table 4.** Multiple Regression Predicting Target-Agreement for Multiple-Choice Laboratory Items from Two EARS and Five Trait-Empathy Variables<sup>1</sup>

Independent Variables	B	b	$sr^2$	t
<i>EARS</i>				
	-.01	-.06	.00	-.34
	.07	.33	.11	2.33*
<i>Self-report empathy measures</i>				
	.07	.55	.17	2.88†
	-.16	-.44	.11	-2.35*
	-.00	-.00	.00	-.03
	-.05	-.18	.02	-1.01
	-.05	-.15	.02	-.91
			$R^2=.36$	
			$R^2$ (adjusted)	
			= .22	
			$R=.60$	

<sup>1</sup> Note that target and consensus scoring procedures are conceptually, mathematically, and empirically distinct.

<sup>2</sup> (Mehrabian & Epstein, 1972); 3 (Davis, 1983).

\*  $p<.05$

†  $p<.01$

The results from the analysis employing the laboratory empathy measure multiple-choice scale scoring for target agreement are summarized in Table 4. The Multiple  $R$  was significantly different from zero  $F(7, 32)=2.60, p<.05$ . The results from the analysis employing the laboratory empathy measure multiple-choice scale scoring for consensus agreement were virtually identical to the results from the previous analysis described in Table

4. For this final regression analysis, a significant amount of variability was explained by the set of predictor variables ( $R=.64$ ;  $F(7, 32)=3.24$ ,  $p<.05$ ). The three significant predictor variables in this analysis were the EARS scale scoring for consensus agreement, the Davis empathic concern scale, and the Mehrabian and Epstein trait-empathy scale.

## DISCUSSION

Emotional intelligence has garnered much attention both in academic (e.g., Salovey & Mayer, 1990) and popular (e.g., Goleman, 1995) domains. Some researchers have found that emotional intelligence is predictive of a variety of positive outcomes (e.g., Martinez-Pons, 1997). However, valid and accepted measures of emotional intelligence are still needed for researchers in this field. The development of such measures should, potentially, allow researchers to develop a more coherent conceptualization of emotional intelligence. The EARS was designed to be such a measure.

Several self-report measures of emotional intelligence have been developed (e.g., Schutte et al., 1998). These measures have demonstrated validity. However, as they are self-report measures, they are somewhat limited. For instance, self-report measures of emotional intelligence may be particularly subject to response biases. The EARS, on the other hand, is a performance measure of the empathic component of emotional intelligence. Being a performance measure gives the EARS the benefit of tapping an actual ability while, concurrently, being relatively free of response bias.

In previous research (Mayer & Geher, 1996), the EARS demonstrated both convergent and discriminant validity. However, that previous study did not adequately address the internal reliability nor the predictive validity of the EARS. The current research primarily addressed these issues. After completing a careful item analysis, *target agreement* and *consensus agreement* scores for the EARS increased from .24 to .75 and from .53 to .80 respectively. Thus, this revised version of the EARS is clearly more internally reliable.

Further, the current data suggest that the EARS has demonstrated reasonable predictive validity. Scores on the EARS were generally more predictive of empathy in the laboratory task compared with scores on the self-report empathy measures. Thus, the EARS seems to tap how well people can understand the emotions of others in a context that includes information across several sensory modalities (including both visual and auditory stimuli). These data suggest that the EARS is predictive of emotional detection in a relatively real-life setting.

The current data also raise questions regarding the differential importance of consensus agreement and target agreement as criteria of the empathic component of emotional intelligence. The multiple regression analyses in the current research have several implications regarding distinctions between consensus agreement and target agreement. Except for in the first such analysis, where none of the predictor variables accounted for a significant amount of variability in the criterion variable, the EARS consensus agreement score consistently appeared as a significant predictor of the different criterion variables. Additionally, in predicting one of the target agreement variables, both the target-agreement and consensus-agreement EARS scales accounted for a significant amount of variability in the criterion variable. Further, in predicting the multiple-choice scale scoring for target agreement and in predicting this same scale scoring for consensus agreement, a combination



of the EARS consensus-agreement score and the two general empathy trait scores accounted for significant amounts of variability in the criterion variables.

These findings imply that the ability to agree with a consensus as to an individual's emotions may be beneficial in understanding emotions from the perspectives of both actors and observers. However, the ability to accurately assess an actor's own self-report seems to be, in general, less related to understanding one's emotion from the perspective of both actor and observer. Additionally, the tendency for a person to report that he or she is an empathic person may be related to understanding the emotions of a given target in terms of both how that target describes his or her own emotion as well as how other observers judge that emotion. Thus, self-reported measures of emotional intelligence clearly have demonstrated utility. However, as is most clearly indicated by the zero-order correlations reported in Table 3, this tendency to self-report high levels of empathy seems to be, in general, less related to accurate emotion detection than the tendency to score high on the EARS.

The fact that the EARS was more predictive of the performance-based emotion perception measures than the self-report empathy scales may, to some degree, result from shared method variance. Importantly, the fact that some of the measures in this research are performance-based and some of the measures in this research are self-report-based is by design. The fundamental impetus for this research, initially, was a skepticism regarding self-report empathy measures as predictors of relevant behaviors (such as being able to actually discern others' emotions). For such self-report measures, which are very widely employed, to be useful they would need to demonstrate predictive validity. In other words, for such measures to be useful, they would need to, ultimately, predict scores on some performance-based dependent measure. Accordingly, the shared-method variance between the EARS and the dependent measures in the current research exists by design. The main point here is that performance measures of the emotion-discernment component of empathy can be just as easy to administer as self-report empathy measures with the added luxury of actually predicting how well people can perceive others' emotions. Ultimately, the reasoning goes, work in this field would benefit from researchers using performance-based measures of constructs such as emotion perception whenever possible.

The findings in the current research also speak to the differential *meaning* of target and consensus scoring criteria. The correlation between target scores on the EARS and Consensus scores on the EARS was .02. This finding suggests that the ability to know a person as he or she knows him or herself and the ability to agree with others as to what a target is feeling are largely independent dimensions.

This finding, that emotion perception based on target scoring and emotion perception based on consensus scoring are somewhat independent, seems to pertain to Nisbett and Wilson's (1977) work on our ability to report on our own mental processes. Nisbett and Wilson argue that, more often than they believe, people are not privy to the actual reasoning underlying their decisions. To the extent that this ignorance regarding people's own mental processes exists, it makes sense that knowing what an individual thinks he or she feels may be quite different from knowing what others think that same individual feels. In other words, to the extent that humans are not adept at knowing the actual reasoning underlying their behaviors, it makes sense that consensus and target measures would yield different kinds of results and, further, that such measures would be orthogonal to one another.

With regard to target and consensus scoring, both kinds of scoring procedures have their costs and benefits. Target scoring makes sense in that it theoretically gets at one of the

fundamental aspects of emotional intelligence: How well does one know the feelings of another. However, target scoring has its problems. Targets may not be accurate reporters of their own emotions due to response biases and/or poor self-understandings. Consensus scoring procedures, on the other hand, are more reliable. Further, in the current research, the EARS consensus score was most predictive of other indices of emotional intelligence suggesting that consensus scoring may be more useful than target scoring as well. This pattern, for consensus scores to be more reliable predictors of relevant outcomes, matches other research in this area which also has found consensus scores to be (a) more internally reliable and (b) more predictive of relevant criterion variables (e.g., Mayer et al., in press).

## Future Research

While the current research demonstrates internal reliability and predictive validity of the EARS, this research also raises questions for future research. For instance, these findings are somewhat unclear regarding the relationship between target and consensus agreement. In the current sample, target agreement and consensus agreement were uncorrelated. In other words, being able to accurately assess a target's self-report may be unrelated to being able to accurately assess others' perceptions of that target's self-report. This finding suggests two very different abilities that may underlie emotional intelligence.

However, somewhat surprisingly, the target and consensus scores of the EARS were not predictably related to target and consensus scores of the laboratory empathy measure. Specifically, target agreement on the EARS was not consistently related to the target empathy scores while, similarly, consensus agreement on the EARS was not consistently related to consensus empathy scores (See Table 3). For instance, while consensus agreement on the EARS was positively related to target agreement on the multiple choice section of the laboratory measure, target agreement on the EARS was uncorrelated with this same criterion variable. Further research needs to more specifically address the differential correlates of target agreement ability and consensus agreement ability.

One potentially relevant new development speaking to this issue of the different kinds of criteria used in EI research pertains to the inclusion of what Mayer et al. (in press) call an *expert criterion*. In a recent study, they had two sets of participants complete the MSCEIT. These classes of participants included over 2,000 college students and 21 experts on emotion. In terms of some indices, using the emotion experts' responses as criteria of *correct* responses yielded more reliable results. As these authors indicate, more research examining the differential patterns of correlates between general consensus criteria and expert criteria warrant further research. Given the extensive data pertaining to the EARS, a study examining differential correlates of the EARS based on general consensus versus expert criteria may be useful in elucidating the empirical distinctions between these two criteria.

A final implication of the work outlined here speaks to the utility of including narrative-based tasks in measures of EI. While the MEIS included the *stories* task, which is quite similar to the EARS, the more current MSCEIT does not include such a narrative task. Partly, the omission of such a task in the MSCEIT derives from the fact that such tasks are quite verbal and, as such, it is difficult to separate variability in responses pertaining to EI versus verbal ability. However, the current data provide evidence that the EARS is predictive of emotion perception in a life-like setting. This information is highly relevant and suggests that

future EI researchers would benefit from creating measures similar to the EARS in their conceptual underpinnings and in their format. Similarly, the evidence for the predictive validity of the EARS may serve as a model for validation of other EI measures. Validating other measures in the context of the degree to which such measures predict actual responding in real-life settings would be useful in bolstering the utility of such measures.

## REFERENCES

- Bar-On, R. (1997). The emotional quotient inventory (EQ-i): Technical manual. Toronto: Multi-Health Systems.
- Batson, C.D., & Coke, J.S. (1983). Empathic motivations of helping behavior. In J.T. Cacioppo and R.E. Petty (Eds.), *Social psychophysiology: A sourcebook* (pp. 417-433). New York: Guilford.
- Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J.M. Haviland-Jones (Eds.), *Handbook of Emotions*. New York: Guilford.
- Crowne, D.P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Darwin, C. (1872). *The expression of emotions in man and animals*. Chicago: University of Chicago Press, 1965.
- Davis, M.H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126.
- Ekman, P. & Friesen, W.V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17, 124-129.
- Geher, G. (1994). *Assessing the validity of a scale designed to measure empathic accuracy*. Paper submitted as partial requirement for the Master's degree, University of New Hampshire, Durham, NH.
- Geher, G., & Mayer, J.D. (1998). *Emotional accuracy research scale and scoring manual*. Previously unpublished manuscript.
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam Books, Inc.
- Ickes, W., Stinson, L., Bissonette, V., & Garcia, S. (1990). Naturalistic Social Cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 54, 730-742.
- Levenson, R.W., & Ruef, A.M. (1992). Empathy: A physiological substrate. *Journal of Personality and Social Psychology*, 63, 234-246.
- Martinez-Pons, M. (1997). The relation of emotional intelligence with selected areas of personal functioning. *Imagination, Cognition, and Personality*, 17, 3-13.
- Mayer, J.D., Caruso, D., & Salovey, P. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27, 267-298.
- Mayer, J.D. & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, 22, 89-113.
- Mayer, J.D., Salovey, P., & Caruso, D.R. (1999). *The Emotional IQ test*. [on-line]. Available: <http://www.virtent.com/eq.html>.
- Mayer, J.D., Salovey, P., & Caruso, D.R. (2000). Models of Emotional Intelligence (pp. 396-420). In R. Sternberg (Ed.), *Handbook of intelligence*. Cambridge, UK: Cambridge University Press.

- Mayer, J.D., Salovey, P., Caruso, D.R., & Sitarenios, G. (in press). Measuring emotional intelligence with the MSCEIT V2.0. Emotion.
- Mayer, J.D., Salovey, P., Gomberg-Kaufman, S., & Blainey, K. (1991). A broader conception of mood experience. *Journal of Personality and Social Psychology*, 40, 100-111.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, 40, 525-543.
- Nisbett, R.E., & Wilson, T.D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Rosenthal, R., Hall, J.A., DiMateo, M.R., Rogers, P.L., & Archer, D. (1979). *Sensitivity to non-verbal communication: The PONS test*. Baltimore: Johns Hopkins University Press.
- Salovey, P. & Mayer, J.D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, 9, 185-211.
- Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25, 167-177.
- Wise, L.G. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of Personality and Social Psychology*, 50, 314-321.
- Yerkes, R.M. (1921). Psychological examining in the United States. *Memoirs of the National Academy of Sciences* (15:890 pp. entire).

## APPENDIX A.

### THE EMOTIONAL ACCURACY RESEARCH SCALE AND SCORING INFORMATION (GEHER & MAYER, 1998)

Below are eight people's descriptions of situations which caused them to feel various emotions. The stories that they wrote were edited in order to omit any personally identifying information. Each description is followed by twelve pairs of words or phrases. In each case, the person who wrote the description reported that he or she felt one way more than the other. For example, the pair "My teeth clenched – Challenged" might follow a description. Your task is to guess which word or phrase best describes how the person who wrote the description reported feeling. For instance, in this example, if you felt the author of the description reported feeling angry, you would circle "My teeth clenched" for that pair. Please read each vignette and circle the word or phrase in each pair that you think the author of the vignette reported feeling most strongly. Try to incorporate all of the situations in a description when deciding on your answers.

#### Vignette #1: Written by a Twenty-year-old Male

I studied for an hour and a half before this experiment for a biology test I have tomorrow because I have procrastinated on all the work I have to do in this class. My test is tomorrow and I know little to nothing about the material covered and I have a short time to learn it. Because of this situation I am tense about how much I have to do and pissed off that I didn't do most of it before today.

I took a bus and came to campus today. I spent all day on campus going to boring classes and trying to get some work for biology done. I am psyched to go home after this experiment.

This friend of mine said she'd cook dinner for me ages ago. Kathy came up to me in the library and asked if she could come to my apartment and cook dinner. I'm happy 'cuz this girlfriend of mine said she'd cook me and my roommates dinner, but I don't see how I'll get any work done.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations he described.

1. Nervous on another's behalf – Shield myself from attack
2. Delighted – Go with my feeling
3. Chuckling – Could kill someone
4. Attacked – Feeling is exploding
5. Powerful – Happy for someone else
6. Swear at someone – Withdraw
7. My patience is running out – Rewarded
8. Act as if no problem – Angry
9. Hide – Give approval
10. Alone – Do nice things
11. Distance self from feeling – Apart from others
12. Happy along with another – Terrified

### **Vignette #2: Written by a Forty-year-old Female**

I signed up for a history telecourse as a non-traditional student. Tonight is exam night and I'm nervous because I want to do well. At the same time, I signed up for this experiment and what I really should be doing is studying for the exam. This has made me both nervous and anxious because I feel unprepared for the exam. When I came for this experiment, I was greeted by a friend and another person I know from a course of mine last semester. Right now, I feel happy to see them both. My friend's a great person and it's good to know he's doing well here at the university. I see my other friend every day and I am always glad to have someone to talk to.

Someone in my office went behind my back and said things about me which forced my supervisors to confront me. She "ratted on me" instead of coming directly to me and talking it out, and that made me angry. Finding out she did that makes me distrust her and I always feel I have to watch every move I make so I don't get "stabbed in the back." I feel resentful towards her.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations she described.

1. Hair on end – Appreciated
2. Act as if no problem – Chuckling
3. Angry – Lively

4. Rewarded – Attacked
5. My patience is running out – Dared
6. My fists clenched – Sharing another's resentment
7. Could kill someone – Sharing another's joy
8. Pretend everything is okay – Furious
9. Admired – Trembly voice
10. Go with my feelings – Terrified
11. Nervous – Open to my feelings
12. Hostile – Smiling

### **Vignette #3: Written by a Nineteen-year-old Male**

This situation was created by two other tests and a paper that was due yesterday. I wanted to do well on the test that didn't matter as much as the other except for reasons of pride. I wanted to do better than my friends on the test so I studied more for it and less for the more important test. One of my friends did better than me on the test I studied hard for. This hurt my pride and I am currently frustrated right now.

I had a heavy schedule of classes today. After two days of taking tests and writing a paper I now had to face four hours of classes in one day. Since I am nearly finished with the long day, I feel relieved.

I heard an announcement for a party on Thursday night. Since I have had a rough week I am looking forward to partying and Thursday is a great day for me. This situation brought about anticipation of a fun event.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations he described.

1. Lively – Could kill someone
2. Share feelings with others – Sharing another's joy
3. Dance around – Angry
4. Delighted – Happy
5. Furious – Distance self from feeling
6. Kick something – My fists clenched
7. Hostile – Be by myself
8. Praised – Attacked
9. Mad – Go with my feelings
10. Trembling – Celebrate
11. Healthy – Hair on end
12. Chuckling – Act as if no problem

### **Vignette #4: Written by a Twenty-year-old Female**

My best friend's mother died this weekend. She had diabetes for a long time, and as she got older her health grew worse and worse. I went to her funeral on Saturday. Many of my

friends from high school were also there because we all wanted to be there for our friend and because we all knew and liked her mother. It made me realize how lucky I am to have younger, healthy parents when I saw my friend standing there crying. Just watching her huge family come pouring into the church also made me sad.

My roommate has been kind of blowing off her boyfriend. She told him she did not want to see him until Spring Break. He is hurt because he thinks she does not like him anymore, and he wants to come up here to see her this weekend. I have been gone almost every weekend since school started, giving her plenty of opportunities to have him up here while I am gone, and now I'm finally getting to stay here for the weekend and he might be coming up. (Why can't she go visit him instead?!?!)

I got to know this girl a little because she was in one of my classes. We happened to be talking about housing for next year and I mentioned wanting to live in the Knollwood apartments. She asked me if I would like to try to get into Knollwood with her and some of her friends. Without giving it very much thought, I said "yes." I gave it more thought and I am not sure if I can handle living with a smoker. I don't like where I'm living this semester, and if I get stuck somewhere I don't like next year too, I'm GOING TO GO CRAZY.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations she described.

1. Be by myself – Kick something
2. Stomping feet – Alone
3. Pretend everything is okay – Threaten a fight
4. Angry for someone else – Help a friend
5. Evade feeling – Defiant
6. Sharing another's anger – Threatened with death
7. Hostile – Unhappy for another
8. Fearful – Apart from others
9. Cheated – My teeth clenched
10. Withdraw – Scared for someone else
11. Attacked – Isolate myself
12. Mad – Delighted

### **Vignette #5: Written by a Fifty-year-old Female**

This winter there have been weeks of gray skies, mounds of snow, ice, cold. The February blues sink in and as I haven't been through a "normal" winter for a few years, the grayness of this winter and its precipitation have been hard to take. However today, the sun came out and my friendly shadow appeared. This brings out a "sunnier" disposition in me. I'm not so alone when my shadow follows me.

I am overweight and scared of the thoughts of high blood pressure, arthritis, and coronary disease. I don't like hurting when I do physical activity. I attended an orientation session for medically supervised loss of weight. It seems effective but costly. I am questioning my commitment to the program. The orientation answered all my questions. I feel hopeful it will

work and I can do it. With 'trust' and 'hope' being big issues for me, it feels good to feel them both at the same time.

Dee, Lillith, and I were all in the same class last semester. Dee is a fine young man and I enjoyed his perceptions of life immensely last semester. Dee showed up at this lab. It was good seeing him and chatting about our grades and what he is up to this semester. Dee chose to sit next to me at the lab and discuss our class last semester. As a non-traditional student, it made me feel wonderful that he chose to interact with me again. I guess we have a good rapport.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations she described.

1. Lively – Attacked
2. Kick something – Open to my feelings
3. Evade feelings – Glowing face
4. Act as if no problem – Dance around
5. Powerful – Happy for someone else
6. 6 Shield myself from attack – My patience is running out
7. Rewarded – Share feelings with others
8. Show feeling – joyful
9. Happy along with another – Stomping feet
10. Mad – Alone
11. Nervous – Dared
12. Defiant – Be by myself

### **Vignette #6: Written by a Nineteen-year-old Female**

I just got out of my political science class now where we had to present a panel discussion. There were six of us who discussed abortion. Three for it, three against it saying it was immoral. We presented our cases and then opened it to comments and questions from the audience. This one guy said straight out that he thought our discussion was a waste of time and we had done no real research. This made me so mad because I think we did an incredible job and we've been working on it for weeks.

I started dating this guy Bob a few months ago. Everything has been incredible, we're both in love. But my friends feel I'm blowing them off. Bob lives on the same floor as me in an apartment building, we are always together. My friends have been ignoring me and when they have confronted me, they say I have to change. So lately I've been doing what they say I should, but I feel they aren't meeting me half way. I also feel I've lost trust in them.

I'll have been dating Bob four months as of tomorrow! We are always together because we enjoy each other. But we also respect each other's space. We have really been talking about our future together lately!

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations she described.



1. Share feelings with others – Trembly voice
2. Defiant – Scared
3. Act as if no problem – Frightened for another
4. Fearful – Feeling is exploding
5. Kick something – Alone
6. Admired – Happy along with another
7. Unhappy for another – Do nice things
8. Angry – Lively
9. Appreciated – Nervous on another's behalf
10. Apart from others – My teeth clenched
11. Withdraw – Powerful
12. Scared for someone else – Could kill someone

### **Vignette #7: Written by a Nineteen-year-old Male**

I have had a stressful week and I have had very little sleep or time to relax. I am stressed out about tests and work that has to be completed. I feel a little depressed because I have a feeling I failed two tests I took in the last two days. My day is also never ending with one thing after the other that has to be done.

Someone was harassing my girlfriend. A guy was hitting on my girlfriend and trying to get some play off her. I got a hold of him and told him to lay off because it was upsetting her and pissing me off. She and I talked about the situation last night and I still have this in mind. This may also be adding to my stressful week.

I have been trying to decide whether or not to go camping this weekend. I would like to go camping but I can't drive my car because I have no insurance to cover anyone other than myself. I keep thinking about it and I can't decide whether or not to go to stay here at the university for the weekend. I am feeling pressure from right and wrong.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations he described.

1. Withdraw – Fearful
2. Act as if no problem – Lively
3. Ominous threat – Sharing another's resentment
4. Frightened for another – Mad
5. Happy along with another – Defiant
6. Be by myself – Alone
7. Dared – Isolate myself
8. Threatened with death – Hair on end
9. Chuckling – Angry for someone else
10. Furious – Open to my feelings
11. Pretend everything is okay – Scared for someone else
12. Threaten a fight – Mad on another's behalf

**Vignette # 8: Written by a Nineteen-year-old Male**

I want to go to the city to visit my friends. I was talking to my friend who goes to a college there, and she said she wanted me to come visit her. There is nothing sexual, we are just good friend.

I am a nice guy, I'm not just blowing my own horn, but that is what all of my friends tell me. I have plenty of friends, both male and female, but sometimes I want a female friend to be a little more than just a friend. My lab partner has a girlfriend, he goes home to visit her every weekend, then comes back and tells me about what they did. I have told him about my situation, but he persists in bringing it up. It's probably nothing, but I would like a girl that will be more than a best friend. I know it's just this age, but I'm really horny. I frustrates me that jerks, like my lab partner, have someone in their lives, but I don't.

I have a sociology exam tomorrow. I haven't started studying yet, but for some reason I don't care. I want to do well in the class, but I really could care less about studying. I don't want to fail. If I don't put forth an effort, it is like I didn't try. If I study my ass off and I still fail I feel worse. I don't have an excuse.

For each of the twelve pairs below, choose the word or phrase within that pair which best describes the reported feelings of the person who wrote the above passage across all of the situations he described.

1. Could kill someone – Apart from others
2. Sharing another's resentment – Defiant
3. Pleased on another's behalf – Help a friend
4. Act as if no problem – Sharing another's joy
5. Give approval – My patience is running out
6. Go with my feelings – Delighted
7. Smiling – Mad on another's behalf
8. My teeth clenched – Distance self from feeling
9. Attacked – Apart from others
10. Angry – Praised
11. Be by myself – Hostile
12. Joyful – Trembling

**EARS Scoring Information**

The Emotional Accuracy Research Scale (EARS) was designed as a performance measure of the emotion-perception component of emotional intelligence.

Following are two distinct scoring methods. One of these methods corresponds to *Target agreement* (i.e., the degree to which participants agree with specific targets regarding the emotion the target reported feeling most strongly) the other corresponds to *Consensus agreement* (i.e., the degree to which participants' judgments of targets' emotions corresponded to other participants' judgments for those same targets).

Notice that several items are marked with asterisks (\*). We recommend that you omit these items when computing total scores; these items have been found to decrease scale reliabilities.

The EARS is comprised of 96 pairs of emotions. In each item, the target of the vignette reported feeling one emotion more strongly than the other. For scoring purposes, the first item within each pair is represented by a '1'; the second is represented by a '2.' To compute "TOTAL TARGET SCORE" for the EARS, sum the instances in which participants agreed with the targets.

Notice that for each EARS Target, 12 forced-choice pairs follow his or her description of situations affecting his or her emotion. For scoring purposes, the 96 items correspond to items in the scale as follows: Items 1 -12 correspond to the first target's 12 items in order. Items 2-24 correspond to the items for the second target, and so forth.

### *Ears Target Agreement*

Correct answers (i.e., answers that the actual targets responded feeling more strongly) are as follows:

Item #	Target Answer	Item #	Target Answer	Item #	Target Answer	Item #	Target Answer
1*	2	25	1	49	1	73	1
2	1	26	1	50	2	74*	1
3	1	27	1	51	1	75	2
4*	2	28	2	52	1	76	1
5*	1	29	2	53*	1	77	1
6*	2	30	1	54	2	78	1
7	2	31*	2	55	2	79	2
8	1	32*	1	56	1	80	2
9	2	33	2	57	1	81	2
10	2	34	2	58*	2	82	2
11	1	35	1	59	2	83	2
12	1	36*	2	60	2	84*	1
13*	1	37*	1	61	1	85	2
14*	1	38*	2	62	1	86	2
15*	1	39	1	63	2	87	2
16*	2	40	2	64	2	88	1
17*	1	41	1	65	1	89	1
18	2	42	1	66	2	90*	2
19	2	43	2	67	2	91	1
20*	2	44	2	68	1	92	2
21*	1	45	1	69	2	93	2
22*	2	46	2	70	2	94*	1
23*	1	47	2	71*	1	95	1
24	2	48	1	72*	1	96	1

\* an item that may reduce internal reliability.

### *Ears Consensus Agreement*

The "consensus agreement" score corresponds to the degree to which a participant's judgment about the target's emotion reflects the consensus of other participants (independent of whether the target him or herself agreed). Here, the "consensus answer" for each item corresponds to the emotion within each pair that was chosen by a majority of participants

(based on the data with which we standardized the EARS). To compute a consensus score, simply sum the instances in which a given participant chooses items that agree with the consensual answers.

Item #	Target Answer	Item #	Target Answer	Item #	Target Answer	Item #	Target Answer
1*	2	25*	1	49	1	73*	2
2	2	26	1	50	2	74*	1
3*	2	27*	1	51	2	75	1
4	2	28	2	52*	2	76*	2
5*	1	29*	2	53*	1	77*	2
6	1	30*	1	54*	1	78*	1
7	1	31*	2	55*	2	79*	1
8*	2	32	2	56	2	80	2
9	1	33	1	57	1	81	2
10*	1	34	2	58	2	82*	1
11*	1	35	2	59	1	83*	2
12	2	36	2	60	2	84*	1
13	1	37	2	61*	1	85	2
14	1	38*	1	62	1	86	2
15	1	39*	1	63	1	87	2
16	2	40	2	64	2	88	1
17	1	41*	1	65	1	89	2
18*	1	42*	1	66	2	90	1
19*	1	43*	2	67	2	91*	2
20*	2	44*	1	68*	1	92*	1
21	2	45*	1	69	1	93*	2
22	1	46	1	70*	1	94	1
23	1	47	2	71	1	95*	1
24	1	48	1	72*	1	96	2

For a more sensitive consensus score, we recommend the following: If you have EARS data from a relatively large number of participants, compute the frequencies of answers for each item. For example, for the very first item, determine how many participants chose the first pair in the item and how many chose the second. Then, convert these frequencies to percentages. When computing the total consensus score, add the percentage that agreed with the participant for each item. Thus, for example, if 80% of participants chose the first emotion option in the first item, add .8 to that participants Consensus score. This "weighted" measure is more sensitive to degree of consensus and yields a more reliable and valid score.