

# Emotional Intelligence Meets Traditional Standards for an Intelligence

JOHN D. MAYER

*University of New Hampshire, Durham, NH, USA*

DAVID R. CARUSO

*Work Life Strategies, New Canaan, CT, USA*

PETER SALOVEY

*Yale University, New Haven, CT, USA*

An intelligence must meet several standard criteria before it can be considered scientifically legitimate. First, it should be capable of being operationalized as a set of abilities. Second, it should meet certain correlational criteria: the abilities defined by the intelligence should form a related set (i.e., be intercorrelated), and be related to pre-existing intelligences, while also showing some unique variance. Third, the abilities of the intelligence should develop with age and experience. In two studies, adults ( $N = 503$ ) and adolescents ( $N = 229$ ) took a new, 12-subscale ability test of emotional intelligence: the Multifactor Emotional Intelligence Scale (MEIS). The present studies show that emotional intelligence, as measured by the MEIS, meets the above three classical criteria of a standard intelligence.

Emotions are internal events that coordinate many psychological subsystems including physiological responses, cognitions, and conscious awareness. Emotions typically arise in response to a person's changing relationships. When a person's relationship to a memory, to his family, or to all of humanity changes, that person's emotions will change as well. For example, a person who recalls a happy childhood memory may find that the world appears brighter and more joyous (e.g., Bower, 1981). Because emotions track relationships in this sense, they convey meaning about relationships (Schwarz & Clore, 1983). Emotional intelligence refers to an ability to recognize the meanings of emotions and their relationships, and to reason and problem-solve on the basis of them. Emotional intelligence is involved in the capacity to perceive emotions, assimilate emotion-related feelings, understand the information of those emotions, and manage them (Mayer & Salovey, 1997; Salovey & Mayer, 1990).

---

Direct all correspondence to: Dr. John (Jack) D. Mayer, Department of Psychology, Conant Hall, 10 Library Way, University of New Hampshire, Durham, NH 03824-3567, USA. E-mail: jack.mayer@unh.edu

---

**INTELLIGENCE 27(4): 267–298**  
ISSN: 0160-2896

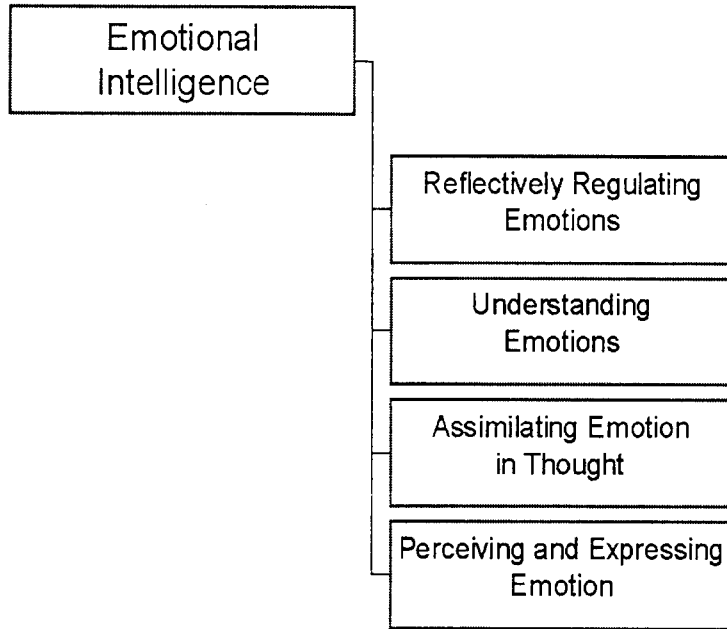
Copyright © 2000 by Elsevier Science Inc.  
All rights of reproduction in any form reserved.

Emotional intelligence can be assessed most directly by asking a person to solve emotional problems, such as identifying the emotion in a story or painting, and then evaluating the person's answer against criteria of accuracy (Mayer, DiPaolo, & Salovey, 1990; Mayer & Geher, 1996). It is worth noting, however, that emotional intelligence, as an ability, is often measured in other ways. Some approaches have asked people their personal, self-reported beliefs about their emotional intelligence. Test items such as, "I'm in touch with my emotions," or "I am a sensitive person," assess such self-understanding (e.g., Mayer & Stevens, 1994; Salovey, Mayer, Goldman, Turvey, & Palfai, 1995). Self-reports of ability and actual ability, however, are only minimally correlated in the realm of intelligence research (e.g.,  $r = 0.20$ ; Paulhus, Lysy, & Yik, 1998) and that appears to hold in the area of emotional intelligence as well (Davies, Stankov, & Roberts, 1998).<sup>1</sup> Self-concept is important, of course, because people often act on their beliefs about their abilities as opposed to their actual abilities (Bandura, 1977). Emotional intelligence as a domain of human performance, however, is best studied with ability measures.

Emotional intelligence has often been conceptualized (particularly in popular literature) as involving much more than ability at perceiving, assimilating, understanding, and managing emotions. These alternative conceptions include not only emotion and intelligence per se, but also motivation, non-ability dispositions and traits, and global personal and social functioning (e.g., Bar-On, 1997; Goleman, 1995). Such broadening seems to undercut the utility of the terms under consideration. We call these *mixed* conceptions because they combine together so many diverse ideas. For example, the Bar-On Emotional Quotient Inventory (EQ<sub>i</sub>) includes 15 self-report scales that measure a person's self-regard, independence, problem solving, reality-testing, and other attributes (Bar-On, 1997). Such qualities as problem solving and reality testing seem more closely related to ego strength or social competence than to emotional intelligence. Mixed models must be analyzed carefully so as to distinguish the concepts that are a part of emotional intelligence from the concepts that are mixed in, or confounded, with it.

General intelligence serves as an umbrella concept that includes dozens of related groups of mental abilities. Most of the smaller subskills studied in this century are related to verbal, spatial, and related logical information processing (see Carroll, 1993, for an authoritative review). Such processing is sometimes referred to as "cold" to denote that its ego- or self-involvement is minimal (Abelson, 1963; Mayer & Mitchell, 1998; Zajonc, 1980). Information processing, however, also deals with "hot," self-related, emotional processing. Emotional intelligence is a hot intelligence. It can be thought of as one member of an emerging group of potential hot intelligences that include social intelligence (Sternberg & Smith, 1985; Thorndike, 1920), practical intelligence (Sternberg & Caruso, 1985; Wagner & Sternberg, 1985), personal intelligence (Gardner, 1993), non-verbal perception skills (Buck, 1984; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979), and emotional creativity (Averill & Nunley, 1992). Each of these forgoing concepts forms coherent domains that partly overlap with emotional intelligence, but that divide human abilities in somewhat different ways.

The ability conception of emotional intelligence was developed in a series of articles in the early 1990s (Mayer et al., 1990; Mayer & Salovey, 1993; Salovey & Mayer, 1990). For example, the first empirical study in the area demonstrated that people's abilities to identify emotion in three types of stimuli: colors, faces, and designs, could be accounted



**Figure 1.** A four-branch model of the skills involved in emotional intelligence (after Mayer & Salovey, 1997).

for by a single ability factor—which we supposed was emotional intelligence (Mayer et al., 1990). Another study examined the understanding of emotion in stories (Mayer & Geher, 1996); this latter study provided further indications that the underlying factor “looked like” an intelligence. Simultaneous with this empirical work, we have honed our definition of emotional intelligence and the abilities involved (e.g., Mayer & Salovey, 1997). The present article represents a culmination of this work, testing our most highly developed conception of emotional intelligence by operationalizing it according to 12 ability tests of emotional intelligence. The present study can help answer important questions about emotional intelligence, among them: whether emotional intelligence is a single ability or many, and how it relates to traditional measures of general intelligence and other criteria.

## STANDARD CRITERIA FOR AN INTELLIGENCE

### Three Criteria for an Intelligence

An intelligence such as emotional intelligence must meet stringent criteria in order to be judged as a true intelligence. For the purposes here, these criteria can be divided into three fairly distinct groups: conceptual, correlational, and developmental. The first, conceptual criteria, includes that intelligence must reflect mental performance rather than simply preferred ways of behaving, or a person’s self-esteem, or non-intellectual attainments

(Carroll, 1993; Mayer & Salovey, 1993; Scarr, 1989); moreover, mental performance should plainly measure the concept in question, i.e., emotion-related abilities. The second, correlational criteria, describe empirical standards: specifically, that an intelligence should describe a set of closely related abilities that are similar to, but distinct from, mental abilities described by already-established intelligences (Carroll, 1993; Neisser et al., 1996).<sup>2</sup> The third, developmental criterion, states that intelligence develops with age and experience, and is based on the groundbreaking work by Binet and Simon at the beginning of century (as reviewed in Fancher, 1985, p. 71; see also, Brown, 1997). These three criteria will be next examined in greater detail.

### Conceptual Criteria for an Intelligence

We have argued elsewhere that emotional intelligence does indeed describe actual abilities rather than preferred courses of behavior. These four broad classes of abilities can be arranged from lower, more molecular, skills to higher, more molar, skills, as is done in Fig. 1 (Mayer & Salovey, 1993, 1997). The lowest level skills involve the perception and appraisal of emotion, e.g., in a facial expression or artwork. The next level up involves assimilating basic emotional experiences into mental life, including weighing emotions against one another and against other sensations and thoughts, and allowing emotions to direct attention. An example includes holding an emotional state in consciousness long enough to compare its correspondences to similar sensations in sound, color, and taste. The third level involves understanding and reasoning about emotions. Each emotion—happiness, anger, fear and the like—follows its own specific rules. Anger rises when justice is denied; fear often changes to relief; sadness separates us from others. Each emotion moves according to its own characteristic rules, like the different pieces on a chessboard. Emotional intelligence involves the ability to see the pieces, know how they move, and reason about emotions accordingly. The fourth, highest level, involves the management and regulation of emotion, such as knowing how to calm down after feeling angry or being able to alleviate the anxiety of another person. Tasks defining these four levels or branches are described in greater detail in the section concerning scale development below.

In considering tasks for an emotional intelligence test, how are we to discriminate right from wrong answers? One common approach drawn from emotions research has been to look for group *consensus* as to the emotional content of stimuli (e.g., Mayer et al., 1990; Wagner, MacDonald, & Manstead, 1986). If the group agrees that a face is happy, say, then that becomes the correct answer. A second possibility is to use *expert* criteria for emotional meanings. An expert could bring a history of philosophy and empirical psychology to bear on judgments about emotional meanings (e.g., Darwin, 1872/1965; Ortony, Clore, & Collins, 1988; Plutchik, 1984; Spinoza, 1675/1984), and this might provide answers similar to, or different from, a consensus criterion. On the other hand, it has been argued that experts simply provide estimates of group consensus, and those estimates are fallible (Legree, 1995). Finally, a *target* criterion is applicable in selected circumstances in which a target individual's emotions or emotional creations are being judged. In such cases, the target can report the emotion he or she was feeling or expressing at the time. The group's consensus, the expert, and the target criteria, represent somewhat different perspectives, and it is therefore unlikely that they would be in complete

agreement. For example, target individuals sometimes report pleasant feelings, perhaps to be socially conforming, when in fact they are perceived by a group as experiencing less pleasant feelings (Mayer & Geher, 1996). Such differences in perspective do not necessarily rule out a general convergence toward a criterion. Such a rough convergence would substantiate the view that emotions convey information, and that emotional intelligence is, in fact, an intelligence.

### **Correlational Criteria for an Intelligence**

#### **The Logic of Correlational Criteria for Intelligence**

Emotional intelligence should define a set of abilities that are moderately intercorrelated with one another. There are many excellent overviews of mental abilities and the criteria for defining their class (e.g., Carroll, 1993; Flanagan, Genshaft, & Harrison, 1997). This logic can be illustrated with an example drawn from the clinical assessment of intelligence. The original Wechsler Adult Intelligence Scales (i.e., WAIS, WAIS-R, WAIS-III; Wechsler, 1958; see Anastasi & Urbina, 1997, for a review of later tests) contained a set of verbal intelligence scales. These consisted of many related mental tests including identifying similarities among concepts, recognizing word meanings (vocabulary), general information, comprehension, and arithmetic. The abilities measured, e.g., vocabulary and information, are moderately intercorrelated—they rise and fall across people at about the  $r = 0.40$  level. The tasks can be summarized by a verbal IQ, where the IQ is based on a person's overall performance on those tasks compared to the performance of other people their age (because ability levels change with age).

The Wechsler tests from mid-century to 1998 typically paired verbal intelligence with performance intelligence. Performance abilities, such as assembling puzzles, identifying missing elements in visual depictions, and ordering picture sequences, also correlate highly with each other. These can be summarized by a performance IQ, similarly based on the person's overall performance on the tasks. The verbal and performance tasks correlate less highly with each other; i.e., the verbal and performance tasks are related to each other, but not quite as closely as skills within each group.<sup>3</sup> They are also related, however, and can be combined to form an overall IQ, which represents the individual's average performance on a broader range of mental tasks.

#### **The Establishment of New Intelligences**

The possibility that there exists one or more additional classes of intelligence, beyond verbal and performance intelligence, has long intrigued researchers. The identification of a new class of intelligence would broaden our contemporary concepts of intelligences. Moreover, adding missing intelligences to an omnibus IQ test can increase the test's fairness by more accurately representing individuals whose abilities were higher on unknowingly omitted tests than on the tests that were present.

The identification of a class of intelligence, such as verbal or performance, however, does not occur all at once. Usually, there proceeds a painstaking process of developing candidate tasks for the intelligence, finding a rationale for correct answers (if not obvious), and then examining their intercorrelations with existing measures of intelligences. For

example, social intelligence was proposed as a third member of the verbal/performance grouping earlier in the century; it was defined as "the ability to understand men and women, boys and girls, to act wisely in human relations" (Thorndike, 1920). Measures of verbal intelligence, however, already incorporate much social thinking; in fact, normal verbal communication is so social that it is difficult to come up with vocabulary ("What is democracy?") or general knowledge questions ("Who was John F. Kennedy?") that do not contain social information. In part, for such reasons, Cronbach (1960) concluded that social intelligence could not be distinguished from verbal intelligence. The search for a third broad intelligence abated for the next several decades, although a number of alternative intelligences have been discussed as possible candidates.

Research on social intelligence has continued, with important work by Sternberg and Smith (1985), Cantor and Kihlstrom (1987) Legree (1995) and others. Much of that work represented important conceptual development of social intelligence; little of that work, however, concerned itself with actual ability measurement in relation to other intelligences (some exceptions are Legree, 1995; Wagner & Sternberg, 1985). In addition, other intelligences have been proposed, e.g., the multiple intelligences of Gardner (1993) which included personal, musical, and other intelligences. Here, too, research on individual differences and their relations to already-existing intelligences was de-emphasized (Sternberg, 1994).

Emotional intelligence represents an alternative grouping of tasks to social intelligence. On one hand, emotional intelligence is broader than social intelligence, including not only reasoning about the emotions in social relationships, but also reasoning about internal emotions that are important for personal (as opposed to social) growth. On the other hand, emotional intelligence is more focused than social intelligence in that it pertains primarily to the emotional (but not necessarily verbal) problems embedded in personal and social problems. For example, reasoning about a sequence of internal feelings, or about the feelings in a relationship, can be readily distinguished from general questions about democracy, or John F. Kennedy, as described above. This increased focus means that emotional intelligence may be more distinct from traditional verbal intelligence than is social intelligence.

### **The Developmental Criterion for an Intelligence**

There remains a third criterion an intelligence must meet: that it develops with age and experience, from childhood to adulthood. That third criterion will be discussed at the outset of Study 2, which is focused on studying developmental issues in emotional intelligence.

### **INTRODUCTION TO THE PRESENT STUDIES**

Widely accepted intelligences share certain features in common: they are abilities, they manifest specific correlational patterns among themselves and in relation to other intelligences, and they develop with age and experience. The two studies described here operationalize emotional intelligence as a set of abilities, study the intercorrelational pattern among those abilities, and examine evidence for their growth between adolescence and early adulthood.

In Study 1, we constructed a set of 12 ability measures drawn from each of the four defined areas of emotional intelligence including perceiving, assimilating, understanding, and managing emotion (Mayer, Salovey, & Caruso, 1997). The test was administered to a large group of adults. We predicted moderate correlations among the 12 tasks, and that a group factor, i.e., one that loads all 12 tasks, can be derived. As in our earlier work, we predicted that a combination of these tasks correlates with traditional forms of intelligence such as verbal intelligence at such a level as to be distinct from such traditional intelligences. Study 1 also examines evidence of whether this emotional intelligence predicts empathy, parental warmth, and cultural pursuits.

Study 2 focused more specifically on whether emotional intelligence meets the developmental criterion for an intelligence. An adolescent sample was given a reduced set of the same group of tasks. The adolescent data are then compared to a subset of the adult data from Study 1 so as to test the hypothesis that adults outperform adolescents on the tasks.

## STUDY 1

### Method

#### Participants

Participants were 503 adults (164 men and 333 women, six unreported) with a mean age of 23 years (range: 17–70), drawn from several sources. One group of individuals (47%,  $N = 235$ ) was comprised of full-time college students who participated to fulfill an introductory psychology course research requirement, or who were paid (US\$15) for their participation. The remainder (53%, 268) were part-time college students, corporate employees, career workshop attendees, and executives in an outplacement setting who volunteered. The full sample was roughly representative of the ethnic composition of the United States census (Self-identified ethnicity/race: African–American, 12% (58); Asian or Asian–American, 6% (31); Hispanic, 6% (32); Native American, 1% (4); White: 68% (340); Other/Not Reported: 7% (38)). The sample was above-average in education: less than 1% (2) had no college; 80% (401) was in college or had been; 12% (59) was college graduates; 7% (34) had advanced degrees; information on the remainder 1% (7) was unreported.

#### The Multifactor Emotional Intelligence Scale (MEIS)

**Overview of Test Organization.** The MEIS consists of 12 tasks, divided into four classes or “branches” of abilities including (a) perceiving, (b) assimilating, (c) understanding, and (d) managing emotion (Mayer & Salovey, 1997; Mayer et al., 1997). Branch 1’s four tests measured emotional perception in Faces, Music, Designs, and Stories. Branch 2’s two tests measured Synesthesia Judgments and Feeling Biases. Branch 3’s four tests examined the understanding of emotion, including in Blends, Progressions, and Transitions between and among emotions, and Relativity in emotional perception. Branch 4’s two tests examined Emotion Management in Self and Others. The content of the subtests and their scoring is described below, as are the three scoring methods employed: consensus, expert, and target.

### Branch 1: Perceiving Emotion

Branch 1 tasks concerned the ability to perceive and identify the emotional content of a variety of stimuli.

**Faces (Eight Stimuli; 48 Items).** The first Branch 1 task, Faces, used as stimuli eight faces from a CD-ROM photographic library and from personal photos, chosen to represent a variety of emotions, and for their authenticity in representing those emotions. Each face was followed by six emotions: *happiness, anger, fear, sadness, disgust, and surprise*. The test-taker was to answer on a five-point scale whether a given emotion (e.g., anger) was “Definitely Not Present” (1) or “Definitely Present” (5). The responses were scored according to two criteria: consensus and expert.

**Consensus Scoring.** The group consensus served as the criterion for this scoring approach. Each participant response was scored according to its agreement with the proportion of the participant group who endorsed the same alternative. For example, if 0.51 of the participant group reported that anger was somewhat present (“4” on the scale), then a participant who chose “4” would receive 0.51 for the item. If the participant believed anger was definitely not present (“1” on the scale), and only 0.06 of the sample agreed, then the individual would receive a 0.06 for the item.

**Expert Scoring.** The first two authors served as experts for the tasks, and went through the test answering questions by bringing to bear, as much as possible, their reading of Western philosophical treatments of emotion, and their reading of contemporary psychological models of emotion. For example, in deciding questions about emotional blends, reference was made to the theory of emotional blends by Plutchik (1984). For each item, the authors identified the best alternative (from 1 to 5) for each response; general agreement with this best response (choosing the selected value, or the integer on either side of it) was scored “1”; otherwise, the individual received a “0.”<sup>4</sup>

**Music (Eight Stimuli; 48 Items).** The second Branch 1 task, Music, was similar to the Faces task. The stimuli consisted of eight brief (5–10 s) original pieces of music composed for this project. Participants heard each piece of music and then rated each one as to its emotional content on a series of mood adjective scales. Each mood adjective was rated from 1 (“Definitely Not Present”) to 5 (“Definitely Present”). The same six mood adjectives were employed as in Faces.

**Target Scoring.** The music test was scored according to the consensus and expert methods used above. In addition, the target scoring method was employed here. Target scoring made use of an additional data set. As the composer–musician worked, he was requested to think about his feelings and the feelings his music conveyed, which he then recorded on a mood scale. Target scoring was scored for agreement with the target’s feelings (in this case, the composer–musician). It was scored as the expert scoring was, with a “1” for a match (give or take 1) and “0” for a non-match. Indeed, the target can be thought of as a second type of expert.

**Designs (Eight Stimuli; 48 Items).** The third Branch 1 task, Designs, was identical to the above except that eight original computer-generated graphic designs served as the



stimuli. The designer was requested to create graphics that portrayed a variety of feelings. As the designer worked, he recorded his feelings on the six-adjective mood scale about what he expressed in the design. Consensus, expert, and target scoring were employed for this task.

**Stories (Six Stimuli; 42 Items).** The fourth Branch 1 task, Stories, was identical to the above tasks except that six stories were employed. The stories were obtained as in Mayer and Geher (1996). Fifteen adult acquaintances of the authors were asked to report on situations or thoughts affecting their moods, including (a) "What led up to the situation?"; (b) "What is the situation, or what you are thinking about?"; and (c) "What happened in this situation which made you feel the way you do?" Immediately thereafter, these 15 supplemental participants recorded their moods on a 30-item mood-adjective checklist, using the five-point rating scale described above (see Faces). The passages were then edited lightly. The six passages were then presented to participants in the main study. An example was as follows.

*This story comes from a middle-aged man. Everything has been piling up at work and I am falling behind. I have been working late many nights and as a result, my wife and daughter are feeling left out. My relationship with them is being stressed. I feel that I am letting them down emotionally. I feel guilty not spending time with them. At the same time, a close family member moved in with us after his divorce and job loss. We have no privacy and I finally told him he has to move out. It was very difficult for me, especially since in the way I was raised, you don't treat a guest this way.*

Each story was followed by a seven-adjective mood scale; the adjectives varied from story to story. They were selected so as to balance adjectives that were applicable to the story and those that were not, as well as to balance positive- and negative-toned adjectives. For the above story, the seven adjectives were, "depressed, frustrated, guilty, energetic, liking, joyous, and happy." The participant's job was to identify the emotion in the story. The responses were scored by consensus, expert, and target criteria.

## **Branch 2: Assimilating Emotions**

Branch 2 tasks concerned the ability to assimilate emotions into perceptual and cognitive processes.

**Synesthesia (Six Stimuli; 60 Items).** The first Branch 2 task, Emotional Synesthesia, measured people's ability to describe emotional sensations and their parallels to other sensory modalities. The analysis of emotions often involves describing their composition in regard to other sense modalities, including movement, touch, pace, and color (Clynes, 1977; de Rivera, 1977). In this task, people imagined an event that could make them feel a particular feeling, which they then described on 10 semantic differential scales. For example, one item asked, "Imagine an event that could make you feel both somewhat surprised and somewhat displeased . . . Now describe your feelings on," each of 10 five-point semantic differential scales, including "warm 1 2 3 4 5 cold," and other scales involving color (yellow or purple) touch (sharp or dull) and so

forth; the scales were invariant across stimuli. This task was scored by consensus and expert criteria.

***Feeling Biases (Four Stimuli; 28 Items).*** The second Branch 2 task, Feeling Biases, asked people to assimilate their present mood into their judgments of *how they felt toward a [fictional] person at the moment*. Thus, one task instructed participants to:

Imagine that Jonathan is one of your relatives. He is a tall, muscular person. Jonathan said something to you that made you feel both guilty and afraid. Feeling both guilty and afraid about Jonathan, how does he seem?

The seven traits following each passage varied so as to be relevant to each passage; in the above example, traits included “sad, trusting, tense, cynical, aggressive, controlling, and hasty.” The traits were rated on a five-point scale (“Definitely Does Not Describe” (1) to “Definitely Does Describe” (5)). The rationale for this task was that people who use their emotions in thinking do so, in part, by analyzing judgmental transformations that occur with mood. This task was scored according to consensus and expert criteria.

### **Branch 3: Understanding Emotions**

Branch 3’s tasks concerned reasoning about and understanding emotions.

***Blends (Eight Stimuli; Eight Items).*** The first Branch 3 task, Blends, concerned the ability to analyze blended or complex emotions. Items were of the following form.

Optimism most closely combines which two emotions?

- (a) pleasure and anticipation
- (b) acceptance and joy
- (c) surprise and joy
- (d) pleasure and joy.

Participants were instructed to select the single best answer. The eight items covered blends of two emotions (four items), blends of three emotions (two items), and blends of four emotions (two items). Scoring was by consensus and expert criteria.

***Progressions (Eight Stimuli; Eight Items).*** The second Branch 3 task, Progressions, concerned people’s understanding of how emotional reactions proceed over time, with a special focus on the intensification of feelings. A sample item read:

If you feel angrier and angrier toward someone so that you are losing control, it would result in (choose one):

- (a) gloating
- (b) resentment
- (c) hate
- (d) rage.

Participants were instructed to identify the single best answer. Items were scored according to consensus and expert criteria.

**Transitions (Four Stimuli; 24 Items).** The third Branch 3 task, Transitions, concerned people's understanding of how emotions (and implicitly, the situations eliciting them) follow upon one another. Items were of the following form:

A person is afraid and later is calm. In between, what are the likely ways the person might feel?

Each item was followed by six alternative feelings. Alternatives for the above item were acceptance, fear, anger, anticipation, surprise, and disappointment. The participant rated each item as "Extremely Unlikely" (1) to have occurred, or as "Extremely Likely" (5). The remaining three items followed the same form.

**Relativity (Four Stimuli; 40 Items).** The fourth Branch 3 task, Relativity, was composed of items depicting conflictual social encounters between two characters. The participant's task was to estimate the feelings of both those characters. One item read:

A dog is chasing sticks outside when he runs out in the street and gets hit by a car. The driver stops when the dog's owner dashes over to check on the dog.

The first items concern the dog-owner's feelings. Participants must decide to what extent the dog owner feels each of five ways, including, "ashamed about not being able to have better trained the dog," or "challenged to protect other dogs from mishaps." Each alternative was rated according to how likely a feeling-reaction was, from "Extremely Unlikely" (1) to "Extremely Likely" (5). Next, the participant made similar judgments as to the second character (the driver, above). In the above example, participants judged whether the driver felt "relief that it was only a dog," or "guilty for not being a more cautious driver," and so on, on the same response scale.

#### **Branch 4: Managing Emotions**

Branch 4 concerns the ability to manage emotions.

**Managing Feelings of Others (Six Stimuli; 24 Items).** The first Branch 4 task, Managing Feelings of Others, examines how participants manage the emotions of others. Participants were asked to evaluate plans of action in response to fictional people, described in brief vignettes, who needed assistance. The task consisted of six vignettes, each followed by four possible courses of action. For example:

One of your colleagues at work looks upset and asks if you will eat lunch with him. At the cafeteria, he motions for you to sit away from the other diners. After a few minutes of slow conversation, he says that he wants to talk to you about what's on his mind. He tells you that he lied on his resume about having a college degree. Without the degree, he wouldn't have gotten the job.

Participants were to rate alternatives such as (for the above vignette):

Ask him how he feels about it so you can understand what's going on. Offer to help him, but don't push yourself on him if he really doesn't want any of your help.

Participants rated responses from “Extremely Ineffective (1)” to “Extremely Effective (5).” Tasks were scored according to consensus and expert criteria.

***Managing Feelings of the Self (Six Stimuli; 24 Items).*** The second Branch 4 task, Managing Feelings of the Self, concerns how a person would regulate his own emotions. This task consisted of six vignettes, each one describing a particular emotional problem. For example:

You have been dating the same person for several months and feel very comfortable. Lately, you are thinking that this relationship may be the one and although marriage hasn't been discussed, you are assuming that it is a real possibility. The last thing you expected was the phone call you received saying that the relationship is over. You have lost the love of your life.

Participants were instructed that not every situation is equally applicable to everyone but to imagine, if in that situation, the effectiveness of given responses. One such response to the above situation was:

The best way to cope with this terrible blow is to do whatever you can to block it out and not let it get to you any more than it has. You would throw yourself into your work or some activity and then try to put it behind you.

Participants rated each response from “Extremely Ineffective (1)” to “Extremely Effective (5).” Tasks were scored according to consensus and expert criteria.

### Criterion Scales

Two classes of criterion scales were employed along with the MEIS. Primary criteria included measures of intelligence and self-reported empathic feeling, both of which have been predicted to correlate with emotional intelligence in the past (Mayer et al., 1990; Salovey & Mayer, 1990). Secondary criteria included measures of several areas in which emotionally intelligent individuals are thought to differ from others. These include higher life satisfaction, a family environment that encourages learning about feelings, and aesthetic perception and participation (Mayer & Salovey, 1990; 1995; Salovey & Mayer, 1990).

#### Primary Criteria

***Intelligence Measure.*** The intelligence criterion was adapted from the Army Alpha test of intelligence (Yerkes, 1921). The Army Alpha was employed because its validity is well-established and its form is ideal for group testing of the sort carried out here. The vocabulary scale was used because that subtest is the strongest component of verbal intelligence (e.g., Wechsler, 1958; Morrison, 1976, pp. 318–325). Thirty of the more difficult vocabulary items from the 50 were selected; more difficult items were favored so as to tailor the test to the participant population, which included mostly college-educated individuals. The Army Alpha vocabulary scale employs four response options for each word to be defined (e.g., “Reply: (1) make, (2) do, (3) answer, (4) come”). Participants were instructed to select the alternative from the list that most nearly meant the same as the target word. The scale had an alpha reliability of  $\alpha = 0.88$  in this data set.

**Empathy Measure.** A 30-item empathy scale (Caruso & Mayer, 1999) was developed with content coverage similar to the Epstein–Mehrabian scale (Mehrabian & Epstein, 1972), but with identifiable factor-based subscales. This newer scale was employed so that overall self-reported empathy and also its subcomponents could be compared to emotional intelligence. The scale’s overall self-reported empathy score has an alpha reliability of  $\alpha = 0.86$ . Because of content overlap with the Epstein–Mehrabian scale, it is likely to perform similarly to it. In contrast to the Epstein–Mehrabian, however, the present scale can be divided into five more specific factor-based scales. The five subscales, their reliabilities, and a sample item from each are: (a) Empathic Suffering,  $\alpha = 0.79$ , “The suffering of others deeply disturbs me”; (b) Positive Sharing,  $\alpha = 0.72$ , “Seeing other people smile makes me smile”; (c) Responsive Crying,  $\alpha = 0.74$ , “I cry easily when seeing a sad movie”; (d) Avoidance [*reversed*],  $\alpha = 0.72$ , “I find it annoying when other people cry in public”; and (e) Feeling for Others,  $\alpha = 0.61$ , “If someone is upset, I get upset too.”

### Secondary Criteria

**Life Satisfaction.** Each person was asked about his satisfaction with his Relationships, Academic Status, and Career and Work Situation, to be reported on a five-point rating scale (from “Not at All Satisfied” to “Extremely Satisfied”). A factor analysis indicated the items were unifactorial although they were only moderately intercorrelated. A single life satisfaction score ( $\alpha = 0.59$ ) was employed, representing the sum of each person’s responses.

**Artistic Skills.** Participants also reported their degree of artistic skill in eight areas (from “1” no or little talent to “3” very talented). A unifactorial Artistic Skill score ( $\alpha = 0.71$ ) indicated overall self-reported artistic skill in those areas, which included sculpture, music, and writing.

**Parental Warmth.** Participants also described their parents’ behaviors on a seven-item scale. A unifactorial Parental Warmth factor ( $\alpha = 0.81$ ) included items reporting that parents were warm, listened, were non-abusive, and (reversed) yelled and were strict.

**Psychotherapy.** Psychotherapy was scored as the number of months a person had psychotherapy (which for some people was zero) multiplied by the number of sessions of psychotherapy per month.

**Life Space Leisure.** Life space scales consist of items that record a person’s environment in terms of discrete, externally verifiable, responses (e.g., “How many pairs of shoes do you own?” “How many times have you attended the theater in the last year?”; Mayer, 1998; Mayer, Carlsmith, & Chabot, 1998). Certain life activities, particularly those involving aesthetic appreciation, have been predicted to involve more emotional intelligence than others (Mayer et al., 1990). Thirty-three items concerning leisure activities were administered to participants covering books read, television watched, and cultural events observed. These items yielded three factor-based scales of leisure activities, based on an unrotated, principal components analysis, using all items

loading on a given factor  $r > \pm 0.45$ . The first, Culture-Seeking, factor scale ( $\alpha = 0.78$ ) loaded the following items: listening to classical music, attending concerts, listening to soul, listening to gospel, listening to country, listening to rap, listening to new-age music, listening to bluegrass, listening to rock, and attending museums. The second, Improvement-Seeking, scale ( $\alpha = 0.60$ ), loaded reading self-help books, how-to books, medical books, business books, and short-stories. The third, Entertainment-Seeking, factor scale ( $\alpha = 0.67$ ), loaded watching action television programs, watching comedy programs, listening to punk music, listening to blues music, and watching televised sports.

### **Procedure**

Participants completed the study in small groups or individually. Each participant received an item and answer booklet that contained all necessary instructions, test items and responses. The test was not timed and the test materials were self-administered, with the exception of the music task, for which a tape of instructions and music was played by the experimenter in group settings.

### **Results**

This section is divided into three parts. First, scoring methods for emotional intelligence are compared. Second, the emotional intelligence tasks are intercorrelated and factor analyzed. Third, emotional intelligence is correlated with various external criteria.

#### **Scoring for Emotional Intelligence**

##### **Consensus, Expert, and Target Criteria for Correct Answers**

Emotional intelligence depends on the idea that certain emotional problems have answers that can be judged correct and incorrect. Convergence among different scoring criteria provides a foundation for such assumptions. The data analysis began by comparing the three different methods for identifying a correct answer: according to (a) the group consensus, (b) expert's identification, and (c) a target's assessment (for three tasks only). We began by examining the degree to which these three methods converged toward a correct response. To the extent that the group consensus (as identified by the modal response) and experts agree as to the best answer, their selections should intercorrelate over the items of a given test. For example, if both the group consensus and experts agree that anger is high in one story ("4" or "5"), but low in another ("1" or "2"), then the correlations should be high.

To test the relation between consensus and expert ratings, we selected four tasks, each drawn randomly from one of the four branches, and calculated the intercorrelation between ratings. The four tasks collectively contain 127 items, representing each branch, and provide a good estimate of the test's overall pattern. For each item, we paired the modal consensus choice with the specific expert selection. In fact, the consensus and expert ratings were fairly highly intercorrelated across tasks: Stories (Branch 1)  $r = 0.70$ ; Feeling Bias (Branch 2)  $r = 0.64$ ; Relativity (Branch 3)  $r = 0.61$ ; and Managing Feelings of Others (Branch 4)  $r = 0.80$ . (All  $r$ s were significant,  $p < 0.0001$  level). This suggests that the two criteria are closely related.

**Table 1.** Means, Standard Deviations, and Reliabilities (Coefficient Alpha) of the Agreement with Consensus, Expert, and Target Criteria

<i>Branch and Task</i>	<i>Scoring Method</i>								
	<i>Consensus</i>			<i>Expert</i>			<i>Target</i>		
	<i>M</i>	<i>S</i>	$\alpha$	<i>M</i>	<i>S</i>	$\alpha$	<i>M</i>	<i>S</i>	$\alpha$
Emotional Identification									
1: Faces	0.40	0.08	0.89	0.64	0.11	0.74	–	–	–
1: Music	0.44	0.11	0.94	0.73	0.10	0.86	0.75	0.12	0.88
1: Designs	0.36	0.08	0.90	0.69	0.11	0.74	0.65	0.14	0.81
1: Stories	0.38	0.07	0.85	0.72	0.11	0.72	0.66	0.10	0.61
Assimilating Emotions									
2: Synesthesia	0.31	0.04	0.86	0.69	0.09	0.66	–	–	–
2: Feeling biases	0.30	0.05	0.70	0.72	0.12	0.60			
Understanding Emotions									
3: Blends	0.49	0.10	0.49	0.60	0.19	0.35	–	–	–
3: Progressions	0.58	0.10	0.51	0.83	0.16	0.50	–	–	–
3: Transitions	0.30	0.04	0.94	0.56	0.11	0.85	–	–	–
3: Relativity	0.30	0.04	0.78	0.56	0.11	0.63			
Managing Emotions									
4: Managing others	0.28	0.04	0.72	0.60	0.12	0.42	–	–	–
4: Managing self	0.27	0.04	0.70	0.55	0.12	0.40	–	–	–

The third, Target criterion, was available only for three Branch 1 tasks. This criterion involves reports by the Targets of their actual feelings as they were creating their artistry (Music and Designs) or telling how they felt (Stories). In general, Consensus correlated a bit more highly with the Target criteria than did the Expert criterion (Music task:  $r = 0.61$  for consensus,  $r = 0.52$  for expert; Designs:  $r = 0.70$  and  $0.60$ ; Stories,  $r = 0.80$  and  $0.69$ ).

Generally speaking, the three criteria appeared to correlate moderately highly, indicating that some answers were “more correct” than others, according to any and all of the scoring methods used: consensus, expert, and target.

### Agreement Scoring for Consensus, Expert, and Target Criteria

A given participant’s performance can be assessed in relation to each of the above three scores: Consensus, Expert, and Target. These agreement scores represent the degree to which a given participant’s responses coincided with those of the criteria. The means, standard deviations, and reliabilities of the participants’ performance, broken down by the three scoring methods, can be seen in Table 1. The means are not directly comparable across consensus, expert, and target because of their substantially different scoring approaches (see above). The figures do indicate, however, the average performance level of the sample, and also that there were no problems of floor or ceiling effects in any of the three scoring methods.

The reliabilities of the agreement scores are also promising, with individual tasks having reliabilities most often between  $\alpha = 0.70$  and  $0.94$  for consensus, and a bit lower for expert agreement. The first two tasks of Branch 3, which were also the shortest, had lower reliabilities,  $\alpha = 0.35 - 0.51$ ; the Branch 4 management tasks were also low, but for expert scoring only. All the reliabilities are satisfactory for this

exploratory study concerning the factorial structure of emotional intelligence and what it predicts.

### Sex Differences in Performance

Women performed somewhat higher than men on the 12 tasks, according to all the scoring procedures. The difference was 0.5 standard deviation for consensus agreement ( $M_{\text{women}} = 0.376$ ;  $SD_{\text{women}} = 0.029$ ;  $M_{\text{men}} = 0.358$ ;  $SD_{\text{men}} = 0.036$ ; Hotelling's  $F(12, 409) = 4.0$ ,  $p < 0.001$ .), and about 0.1 standard deviation for each of expert agreement ( $M_{\text{women}} = 0.664$ ;  $SD_{\text{women}} = 0.048$ ;  $M_{\text{men}} = 0.657$ ;  $SD_{\text{men}} = 0.061$ ; Hotelling's  $F(12, 408) = 4.7$ ,  $p < 0.001$ ), and target agreement ( $M_{\text{women}} = 0.689$ ;  $SD_{\text{women}} = 0.093$ ;  $M_{\text{men}} = 0.676$ ;  $SD_{\text{men}} = 0.079$ ; Hotelling's  $F(3, 482) = 1.34$ , n.s.). This replicates earlier similar findings (cf. Buck, 1984; Mayer & Geher, 1996). We endeavored to understand more about this difference by focusing on the Story task in particular, which was representative of the full test according to subsequent factor analyses (see below). The Story task also showed the greatest sex differences, and contained all three scoring criteria. It is possible that women outperformed men using consensus scoring because the women were using a women's criterion which was different than the men, and the larger number of women in the sample ( $N = 333$  vs. 164) meant that the women's choices were scored with higher values than the men's. This, however, did not account for the women's slightly better performance. Women and men seemed to be employing close to the same criterion. The correlation between women's and men's choices for the emotional content across the 42 story items (six stories, seven items each) was  $r(42) = 0.993$ , indicating a high level of agreement (nor was there any difference in the average emotion-level perceived on an item:  $M_{\text{women}} = 2.67$ ;  $SD_{\text{women}} = 1.19$ ;  $M_{\text{men}} = 2.69$ ;  $SD_{\text{men}} = 1.04$ ;  $t(41) = 0.77$ , n.s.).

We further examined women's and men's performance by employing a two (male participant/female participant) by two (male story character/female story character) by two (male-selected consensus/female-selected consensus) MANOVA on the story data. The MANOVA yielded a main effect representing the women's better consensual accuracy ( $F(1, 495) = 20.08$ ,  $p < 0.001$ ). Women outperformed men under all conditions, even using male-chosen consensus across stories ( $M_{\text{women}} = 0.37$ ;  $SD_{\text{women}} = 0.056$ ;  $M_{\text{men}} = 0.35$ ;  $SD_{\text{men}} = 0.059$ ;  $t(495) = 3.68$ ,  $p < 0.001$ ). A second main effect indicated that the participants, as a whole, were more accurate when using women's consensus criteria over men's ( $F(1, 495) = 928.7$ ,  $p < 0.001$ ). A sex by consensus interaction indicated that women did slightly better using their own consensus criterion ( $F(1, 495) = 67.3$ ,  $p < 0.001$ ). There was also a sex-of-target effect that favored judgments concerning male targets ( $F(1, 495) = 696.3$ ,  $p < 0.001$ ). Collectively, these results indicate that women generally do better than men on these tasks and that the results are not caused by any simple bias in the test materials or how they are scored. Moreover, if one judges by the consensus scoring (which may be fairest, as the expert's criteria were developed by the male authors), the difference between women and men's performance is a moderate 0.5 standard deviation in size.

### Correlations among Consensus, Expert, and Target-scored Tasks

**Consensus and Expert Scoring Considered Individually.** One correlational standard for an intelligence is that it defines a cluster of interrelated abilities (Guttman &



**Table 2.** Intercorrelations of the Consensus Scored Tasks with Reliabilities (Coefficient Alpha) on the Diagonal\*

Branch and Task	1				2			3			4	
	Fa	Mu	De	St	Sy	Fc	Bl	Pr	Tr	Re	Mo	Ms
Emotional Identification												
1: Faces	<b>0.89</b>											
1: Music	0.61	<b>0.94</b>										
1: Designs	0.68	0.60	<b>0.90</b>									
1: Stories	0.54	0.47	0.54	<b>0.85</b>								
Assimilating Emotions												
2: Synesthesia	0.24	0.24	0.26	0.38	<b>0.86</b>							
2: Feeling biases	0.30	0.24	0.35	0.47	0.39	<b>0.71</b>						
Understanding Emotions												
3: Blends	0.07	0.13	0.09	0.24	0.22	0.26	<b>0.49</b>					
3: Progressions	0.10	0.15	0.14	0.25	0.34	0.35	0.41	<b>0.51</b>				
3: Transitions	0.25	0.29	0.29	0.37	0.26	0.34	0.19	0.17	<b>0.94</b>			
3: Relativity	0.30	0.35	0.32	0.41	0.32	0.38	0.30	0.34	0.43	<b>0.78</b>		
Managing Emotions												
4: Managing others	0.20	0.21	0.20	0.28	0.25	0.24	0.16	0.22	0.18	0.37	<b>0.72</b>	
4: Managing self	0.19	0.15	0.14	0.30	0.27	0.22	0.20	0.23	0.17	0.25	0.54	<b>0.70</b>

\*N = 500. Note that correlations above  $r \approx 0.08$  are significant at beyond the  $p < 0.01$  level.

Levy, 1991). The following analyses examine the intercorrelations among the 12 tasks to see if they show a “positive manifold”; i.e., a correlation matrix in which most tasks correlate positively with one another. Correlations among the 12 emotional intelligence tasks were calculated using all three scoring methods. In each case, a positive manifold was evident. Scored by the consensus method, the tasks mostly correlated with one another between  $r = 0.20$  and  $0.50$ , with the full range spanning  $r = 0.07$ – $0.68$ . Scored by the expert method, the tasks mostly correlated with one another  $r = 0.10$ – $0.40$ , with a full range from  $r = 0.00$  to  $0.54$ . In either case, the matrix possesses a positive manifold; almost all the tasks are positively intercorrelated, as expected in regard to a unified intelligence. The consensus-scored tasks (with alpha reliabilities on the diagonal) can be seen in Table 2.

**Comparisons among Consensus, Expert, and Target Scoring.** To further compare scoring methods, we examined participants’ performance on each of the 12 tasks, scored according to a consensus, expert-scoring, or target criterion. For consensus and expert scoring (which were available for all 12 tasks), participants’ performance, scored each way, correlated between  $r = -0.16$  and  $0.95$ , with half the tasks above  $r = 0.52$ . The only negative correlation ( $r = -0.16$ ), which occurred for faces, and other low correlations for the Designs task ( $r = 0.24$ ) may have been a consequence of different color photocopying employed to reproduce the stimuli for the groups and the experts.

The convergence for participant’s consensus and target scores for the Music, Designs, and Stories tasks (where target scoring was available) were  $r = 0.81$ ,  $0.22$ , and  $0.43$ , respectively; the same values for expert and target scoring were  $r = 0.67$ ,  $0.46$ , and  $0.16$ . Subsequent analyses indicated the general superiority of the consensus scoring method in relation to the other alternatives. It yielded higher alpha test reliabilities for every task without exception, clearer factor results, (which were, nonetheless, highly similar to expert

**Table 3.** Three-factor Solutions for the Emotional Intelligence Test Scored According to Consensus and According to Expert Criteria, in Unrotated and Rotated Solutions: Principal Components Factoring<sup>a</sup>

Solution	Branch/Task	Unrotated			Oblique Rotated (Pattern Matrix) <sup>b</sup>		
		I	II	III	I	II	III
Unrotated Solution							
Emotional Identification							
	1: Faces	<b>0.67</b>	<b>-0.48</b>	-0.11	-0.10	<b>0.86</b>	0.04
	1: Music	<b>0.63</b>	<b>-0.34</b>	-0.04	0.02	<b>0.70</b>	0.02
	1: Design	<b>0.69</b>	<b>-0.44</b>	-0.02	0.01	<b>0.82</b>	-0.03
	1: Stories	<b>0.73</b>	-0.09	0.05	0.30	<b>0.52</b>	0.08
Assimilating Emotions							
	2: Synesthesia	<b>0.51</b>	0.19	0.10	<b>0.43</b>	-0.12	0.10
	2: Feeling biases	<b>0.59</b>	0.13	0.21	<b>0.53</b>	-0.20	-0.00
Understanding Emotions							
	3: Blends	<b>0.35</b>	<b>0.32</b>	0.24	<b>0.57</b>	-0.10	-0.01
	3: Progressions	<b>0.43</b>	<b>0.38</b>	<b>0.25</b>	<b>0.64</b>	-0.11	0.02
	3: Transitions	<b>0.48</b>	0.04	0.12	<b>0.35</b>	0.23	0.00
	3: Relativity	<b>0.61</b>	0.18	0.09	<b>0.45</b>	0.20	0.14
Managing Emotions							
	4: Managing others	<b>0.49</b>	<b>0.36</b>	<b>-0.49</b>	-0.05	0.00	<b>0.81</b>
	4: Managing self	<b>0.44</b>	0.36	<b>-0.38</b>	0.03	-0.03	<b>0.68</b>

<sup>a</sup>Loadings above  $\pm 0.25$  are in bold typeface for clarity.

<sup>b</sup>Loadings indicated that all three factors were unipolar (i.e., loadings on a factor above  $\pm 0.25$  all shared the same sign). Rotated factors II and III, however, were negative. To clarify results and facilitate discussion, loadings on rotated Factors II and III were reversed in sign here and in subsequent analyses.

scoring), and higher correlations with criteria. The superiority of consensus scoring has been argued persuasively elsewhere (e.g., Legree, 1995). For that reason, the subsequent analyses focus on the consensus scoring.

## The Structure of Emotional Intelligence

### Factor Structure of the MEIS

Our next question was whether emotional intelligence is best characterized as one or many abilities. Although a highly developed theory of emotional intelligence motivates this article, this represented our own first empirical examination of so many tasks. For that reason, we employed exploratory factor analysis at the outset. We therefore applied principal axis factoring (with communalities on the matrix diagonal) to scores on the 12 MEIS subscales. A joint scree/meaningfulness criterion yielded a three-factor solution (first six eigenvalues: 4.3, 1.6, 1.1, 0.9, 0.8, 0.6). Table 3 (left columns) show the three-factor, unrotated solution for the 12 consensus-scored subscales. We then further analyzed this three-factor solution by rotating it according to an oblique criterion (using an oblimin procedure). The right columns of Table 3 show this result.

In the unrotated solution, the first factor may be interpreted as a general emotional intelligence ( $g_{ei}$ ) because it loads all the tasks without exception. This  $g_{ei}$  apparently represents a group factor of emotional intelligence tasks, suggesting their interrelatedness (below, we explore this question further). The second factor, Managing vs. Perceiving

Emotions, discriminates tasks high in reasoning from those high in simple emotional perception. And the third factor, Managing Emotions, describes the two Branch 4 tasks concerning regulating emotions in oneself and others.

The rotated version of this three-factor solution tells the same story from a different angle. The first factor, Emotional Understanding, loads most of the tasks on Branch 3 (Understanding), along with tasks on Branch 2 (Assimilation). The second factor, Emotional Perception, loads most of the tasks on Branch 1 (Perception). The third, Managing Emotion factor, loads the two Branch 4 (Regulation) tasks, as in the unrotated solution. In this analysis, oblimin factors 2 and 3 had uniformly negative loadings. We reversed the loadings in sign so that a higher score indicated a higher level of ability across tasks. We similarly changed the sign of factor scores and scales based on these two factors. This procedure simplifies the presentation and discussion of results while remaining consistent with the substantive findings.

As a pattern matrix should do, this solution “turns up the contrast” on the loadings, separating the test into three portions: perception, understanding, and managing. This is done, in part, by transferring the common variance shared among the individual tasks to the three factors underlying them. As a consequence, the three factors intercorrelate fairly substantially. Perception correlated  $r = 0.39$  with Understanding and  $r = 0.49$  with Management; the latter two intercorrelated,  $r = 0.33$ .

The above results provide strong empirical support for a three-factor model of the MEIS. Recall that our theoretical model involves a four-branch model. We wondered whether there was also evidence for a four-factor model. To fully investigate this possibility, we modeled the data as a four-factor solution using covariance structural modeling. We used a stringent model in which each task was forced to load only on its hypothesized factor and no other (e.g., Arbuckle, 1997, p. 396). The factors themselves, however, were allowed to intercorrelate as above. The model fit was sufficiently good to be informative, with a Root Mean Square Error of Approximation (RMSEA) of 0.09 with no relaxation of parameters<sup>5</sup> (one rule of thumb is that a RMSEA 0.05 indicates a close fit; Browne & Cudeck, 1993). As appealing as this four-factor model is to us, the drawback is that the model estimates two of the factors, Assimilation and Understanding, to intercorrelate  $r = 0.87$ , which makes them difficult to distinguish from one another. For that reason, we continue to focus on the three-factor model in our analyses, while acknowledging that the four-factor model remains viable.

### **Hierarchical Relations among Factors and the Creation of MEIS Scales**

The first unrotated factor of the MEIS was earlier said to represent a  $g_{ci}$ , or general factor of the test. Such general factors sometimes can arise spuriously due to the nature of principal axis factoring. For that reason, it is often recommended that a hierarchical factor analysis be employed as a secondary check of the existence of a hierarchical factor (e.g., Carroll, 1993; Jensen & Weng, 1994). Obtaining a hierarchical (second-order) factor that loads all the primary factors is generally considered stronger evidence for a general factor because it is based solely on the covariances among the primary factors. A new factor analysis was therefore conducted on the Perception, Understanding, and Managing factor scores. A single hierarchical factor was extracted that loaded Perception, Understanding, and Management at substantial levels ( $r = 0.50, 0.86, \text{ and } 0.75$ , respectively).<sup>6</sup> This

hierarchical factor correlated with the unrotated first factor of the principal axis factoring at  $r = 0.94$ . This final result indicates that general emotional intelligence can be reasonably represented by the first unrotated principal axis factor, and that it loads all the scales studied here.

### The Construction of Factor-based Scales

For our further analyses, we first constructed factor scales for Perception, Understanding, and Managing Emotions factors (and for General Emotional Intelligence,  $g_{ei}$ , based on the first unrotated factor). The scales were constructed by summing  $z$ -scored subscale scores from tasks that loaded on the factors above  $r = \pm 0.35$ . The resulting factor-based scales were then correlated with the original factor scales (based on a weighted sum of all 12 scales) to ensure that they represented the original scales adequately. The three factor-based scales representing perception, understanding, and managing, correlated very highly with their respective factor scales ( $r = 0.98, 0.97,$  and  $0.98$ , respectively), and were highly reliable ( $\alpha = 0.96, 0.92,$  and  $0.81$ ). The three factor-based scales were moderately intercorrelated (Perception with Understanding,  $r = 0.44$ ; Perception with Managing,  $r = 0.29$ ; Understanding with Managing,  $r = 0.43$ ). The overall General Emotional Intelligence factor-based scale also correlated with its original factor scale  $r = 0.97$ , and had a reliability of  $\alpha = 0.96$ . These are the scales reported in the rest of the article.

Researchers wishing to retain the four-branch theoretical model (modestly supported by covariance structural modeling above) may wish to employ four, rather than three, factor-based scales. The above three scales may be transformed into four by (a) retaining the Perception and Managing scales as calculated above, and (b) splitting the Understanding scale, above, into two scales. The first of these two scales, the revised Understanding scale, is calculated as the sum of the  $z$ -scores of the Blends, Progressions, Transitions, and Relativity tasks. The second of these two scales, the new Assimilation scale, is calculated as the sum of the  $z$ -scores of the Synesthesia and Feeling Biases tasks.<sup>7</sup> The  $\alpha$  reliabilities of the (unchanged) Perception, (new) Assimilation, (revised) Understanding, and (unchanged) Management factor-based scales are, respectively,  $\alpha = 0.96, 0.86, 0.89,$  and  $0.81$ . The Understanding factor-based scale still correlates with the original factor scale  $r = 0.89$ . The new, Assimilation factor-based scale, correlates with Understanding,  $r = 0.65$ .

### Relation of the Emotional Intelligence Factors to Criterion Measures

The final correlational criterion for an intelligence is that it correlates moderately with intelligences in other domains. The correlation should be high enough to indicate that the new skill is an intelligence, but low enough to illustrate that it says something new about human abilities. Aside from emotional intelligence's correlation with verbal intelligence, emotional intelligence will be important to the degree that it predicts other criteria as well.

Table 4 shows the correlation of the emotional intelligence factors with various criteria. The central correlations to examine are those with the  $g_{ei}$  factor (first column). General emotional intelligence is then divided into subfactors of perception, understanding, and management; correlations with those subfactors are shown in the next three columns.

**Table 4.** Correlations Between Individual Tasks and Selected Criterion Variables

Criterion Variables	Overall Score	Subfactor Scores		
	$g_{ei}$	Perception	Understanding	Management
<b>Primary Criteria</b>				
Ability				
Verbal IQ	0.36**	0.16**	0.40**	0.20**
Empathy				
Overall	0.33**	0.20**	0.25**	0.34**
Suffering	0.35*	0.18**	0.28**	0.37**
Positive Sharing	0.26**	0.12**	0.16*	0.36**
Crying	0.14**	0.10*	0.10*	0.13**
Avoidance	-0.26**	-0.20**	-0.23**	-0.15**
Feeling for others	0.16**	0.08	0.09*	0.24**
<b>Secondary Criteria</b>				
Life satisfaction	0.11*	0.01	0.11*	0.13**
Artistic skills	0.05	0.03	0.07	0.00
Parental warmth	0.23**	0.20**	0.18**	0.15**
Psychotherapy	0.03	0.04	0.14*	0.02
Leisure (life space)				
Culture-seeking	0.00	-0.07	0.01	0.03
Self-improvement	-0.16**	-0.07	-0.22**	-0.05
Entertainment	-0.02	0.09*	-0.04	0.05

\* $p < 0.05$  level.\*\* $p < 0.01$  level, two-tailed tests.

The correlation between the General Emotional Intelligence factor-based scale and verbal intelligence is  $r = 0.36$ ,  $p < 0.001$ . This is the moderate level at which one would hope that a new domain of intelligence would be correlated with existing domains. In addition, emotional intelligence has a number of interesting correlations with other variables. The  $g_{ei}$  factor-based scale correlates  $r = 0.33$ ,  $p < 0.001$  with overall empathy, also as expected (Salovey & Mayer, 1990; Mayer et al., 1990; Mayer & Geher, 1996), and possesses a number of significant correlations with subtypes of empathy as well, correlating positively and at similar levels with Suffering, Positive Sharing, and negatively with Avoidance. Emotional intelligence had a positive correlation with parental warmth,  $r = 0.23$ ,  $p < 0.01$ , and a negative correlation with pragmatic attempts at self-improvement  $r = -0.16$ ,  $p < 0.01$ , including reading self-help books, books on business methods, and the like. The subfactor scales further qualify the relations, suggesting that Understanding is most closely related to verbal intelligence among the three subfactors, and that Management most accounts for empathy; all three subfactors are related to Parental Warmth.

An extremely stringent test would partial verbal IQ and self-reported empathy out of the correlation between emotional intelligence and the six secondary criteria. Doing this may remove variance that legitimately belongs to emotional intelligence, but it also ensures that emotional intelligence contributes unique variance in predicting criteria. Partialing out the influence of intelligence and empathy yielded a  $g_{ei}$  that maintained its significant negative correlation with attempted self-improvement ( $r = -0.10$ ,  $p < 0.05$ ), and added a negative relation to culture-seeking ( $r = -0.09$ ,  $p < 0.05$ ), although it no longer correlated with life satisfaction or parenting.

Although the statistical relations between emotional intelligence and the life space criteria may seem low, two things are worth noting about them. First, our central focus has been on understanding the structure of emotional intelligence. The few secondary criterion scales included here were exploratory and brief. Even these crude measures, however, demonstrate that emotional intelligence predicts criteria independent of the influence of both verbal intelligence and empathy. Second, it is worth recalling that personality relations tend to be small but consistent over the years. This small but consistent influence can substantially change a person's position in life, just like a slow but steady current can move a boat a considerable distance across a lake over time.

### Summary and Discussion of Study 1

The results from Study 1 indicate that emotional intelligence shows a pattern that is consistent with a new domain of intelligence. Emotional intelligence can be operationalized as sets of abilities, and better answers can be distinguished from worse answers, as indicated by the convergence of three scoring methods. The 12 tasks also intercorrelate with one another, independent of which scoring method is employed. The scale yields four scores: A first, superordinate factor of general emotional intelligence that provides one excellent and economical method for representing the concept. The General Emotional Intelligence factor can be divided in turn into three subscales: Perception, Understanding, and Managing (thus reducing our four-branch model to a three-branch model). Finally, emotional intelligence correlates moderately with a measure of verbal intelligence, indicating that it is related to other intelligences without being the same as them. Emotional intelligence shows promise as a predictor of other qualities such as empathy, (retrospective) parenting style, and life activities.

### STUDY 2

Thus far, emotional intelligence has met two of three important criteria of a traditional intelligence. First, it has been operationalized as a set of abilities. Second, it has shown a pattern of correlations consistent with the existence of such an intelligence. The third criterion is that intellectual capacities grow with age and experience from childhood to early adulthood (Brown, 1997; Fancher, 1985). The importance of age to intelligence was first recognized by Binet. As Fancher (1985, p. 71) describes it:

Gradually . . . a key insight developed—one which seemed perfectly obvious once recognized, but which nevertheless had previously eluded Binet and other investigators of intelligence. *Age* was a crucial factor to be considered: both subnormal and normal children might learn to pass the same tests, but normal children did so at a younger age.

Fancher attributes Binet's success in measuring intelligence, in comparison to the failures of his contemporaries, to the realization that mental abilities grow with age and experience.

For emotional intelligence to behave as does a standard intelligence, it should be shown to increase with age. To test whether this actually occurs, several portions of the scale employed in Study 1 were administered to a young adolescent sample (ages 12–16) in Study 2. The performance of the adolescents was then compared to the

performance of an adult subsample drawn from Study 1. The use of two samples close in age ensures that the same test items can be used and understood by both groups. It also provides a challenging test of the developmental hypothesis because proximity in age should yield only small differences in performance between the two groups. We hypothesized that the adult sample would significantly outperform the adolescents on the scale.

## Method

### Adolescent Sample

Participants were 229 adolescents (125 young men, 101 young women; 3 unidentified) with a mean age of 13.4 (range 12–16) who were recruited from two independent secondary schools and a religious youth group. These were split among 35% (81) 7th graders, 36% (83) 8th graders, 9% (20) 9th graders, 12% (27) 10th graders, 6% (13) 11th graders, and 1% (2) 12th graders; (percentages add to 99% due to rounding error). The sample deviated somewhat from the ethnic composition of the United States census in under-representing minority groups (Self-identified ethnicity/race: African–American, 5% (12); Asian or Asian–American, 3% (6); Hispanic, 3% (7); Native American, 0% (0); White: 79% (177); Other/Not Reported: 9% (27).

### Adult Sample

The adult sample from Study 1 was again used in Study 2. Here, however, the adult sample was divided on the basis of subject number into two equal-sized samples: the “Independent Adult Sample” and the “Consensus Sample.” The Independent Adult Sample served as the comparison group for the adolescent group. The Consensus Sample was used to calculate a consensus score to which the first, “Independent Sample” had not contributed.

### Materials

For reasons of time and age-appropriateness, only a subset of the scales administered to adults was administered to the developmental sample. These included Faces, Music, Designs, and (age-appropriate portions of) Stories from Branch 1,<sup>8</sup> Synesthesia from Branch 2, and Blends and Relativity from Branch 3. In addition, the Army Alpha Vocabulary scale and the Empathy scale were administered as criteria.<sup>9</sup>

### Procedure

Parental consent was first obtained for each participant in the adolescent group, and then informed consent obtained from each subject. All data were collected anonymously; no names were requested. Furthermore, subjects were explicitly instructed not to answer any questions that made them uncomfortable.

Participants in the developmental sample were tested in a similar manner to the adults. They completed the materials in small groups. Each participant received an item and answer booklet that contained all necessary instructions, test items and responses. For the music task, a researcher (or a classroom teacher) played the cassette tape that included all

**Table 5.** A Test of the Developmental Hypothesis: Means (and Standard Deviations) of Adult vs. Adolescent Performance on Selected Consensus-scored Scales of Emotional Intelligence

Branch/Task	Adult		Adolescent		F
	M	SD	M	SD	
Emotional Identification					
1: Faces	0.400	0.078	0.384	0.075	5.2*
1: Music	0.445	0.092	0.438	0.074	0.5
1: Design	0.359	0.086	0.353	0.077	0.7
1: Stories	0.328	0.069	0.323	0.061	1.5
Assimilating Emotions					
2: Synesthesia	0.306	0.045	0.295	0.047	7.8**
Understanding Emotions					
3: Blends	0.491	0.087	0.424	0.105	52.2**
3: Relativity	0.307	0.053	0.304	0.058	0.2
Combined tests	0.378	0.046	0.359	0.048	25.6**

\* $p < 0.05$ .\*\* $p < 0.01$ .

necessary instructions as well as the musical selections. Students required 45–75 min to complete the test booklet.

## Results

### Scoring

Three scoring procedures were employed as in Study 1: agreement with consensus, expert ratings, and target reports. Some modifications in the consensus scoring were necessary for this study. Using the adult consensus as in Study 1 would plainly favor adults because each adult's score contributed to the consensus. To control for this, the adult sample was divided in half (on the basis of odd/even subject number). Next, new adult consensus scores were calculated for the even half of the sample only (the consensus sample). This left the odd half of the adult sample with responses that were independent of the adult consensus (the independent sample). It was this "independent" adult sample whose consensus scores were compared to the adolescent's consensus scores. Expert-scoring and target-scoring were the same as in Study 1.

### Adult–Adolescent Comparisons

The central purpose of Study 2 was to examine whether adults functioned at a higher level of emotional intelligence than adolescents. This hypothesis was tested via a two (Age-Group) by seven (Task) ANOVA, where the seven tasks were within-subjects variables. As the developmental hypothesis predicted, scores were higher for adults than for adolescents for consensus agreement (Grand Mean = 0.38 vs. 0.36;  $F(1,713) = 23.8$ ,  $p < 0.001$ ), for expert agreement (Grand Mean = 0.66 vs. 0.64;  $F(1,709) = 22.3$ ,  $p < 0.001$ ), and for target agreement (Grand Mean = 0.69 vs. 0.67; Hotelling's  $F(1,718) = 8.0$ ,  $p < 0.01$ ). Significant Task and Age  $\times$  Task effects were also present for all three scoring methods. Focusing on consensus scoring, there was a significant Age-Group  $\times$  Task interaction for consensus ( $F(6,708) = 12.5$ ,  $p < 0.01$ ). Table 5 shows a more detailed



comparison between the adult and adolescent groups for consensus scoring on the individual tasks they both received, and which tasks showed significant differences in the predicted direction on their own. We did not examine age-performance correlations beyond demonstrating these average differences. First, the two samples were not strictly comparable, as the adolescents were of slightly higher social class (and therefore would be expected to perform more highly than average). More seriously, the restriction of age-range in both samples would render the correlation impossible to assess. We did, however, determine that the adolescents' scores showed the same pattern of correlations with verbal intelligence ( $r(220) = 0.45, p < 0.001$ ) and empathy ( $r(227) = 0.37, p < 0.001$ ) as did the adults.

### Discussion of Study 2

Study 2 tested whether emotional intelligence met the third of three criteria for a standard, conventional intelligence: whether ability levels increase with age. As predicted, adults performed at higher ability levels than do adolescents. In addition, emotional intelligence in adolescence shows the same relations to verbal intelligence and empathy as with adults.

## GENERAL DISCUSSION

Three major criteria for a standard intelligence are that it consists of mental abilities, that those abilities meet certain correlational criteria, and that the abilities develop with age. In the tests conducted here, emotional intelligence met all three criteria. First, emotional intelligence could be operationalized as a set of ability tests. Second, performance on those ability tests was intercorrelated and partly distinct from verbal intelligence, against which they were compared. Third, emotional intelligence was shown to grow from early adolescence to young adulthood. Collectively, these findings bring us a major step forward toward demonstrating a plausible case for the existence of this intelligence. The data also tell us about the structure of emotional intelligence, and what it might predict.

### The Nature of Emotional Intelligence and Its Measurement

Our factor analyses of the 12 MEIS tasks suggest that one can best conceptualize emotional intelligence as involving three primary factors, and a higher order, General Emotional Intelligence factor that combines the three. The three primary factors involve Perception, Understanding, and Managing of emotion. Perception skills include those drawn from the first branch of the model, including recognizing emotions in Faces, Music, Designs, and Stories. Understanding skills include those drawn from the Assimilation and Understanding branches of the model, including Synesthesia, Feeling Biases, Blends, Progressions, Transitions, and Relativity. Finally, Managing emotions represent skills drawn from the fourth branch of the model, including Managing Others and Managing the Self. The three primary factors, in other words, could be said to capture the four branches of our most recent model of emotional intelligence (Mayer & Salovey, 1997): The Perception branch was captured by the Perception factor scale, the Assimilation and Understanding branches were combined into a single Understanding factor scale, and the Managing Branch was captured by the Managing factor scale. This three-branch measurement approach is also broader than our original 1990 model (which omitted the Understanding branch). The three primary facets of Perception,

Understanding, and Management clearly emerged from the data, and although one still might possibly develop a four-branch measure, it would apparently require developing substantially different Branch 2 tasks than the ones employed here, so as to better distinguish them from the Understanding branch.

The three primary abilities appear to be differentially related to traditional intelligence, with Perception least related ( $r = 0.16$ ), Management moderately related, and Understanding most related ( $r = 0.40$ ). This is consistent with other findings that scales of nonverbal perception, such as the PONS (Profile of Nonverbal States; Rosenthal et al., 1979) which appear to be loaded on emotional perception, are relatively unrelated to intelligence, whereas problem solving of the sort covered on the Understanding branch plainly resembles traditional test items for intelligence more closely. Thus, skills representing emotional intelligence can be ordered along a continuum from those least to those most related to general intelligence. At the same time, the three tasks apparently share a common core of emotionally intelligent processing, as indicated by a more general, overall emotional intelligence factor.

As just noted, a single factor of emotional intelligence incorporates all the tasks studied here. This factor arises as a hierarchical factor obtained from factor-analyzing the three primary factor-based scales of Perception, Understanding, and Management (which are obtained by an oblimin rotation of the 12 tasks). Those three primary scales are fairly intercorrelated and factor-analyzing them yields a single overall factor that summarizes performance across them all. This hierarchical factor is essentially identical to the first, unrotated principal factor of the 12 scales (their  $r = 0.94$ ). The global factor indicates that it makes sense to talk about a single, unified emotional intelligence and a single emotional intelligence score. Such a score provides a reasonable first approximation of a person's ability level in the domain of emotional intelligence. As with any generalization, however, this overall score neglects variations in three subsidiary aspects of emotional intelligence, which can provide further clarification of any overall score.

### **Alternative Representations**

It should be noted that the 12 tasks employed here do not exhaust the universe of emotionally intelligent abilities. As other tasks are developed, it is possible that more factors will be identified. One sort of task, in particular, that does not lend itself to group testing but that may form a separate factor, is ability at expressing emotion (Branch 1 skills). It may also be that a factor better encompassing assimilating emotions (Branch 2 of our model, which merged into Branch 3) might emerge as a more independent factor were it operationalized in tasks different than the ones used here.

### **Sex Differences**

The identification of a new intelligence should increase the fairness of mental ability tests on average. That is because measures of the new intelligence help assess more of the total domain of intelligence, thereby giving any previously neglected capacities their fair consideration. Women and men appear to perform about the same on most intelligence-related mental tests, with most mean differences between 0.15 and 0.30 of an estimated population standard deviation (Hedges & Nowell, 1995). There are, however, some regular differences in the profiles of the two groups. Women are somewhat better on tests of reading comprehension, perceptual speed, associative

memory, and composition. Men are somewhat better in mathematics, social studies, and in scientific knowledge.

To the list of tasks at which women are somewhat better may be added emotional intelligence. Women performed about 0.5 standard deviation higher than men in the present study. The fact that women are slightly superior to men in perceiving emotion has been known for some time, through tests of nonverbal perception (that include emotion) such as the PONS (Rosenthal et al., 1979), as well as through earlier-developed tests of emotional intelligence (Mayer & Geher, 1996). One possible explanation for this is that women must read emotions more carefully because they possess less power in society than do men (LaFrance & Hecht, in press). It is women in more powerful positions rather than less, however, who exhibit the greater emotional accuracy (Hall & Halberstadt, 1994). Such findings suggest that emotional intelligence operates like other areas of intelligence, potentially raising the occupational status of an individual. Issues of power and status aside, women may be socialized to pay more attention to emotions, or they may be better biologically prepared to perform at such tasks; our research does not address the relative contributions of the two (cf., LaFrance & Banaji, 1992).

### **Emotional Intelligence, Intelligence, and Empathy**

The findings here also concern what emotional intelligence predicts. From the outset, emotional intelligence has been hypothesized to correlate with both intelligence and self-reported empathy (Mayer et al., 1990; Salovey & Mayer, 1990). Overall emotional intelligence,  $g_{ei}$ , correlated with verbal intelligence at a low-to-moderate level, as predicted. This replicates some of our earlier work as well (Mayer & Geher, 1996).

Overall emotional intelligence,  $g_{ei}$ , also correlates with self-reported empathy. This, too, replicates earlier studies (Mayer & Geher, 1996; Mayer et al., 1990). Emotional intelligence appears to correlate reliably with self-report empathy scales that share content overlap with the Epstein–Mehrabian scale (Mehrabian & Epstein, 1972). Such scales, including the one used here and the Davis (1983) empathy subscales of “empathic concern” and “emotion-related fantasy,” involve a view of oneself as emotionally responsive and concerned about the feelings of others. The new scale employed here divides that same content domain into a variety of subfactors including Empathic Suffering, Positive Sharing, Responsive Crying, (reversed) Avoidance, and Feeling For Others. Emotional Intelligence correlated with each of these criteria in the expected direction.

### **Emotional Intelligence and Other Intelligences**

The above demonstrations indicate that Emotional Intelligence, as measured by the MEIS, meets the most essential criteria for a standard intelligence. Our results illustrated that emotional intelligence does relate to general intelligence (via its proxy, verbal intelligence). The results, however, provide only the roughest idea of the relation between emotional intelligence and other intelligences. For example, traditional, academic intelligences can be divided into fluid and crystallized intelligences, or verbal and performance intelligences, or divided in many other ways (e.g., Carroll, 1993; Flanagan et al., 1997; Horn & Noll, 1994). Intelligence researchers will want to examine emotional intelligence and those various breakdowns in greater detail. The relation between emotional intelligence and other potentially similar intelligences such as social intelligence and personal intelligence, and the like, are similarly yet-to-be explored. As stated at the outset, any final

choice between emotional intelligence and such alternatives as social intelligence, will depend upon the relative clarity of their operationalizations, their relative relations to general intelligence, and what criteria they predict. It is too early to make this comparison as of yet. Only one of the competing intelligences (social intelligence) has been operationalized well enough (e.g., a minimum of three or four ability tasks) to compare to emotional intelligence. Other alternative intelligences, however, such as personal intelligence, could move in that direction in the future. As alternative intelligences become operationalized, it will be of interest to see how they compare. Finally, emotional creativity (Averill & Nunley, 1992) emphasizes generative, divergent thinking rather than the reasoning and problem solving of emotional intelligence. Emotional intelligence can be thought of as bearing the same relation to emotional creativity as general intelligence bears to general creativity. The intelligence-creativity relations are likely to be complex, but the retention of both concepts likely will be useful.

### **Emotional Intelligence and Other Criteria**

A crucial job of the field is to relate internal characteristics of personality—including abilities such as emotional intelligence—to other psychological tests, and ultimately, to criteria in the life space (Mayer, 1998; Mayer et al., 1998). The test developed here has not yet been correlated with other personality scales such as the Big Five (McCrae & Costa, 1997), and that would be a desirable future direction. Instead, we moved modestly into comparing these internal abilities with actual life criteria: parental warmth, life satisfaction, psychotherapy, artistic ability, and leisure activities related to culture. The findings indicate that emotional intelligence is related to (self-reported) parental warmth and support, and, to a lesser extent, to life satisfaction. Emotional intelligence was also related to leisure pursuits including, negatively, to reading a large number of self-help books. Although these findings are preliminary, and better criteria are desirable, they are suggestive of the fact that emotional intelligence will be of use in predicting particular life criteria.

### **Future Research**

We are presently at the beginning of the learning curve about emotional intelligence. Many questions remain unanswered. Some still concern the factorial structure of emotional intelligence: With the development of more tasks, will there be an additional factor of emotional expressiveness, or of assimilating emotion? Are there nonverbal tests that should be developed? Other questions concern the relation between emotional intelligence and other intelligences: How highly does emotional intelligence correlate with social intelligence, or with performance intelligence, or with spatial intelligence? More generally, how will it relate to the multitude of traditional cognitive abilities reviewed by Carroll (1993) and Horn & Noll (1994)?

Many of the questions of greatest interest to people, however, are those raised (as claims, rather than questions) by members of the press (e.g., Gibbs, 1995; Goleman, 1995). Specifically, these claims included that emotional intelligence accounts in some large part for an individual's success, perhaps more so than conventional analytic intelligence (IQ). Despite the fact that certain among these claims appeared in reputable magazines and newspapers, there has been little or no direct evidence to support them (Mayer & Salovey, 1997; Mayer, Salovey, & Caruso, in press). Until the present article, in fact, there has been no widespread, systematic attempt to understand the measurement of

emotional intelligence as an ability, although self-report mixed-model scales are proliferating (e.g., Bar-On, 1997). The present results indicate that emotional intelligence does play some role in everyday life. It is our hope that the field can move forward employing measures such as the MEIS. Measures such as the MEIS can provide serious answers to the questions above, as well as those that will arise in the future.

There are some matters that are clearly important about emotional intelligence already. Although emotions often have been regarded with respect in the West, there also exists a widespread negative view of people who think emotionally (Payne, 1986). Emotional thinkers have been referred to over the centuries variously as "overly emotional," romantics (or hopeless romantics), people who think with their hearts (instead of their heads), people swayed by emotions, or "biased" by emotions. Such labeling does accurately capture a kind of person who is overwrought with unthinking emotionality. What the existence of emotional intelligence tells us, however, is that there exists another type as well: the emotional, romantic, thinker-with-a-heart, who is engaged in sophisticated information processing, and who, in such a manner, contributes importantly to our lives and culture.

### Conclusion

Measures of intelligence focused on verbal and performance intelligence have been developed over the century. Although verbal, performance, and other similar intelligences have taken us far (cf., Ree & Earles, 1992), there has also been a dissatisfaction with such limited conceptions of mental abilities. Over the century, many have sought out broader sets of mental capacities (e.g., Gardner, 1993; Guilford, 1967; Sternberg, 1988; Thorndike, 1920), or depicted a system of mental abilities (Detterman, 1986). Emotional intelligence represents, to us, an important candidate to enlarge the group on which general intelligence is based. Perhaps a general intelligence that includes emotional intelligence will be a more powerful predictor of important life outcomes than one that does not.

**Acknowledgements:** The research reported here was supported, in part, by the University of New Hampshire. We are grateful for the assistance we received from a number of individuals. Peter Legree read and commented on this manuscript several times and contributed substantially to its quality. Scott Formica was also instrumental in identifying several ambiguities in the near-final manuscript that we were able to correct. In respect to developing the MEIS test itself, the music selections used in the test were composed and recorded by Richard Viard. Graphic designs used in the test were created by David Silverman and Jeremiah Washburn. Finally, Amy Van Buren and Heather Chabot conducted a portion of the data collection. To these and all others who contributed to this work, we extend our heartfelt thanks.

### NOTES

1. The Davies et al. (1998) article critiques early scales in the emotional intelligence literature. The present article was essentially completed before the Davies et al. work was published, and so we do not comment specifically on those authors' criticisms of emotional intelligence scales developed before this one. Nonetheless, it is our hope that the test results here will put to rest certain of the Davies et al. concerns, such as those related to the reliability of emotional intelligence tests.

2. One very different set of criteria, suggested by Howard Gardner, includes requirements that intelligence be identified with a specific brain region or structure and be a culturally valued measure characteristic. Intelligences that are valid according to Gardner's criteria alone are definitely worth studying and may provide information for a next generation of intelligence tests. Still, intelligences that fit his criteria but that are indistinguishable from general intelligence at a behavioral level plainly cannot assist in predicting criteria such as academic success. For that reason, correlational approaches remain of the greatest pragmatic concerns for now.

3. If two intelligences are entirely unrelated, however, we may want to raise the question as to whether one of them is a real intelligence, because mental abilities are generally related to one another. In fact, the "First Law of Intelligence" of Guttman and Levy (1991) states that all mental ability measures are positively correlated.

4. If the expert-selected value was "3" on the five-point scale, responses from 2 to 4 were assigned a value of 1 (correct). If the expert-selected value was "1," then 1-2 would be correct; if the expert value was "4" then 4-5 would be correct, etc.

5. Simply allowing the Stories task to load on the Understanding (as well as Perception) moves the RMSEA index to 0.077.

6. As with Oblique factors II and III, we reversed the sign of the hierarchical factor so that a higher score reflected better ability.

7. These two tasks had estimated loadings of 0.51 and 0.59, respectively, on the assimilation factor, of the oblique four-factor model.

8. Adults and adolescents were compared on two of the eight stories as six stories were deemed potentially unsuitable to adolescents, using extremely cautious criteria, due to their content.

9. Several additional scales were administered that had been rewritten for a younger age group. Reports of the downward extension of the emotional intelligence test can be found elsewhere (see Caruso, Van Buren, Mayer, & Salovey 1998). Only those tests that were identical across groups are examined here because only those are relevant to the developmental hypothesis examined here.

## REFERENCES

- Abelson, R. P. (1963). Computer simulation of "hot cognitions." In S. Tomkins & S. Messick (Eds.), *Computer simulation of personality* (pp. 277-298). New York: Wiley.
- Anastasi, A., & Urbina, S. (1997). *Psychological Testing*. Upper Saddle River, NJ: Prentice-Hall.
- Arbuckle, J. L. (1997). *Amos Users' Guide (Version 3.6)*. Chicago, IL: SmallWaters Corporation.
- Averill, J. R., & Nunley, E. P. (1992). *Voyages of the heart: Living an emotionally creative life*. New York: Free Press.
- Bandura, A. (1977). Self-efficacy: Toward a unified theory of behavioral change. *Psychological Review*, 1, 191-215.
- Bar-On, R. (1997). *Bar-On emotional quotient inventory: A measure of emotional intelligence*. Toronto, ON: Multi-Health Systems Inc.
- Bower, G. H. (1981). Mood and memory. *American Psychologist*, 36, 129-148.
- Brown, B. (1997). Raw scores of cognitive ability are real psychological variables: IQ is a hyperspace variable. V. C. Shipman (Chair). *IQ or cognitive ability?* Symposium presented to the 9th Annual Convention of the American Psychological Society, Washington, DC.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Buck, R. (1984). *The communication of emotion*. New York: The Guilford Press.
- Cantor, N., & Kihlstrom, J. F. (1987). *Personality and social intelligence*. Englewood Cliffs, NJ: Prentice-Hall.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Caruso, D. R., & Mayer, J. D. (1999). *A quick scale of empathy*. Manuscript in preparation.
- Caruso, D. R., Van Buren, A., Mayer, J. D. & Salovey (1998). *An extension of the emotional intelligence test to an adolescent sample*. Manuscript in preparation.
- Clynes, M. (1977). *Sentics: The touch of emotions*. Garden City, NY: Anchor Press/Doubleday.
- Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd edn.). New York: Harper and Row.

- Darwin, C. (1965). *The expression of emotions in man and animals*. Chicago: University of Chicago Press. (Original work published 1872).
- Davies, M., Stankov, L., & Roberts, R. D. (1998). Emotional intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology*, *75*, 989–1015.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126.
- de Rivera, J. (1977). *A structural theory of the emotions*. New York: International Universities Press.
- Detterman, D. K. (1986). Human intelligence is a complex system of separate processes. In R. J. Sternberg & D. K. Detterman (Eds.), *What is intelligence? Contemporary viewpoints on its nature and definition* (pp. 57–61). Norwood, NJ: Ablex.
- Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy*. New York: W.W. Norton.
- Flanagan, D. P., Genshaft, J. L., & Harrison, P. L. (1997). *Contemporary intellectual assessment: Theories, tests, and issues*. New York: Guilford Press.
- Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. (10th Anniversary edn.). New York: Basic Books.
- Gibbs, N. (October 1995). The EQ factor. *Time*, 60–68.
- Goleman, D. (1995). *Emotional intelligence*. New York: Bantam.
- Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.
- Guttman, L., & Levy, S. (1991). Two structural laws for intelligence. *Intelligence*, *15*, 79–103.
- Hall, J. A., & Halberstadt, A. G. (1994). "Subordination" and sensitivity to nonverbal cues: A study of married working women. *Sex Roles*, *31*, 149–165.
- Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41–45.
- Horn, J., & Noll, J. (1994). A system for understanding cognitive capacities: A theory and the evidence on which it is based. In D. K. Detterman (Ed.), *Current topics in human intelligence: Theories, tests, and issues* (Vol. 4). New York: Guilford Press.
- Jensen, A. R., & Weng, L. (1994). What is good *g*? *Intelligence*, *18*, 231–258.
- LaFrance, M., & Banaji, M. (1992). Towards a reconsideration of the gender-emotion relationship. *Review of Personality and Social Psychology*, *14*, 178–201.
- LaFrance, M., & Hecht, M. A. (in press). Option or obligation to smile: The effects of power and gender on facial expression. In P. Philippot, R. Feldman, & E. Coates (Eds.), *The social context of non-verbal behavior*. Cambridge: Cambridge University Press.
- Legree, P. J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence*, *21*, 247–266.
- Mayer, J. D. (1998). A systems framework for the field of personality. *Psychological Inquiry*, *9*, 118–144.
- Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, *22*, 89–113.
- Mayer, J. D., & Mitchell, D. C. (1998). Intelligence as a subsystem of personality: From Spearman's *g* to contemporary models of hot-processing. In W. Tomic & J. Kingma (Eds.), *Advances in cognition and educational practice: Conceptual issues in research on intelligence* (Vol. 5, pp. 43–75). Greenwich, CT: JAI Press.
- Mayer, J. D., & Salovey, P. (1993). The intelligence of emotional intelligence. *Intelligence*, *17*, (4), 433–442.
- Mayer, J. D., & Salovey, P. (1995). Emotional intelligence and the construction and regulation of feelings. *Applied and Preventive Psychology*, *4*, 197–208.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3–31). New York: Basic Books.
- Mayer, J. D., & Stevens, A. (1994). An emerging understanding of the reflective (meta-) experience of mood. *Journal of Research in Personality*, *28*, 351–373.
- Mayer, J. D., DiPaolo, M. T., & Salovey, P. (1990). Perceiving affective content in ambiguous visual stimuli: A component of emotional intelligence. *Journal of Personality Assessment*, *54*, 772–781.
- Mayer, J. D., Salovey, P., & Caruso, D. R. (1997). *Emotional intelligence test*. Needham, MA: Virtual Knowledge [Producer and Distributor] [CD-ROM].
- Mayer, J. D., Carlsmith, K. M., & Chabot, H. F. (1998). Describing the person's external environment: Conceptualizing and measuring the life space. *Journal of Research in Personality*, *32*, 253–296.

- Mayer, J. D., Salovey, P., & Caruso, D. R. (in press). Competing models of emotional intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (2nd edn.). New York: Cambridge.
- McCrae, R. R., & Costa, P. T. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509–516.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*, 525–543.
- Morrison, D. F. (1976). *Multivariate statistical methods* (2nd edn.). New York: McGraw-Hill.
- Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Lochlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77–101.
- Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.
- Paulhus, D. L., Lysy, D. C., & Yik, M. S. M. (1998). Self-report measures of intelligence: Are they useful as proxy IQ tests? *Journal of Personality*, *66*, 525–554.
- Payne, W. L. (1986). A study of emotion: Developing emotional intelligence; self-integration; relating to fear, pain and desire. *Dissertation Abstracts International* (University Microfilms No. AAD9-5947), *47*, (1-A), 203A.
- Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 197–219) Hillsdale, NJ: Erlbaum.
- Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science*, *1*, 86–89.
- Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P., & Archer, D. (1979). *Sensitivity to nonverbal communication: A profile approach to the measurement of individual differences*. Baltimore: Johns Hopkins University Press.
- Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition and Personality*, *9*, 185–211.
- Salovey, P., Mayer, J. D., Goldman, S., Turvey, C., & Palfai, T. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. In J. W. Pennebaker (Ed.), *Emotion, disclosure, and health* (pp. 125–154) Washington, DC: American Psychological Association.
- Scarr, S. (1989). Protecting general intelligence: Constructs and consequences for interventions. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy*. Urbana: University of Illinois Press.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.
- Spinoza, B. (1984). "Ethics (Part III): On the origin and nature of emotions." Reprinted in C. Calhoun & R. C. Solomon (Eds.), *What is emotion?* Oxford: Oxford University Press.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Penguin Books.
- Sternberg, R. J. (1994). Commentary: Reforming school reform: Comments on "Multiple intelligences: The theory in practice." *Teachers College Record*, *95*, 561–569.
- Sternberg, R. J., & Caruso, D. R. (1985). Practical modes of knowing. In E. Eisner (Ed.), *Learning and teaching the ways of knowing: 84th Yearbook of the National Society for the Study of Education* (Part II, pp. 133–158). Chicago: University of Chicago Press.
- Sternberg, R. J., & Smith, C. (1985). Social intelligence and decoding skills in nonverbal communication. *Social Cognition*, *3*, 168–192.
- Thorndike, E. L. (1920). Intelligence and its uses. *Harper's Magazine*, *140*, 227–235.
- Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*, 436–458.
- Wagner, H. L., MacDonald, C. J., & Manstead, A. S. R. (1986). Communication of individual emotions by spontaneous facial expression. *Journal of Personality and Social Psychology*, *50*, 737–743.
- Wechsler, D. (1958). *The measurement and appraisal of adult intelligence* (4th edn.). Baltimore, MD: Williams and Wilkins.
- Yerkes, R. M. (1921). Psychological examining in the United States. *Memoirs of the National Academy of Sciences* (15: 890 pp. entire).
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*, 151–175.