Emotion
2001, Vol. 1, No. 3, 232–242

1528-3542/01/$5.00 DOI: 10.1037//1528-3542.1.3.232

Emotional Intelligence as a Standard Intelligence

John D. Mayer
University of New Hampshire

Peter Salovey
Yale University

David R. Caruso
Work-Life Strategies

Gill Sitarenios
Multi-Health Systems, Inc.

The authors have claimed that emotional intelligence (EI) meets traditional standards for an intelligence (J. D. Mayer, D. R. Caruso, & P. Salovey, 1999). R. D. Roberts, M. Zeidner, and G. Matthews (2001) questioned whether that claim was warranted. The central issue raised by Roberts et al. concerning Mayer et al. (1999) is whether there are correct answers to questions on tests purporting to measure EI as a set of abilities. To address this issue (and others), the present authors briefly restate their view of intelligence, emotion, and EI. They then present arguments for the reasonableness of measuring EI as an ability, indicate that correct answers exist, and summarize recent data suggesting that such measures are, indeed, reliable.

In 1999, we published "Emotional Intelligence Meets Traditional Standards for an Intelligence" in the journal *Intelligence* (Mayer, Caruso, & Salovey, 1999). In that article, we presented a new scale of emotional intelligence (EI), the Multi-Factor Emotional Intelligence Scale (MEIS; Mayer, Caruso, & Salovey, 1999) that was based on a decade of theoretical and empirical work. We argued, on the basis of the MEIS and findings with it, that EI was a lot like a traditional intelligence. First, it could be measured as an ability for which there were correct answers. Second, the domain of EI was sizeable in that we could come up with 12 fairly diverse tasks to measure it—everything from recognizing emotion in faces to understanding how emotions are likely to change over time. Third, after administering the test to 503 adults and 229 adolescents, we found that those 12 diverse tasks were positively correlated. A factor analysis of those tasks indicated that they could be defined by one general factor and that they also fell into three or four subgroups of skills roughly corresponding to our model of EI (Mayer & Salovey, 1997). Finally, EI ability increased with age, at least across the age ranges that we explored in cross-sectional studies.

Roberts, Zeidner, and Matthews (2001) have questioned what kind of an intelligence EI may be—if it is an intelligence at all. One should start with the fact that there is a great deal of agreement between their findings and our own. In their words (pp. 223–224):

> Some features of the psychometric analyses support Mayer, Caruso, and Salovey's (1999) claim that EI meets criteria for an intelligence. We replicated the finding of a positive manifold between subtests of the MEIS, and, generally, the pattern of correlations corresponded well to the Mayer, Caruso, and Salovey (1999) findings. Exploratory and confirmatory factor analyses showed broad similarities with Mayer et al.'s factor solutions, although there were some differences in detail, and, in exploratory analyses, subscale communalities were often low. In fact, the confirmatory analyses tend to support Mayer et al.'s initial conception of four branches of EI, rather than the three-factor model that has subsequently been derived.

In fact, the MEIS represented a great step forward for us in relation to our earlier ability scales (e.g., Mayer, DiPaolo, & Salovey, 1990; Mayer & Geher, 1996). The MEIS included many tasks that Roberts et al. (2001), and we, found intercorrelated well. The overall Cronbach's alpha of the factor-based scale

John D. Mayer, Department of Psychology, University of New Hampshire; Peter Salovey, Department of Psychology, Yale University; David R. Caruso, Work-Life Strategies, New Canaan, Connecticut; Gill Sitarenios, Research and Development, Multi-Health Systems, Inc., Toronto, Ontario, Canada.

Correspondence concerning this article should be addressed to John D. Mayer, Department of Psychology, University of New Hampshire, Conant Hall, 10 Library Way, Durham, New Hampshire 03824. Electronic mail may be sent to jack.mayer@unh.edu.

representing the entire test was .96 (Mayer et al., 1999, p. 286). Also, one of the great advances of the MEIS over our earlier, more limited ability measures, was the first attempt to introduce an "expert" criterion for deciding on a correct answer, although at the time, we only had two such experts. In earlier studies, we relied on general consensus scoring; basically, the degree to which the individual agreed with the response of the general group was the index of correctness. In expert scoring, two of the authors provided their own estimation of the optimal answers to the test. (We also examined a target form of scoring in which, e.g., a person whose face was being examined with regard to the emotions it expressed, reported how she or he felt.)

Still, in reviewing the psychometrics of the MEIS, Roberts et al. (2001) found much to be concerned about. Again, in their words (p. 224):

> . . . other aspects of the data render many of the EI concepts more problematic than is acceptable for ability measures . . . In particular, the reliability of subtests that form the highest branches of the model, and are thus probably the most important components of the MEIS for prediction of real-world social behaviors (e.g., progressions, managing others), is among the poorest in this battery. In addition intercorrelations between subtests, although resulting in positive manifold, are notably lower than is common in research involving ability measures (compare, for example, data presented here with various data sets presented in Carroll, 1993). Further, various factor analyses indicate a structure that is relatively unstable, certainly when compared with similar analyses that have been conducted with intelligence and personality measures.
>
> Perhaps the most severe psychometric difficulty is the lack of convergence between expert- and consensus-scored dimensions. There are instances of agreement, especially for the Blends and Progressions tests, but in general, cross-correlations are too small to suggest convergence. The correlation between the general factors extracted from each of the two data sets was only .26.

Their findings led them to several conceptual points of importance, raised mostly in their discussion. Roberts et al. (2001) concluded that there may be no objective answers to EI tests, and because correct answers are scored on the basis of group consensus, EI does not qualify as an intelligence. Perhaps, the authors suggest, EI actually measures some form of conformity in relation to the group. As previously noted, they were also concerned about the reliability of such scales. We first briefly restate our view of EI. Then, we address Roberts et al.'s concerns by considering the following: (a) Are there really correct answers on tests of EI? (b) If, as we argue, consensus scoring is ideal, is EI merely a measure of conventionality? (c) Finally, are the tests reliable, and if not, can they be made to be reliable? We continue to assert that EI is, indeed, a traditional intelligence. Some of our arguments are theoretical, others rely on further findings from new studies we have conducted (Mayer, Salovey, Caruso, & Sitarenios, 2001).

An Overview of the Concept of EI

The Nature of Intelligence

Symposia on intelligence over the years have repeatedly concluded that the first hallmark of intelligence is abstract reasoning (Sternberg, 1997). That is, intelligence involves such capacities as seeing the similarities and differences among objects, being able to analyze parts and appreciate their relation to each other and as a whole, and generally, being able to reason validly within and across content domains. Abstract reasoning, although the core aspect, is assisted by several other functions. Several such adjunct areas are of relevance here: input, knowledge base, and meta-strategies (for a more detailed view, see Mayer & Mitchell, 1998). These are enumerated in Table 1.

First, abstract reasoning cannot take place without an input function. Different intelligences are often defined according to what is input and processed. For example, verbal intelligence pertains to reasoning about language; spatial intelligence pertains to reasoning about the position and movement of objects in space. Whatever the area, something must get the information—be it verbal, spatial, or emotional—into the system. Second, abstract reasoning is assisted by a well-organized, related body of knowledge: the knowledge base. This was what Cattell originally referred to as *crystallized intelligence* (Ackerman, 1997; Cattell, 1943). Third, there are *meta-cognitions*—basically, strategies for operating with an intelligence in the context of broader mental life. A meta-cognition may be that it helps, in analyzing a problem, to write down portions of the problem so that not everything needs to be kept in short-term memory.

The Nature of Emotion and Emotional Information

There is considerable diversity of opinion as to what emotion is (Frijda, 2000; Solomon, 2000). A reasonably canonical definition, however, might be that an emotion is an organized mental response to an event that includes physiological, experiential, and

Table 1
A Summary Overview of Parts of Intelligence

Aspect of intelligence	Examples from verbal intelligence	Examples from emotional intelligence
Meta-processing (adjunct)	Knowing that writing something down can help one remember it	Knowing that helping someone may make oneself feel better
Abstract understanding and reasoning (core)	Being able to identify the protagonist of a story and compare the individual with other people	Being able to analyze an emotion and identify its parts and how they combine
Knowledge base processing (adjunct)	Having knowledge (and remembering analyses) of prior instances of stories	Having knowledge (and remembering analyses) of prior instances of feelings
Input processing (adjunct)	Being able to keep long sentences in memory	Being able to perceive emotions in faces

Note. This summary overview follows Mayer & Mitchell, 1998; Table 1.

cognitive aspects, among others. Of particular importance to us is that emotions typically occur in the context of relationships (cf. Lazarus, 1991). One is angry if blocked from attaining a goal, happy if loved by someone who one loves in return, afraid when threatened, and the like. These relationships may be entirely internal, as when one is afraid of what one might do, or external, as when one admires another person. If emotions often arise in relationships, then emotional information is information about certain forms of these relationships.

One critical aspect of emotional information, is its consistency across people. Compelling cross-cultural research by Ekman (1973) has supported Darwin's hypothesis that emotional expression has evolved across species (Darwin, 1872/1965). This strongly implies that emotional information—and the capacity to read it—would show some universality across human beings and even closely related mammalian species. Ekman argued that recognition of facial emotional expression was universal. Any apparent differences in human emotional expression from culture to culture could be attributed to the fact that different societies teach different display rules about appropriate moments to express certain feelings. Additional evidence for the regularity of emotional information can be found in the area of artificial intelligence, where cognitive scientists created expert systems that could understand emotions in rudimentary stories (Dyer, 1983). We deal with the nature of emotional information in greater detail below, as we discuss the issue of the "correct answer" to an EI test item.

The Nature of EI

Our model of EI begins with the idea that emotions contain information about relationships. (Other models of EI exist as well; see Mayer, Salovey, & Caruso, 2000b, for a review.) When a person's relationship with another person or an object changes, so do their emotions toward that person or object. A person who is viewed as threatening is feared, an object that is favored is liked. Whether these relationships are actual, remembered, or even imagined, they are accompanied by the felt signals called emotions. EI, in turn, refers to an ability to recognize the meanings of emotions and their relationships and to use them as a basis in reasoning and problem solving. It further involves using emotions to enhance cognitive activities (Mayer et al., 1999).

Our own analysis of emotion-related abilities led us to divide EI into four areas of skills (e.g., Mayer & Salovey, 1995, 1997; Salovey & Mayer, 1990). We call these areas *branches* in reference to the diagrams in which they were first introduced. The four-branch model that we now use divides EI into four areas: accuracy in (a) perceiving emotions, (b) using emotions to facilitate thought, (c) understanding emotions, and (d) managing emotions in a way that enhances personal growth and social relations. We view a distinction between the second branch (using emotions to facilitate thought) and the other three. Whereas Branches 1, 3, and 4 involve reasoning about emotions, Branch 2 uniquely involves using emotions to enhance reasoning. Finally, we view the four branches as forming a hierarchy, with emotional perception at the bottom and management at the top. This four-branch model serves as a basis of our current reviews of the field (e.g., Mayer, 2001; Salovey, Bedell, Mayer, & Detweiler, 2000; Salovey, Mayer, & Caruso, in press; Salovey, Woolery, & Mayer, 2001). With this thumbnail sketch of our model, let us proceed to Roberts et al.'s (2001) concerns.

Which Areas of EI Should Correlate Highest With Cognitive Intelligence?

A clarification of one aspect of our EI model can explain a result obtained by Roberts et al. (2001) that they found to be unexpected. Recall that our four-branch model of EI is hierarchical in the context of an individual's personality. The four branches are briefly described in Table 2. There, emotional understanding is most allied with cognitive processing and abstract reasoning; it is most cognitively saturated. Emotion management, although the highest branch, creates an interface between the cognitive system and the more general personality system. As such, emotion management is actually less cognitive than emotional understanding, because it must balance many factors including the motivational, emotional, and cognitive (Mayer, 2001). Roberts et al. expected the MEIS's emotion management score to have the highest correlation with general IQ, because it should be "most cognitive." In fact, however, our model supposes that the third branch, understanding, is most cognitive and should have the highest relation to abstract reasoning. That, in fact, is what the data show; the third branch does correlate most highly with IQ.

Are There Correct Answers to Our EI Tests?

Roberts et al. (2001) identified "the most severe psychometric difficulty" with our work to be "the lack of convergence between expert- and consensus-scored dimensions" (p. 224). Let us begin in earnest with this issue for two reasons. First, it is primary to whether there can be a correct answer to a test of EI, and hence, whether EI is a standard intelligence. Second, it is, by their own description, their most important criticism. As their own concern began with the empirical finding that consensus and expert data correlated only $r = .26$ (p. 224) in their sample, let us start with the empirical issue. Roberts et al. noted, "The discrepancies are sufficiently large that they imply that one or the other scoring method should be discarded, in that it is hard to envisage modifications that would bring factors that are correlated at less than .50 into alignment" (p. 226).

Empirical Concerns

Roberts et al.'s (2001) findings suggested the unlikelihood of any eventual convergence between scoring methods. At the same time, the MEIS was an experimental measure, a first attempt to operationalize our full model of EI. The addition of rudimentary expert scoring was intended as an exploration of the possibility of another criterion of correctness and was not intended as a final expert criterion. We recognized that two authors answering a long test would be unlikely to create the most optimal version of the expert criterion possible. Indeed, Legree (1995) pointed out that individual experts are typically unreliable. His own research indicated that as experts are aggregated they might be expected to approach the general consensus in this domain. At the time our first article on the MEIS appeared, we argued that expert scoring and consensus scoring converged to some degree and that because the general group consensus appeared more reliable, and yielded better test-factor structure, it should be used (Mayer et al., 1999, pp. 283–284).

Table 2
Overview of the Four-Branch Model of Emotional Intelligence, With a Focus on Its Relation to Intelligence and Personality

Branch	Description of measure	Relation to intelligence and personality
4: Managing emotion	Ability to manage emotions and emotional relationships for personal and interpersonal growth	Interface with personality and personal goals
3: Understanding emotion	Ability to comprehend emotional information about relationships, transitions from one emotion to another, linguistic information about emotions	Central locus of abstract processing and reasoning about emotions and emotional information
2: Facilitating thought with emotion	Ability to harness emotional information and directionality to enhance thinking	Calibrates and adjusts thinking so that cognitive tasks make use of emotional information
1: Perceiving emotion	Ability to identify emotions in faces, pictures	Inputs information to intelligence

More recently, we have developed a new test of EI, the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT; Mayer, Caruso, & Salovey, 2000) that attempts to improve on the psychometric qualities of the MEIS. Findings from this new scale are reported in an unpublished manuscript presently under review for this journal (Mayer et al., 2001). That manuscript, a draft of which is currently available from the authors, reports two large-sample psychometric studies. The second study concerns the MSCEIT V2.0, a 141-item ability scale that uses tasks similar to those of the MEIS to measure the four branches of EI. Like the MEIS, the MSCEIT can be scored according to a general consensus criterion. That is, if .56 of the sample says that there is a moderate amount of happiness in a face, and a participant agrees, his or her score is incremented by .56.

Also as with the MEIS, we used expert scoring for the MSCEIT. Rather than use two authors as experts, however, we asked 21 members of the International Society of Research in Emotion (ISRE) to answer the MSCEIT questions. We then scored the MSCEIT according to an expert-consensus criterion on the basis of the proportion of experts from ISRE who answered each item in a particular way. As reported in our manuscript (2001), when over 2,000 participants' scores on the MSCEIT were calculated by general and expert-consensus scoring, the intercorrelation between the two sets of scores was $r = .98$. That figure is, obviously, well above the $r = .26$ figure that Roberts et al. (2001) featured in their discussion. Perhaps it is less surprising when one realizes that the convergence between general and expert-consensus scoring on the MEIS was between $r = .43$ to .78 for three of the four branches in their data set, with Branch 1's poor convergence ($r = .02$) bringing down the other branches to an overall convergence of $r = .48$ (p. 214). The $r = .26$ figure Roberts et al. (2001) featured in their discussion was apparently a consequence of further transformation of the data into factor scales; see p. 217.

Theoretical Concerns

The finding that general- and expert-consensus scoring effectively converge does not, by itself, solve whether there are correct answers to questions assessing EI. The finding does, however, greatly simplify the issue. If expert scoring is very close or even the same as consensus scoring, the question can now be refocused as "What does that consensus mean?" and "Is this form of determining a correct answer much different than that used in cognitive intelligence tests?"

It is worth noting at the outset that cognitive IQ tests have items that are "objectively scored," freely intermixed with tests that are scored by a (presumably expert) consensus. An objectively scored test would be something like "digits forward" on the Wechsler Adult Intelligence Scale (3rd ed.; WAIS-III; Psychological Corporation, 1997) for which participants hear a series of digits (3 . . . 5 . . . 2) and simply must repeat them. Other WAIS subtests, however, require some discussion to arrive at the correct answer. As the WAIS-III manual puts it,

> To refine the scoring criteria of those subtests for which many acceptable responses are possible (Vocabulary, Similarities, Information, and Comprehension on the WAIS-III . . .), the development team conducted several scoring studies . . . . Two team members independently coded each response, identified discrepancies between the code assignments, and resolved the differences so that each response had only one code. At this point, team members had to agree on the grouping of responses and the assignment of codes but not on what score value to assign a code. . . . After the codes were assigned, the team evaluated the quality of the responses and assigned a score value (0, 1, or 2) to each code on the basis of the accuracy of the response." (Psychological Corporation, 1997, p. 37)

The issue then is not whether experts need to be used, but rather the nature of emotional versus cognitive information, and the nature of emotional versus cognitive experts.

Similarities and Differences Between Emotional and Cognitive Information

When thinking about general and expert consensus scoring of emotional and cognitive tasks, there seem to be three issues that come into play: (a) domain of application, (b) general consensus versus expert consensus convergence, and (c) systematization of knowledge.

Differences in domain of application. Emotional and cognitive information differ in regard to their domain of application. Emotional information pertains to the human world: the relations of people (and animals) to one another, and their relation to cultural institutions, ideas, artifacts, and rules of behavior. Put another way, emotional information applies primarily to matters of how human beings and their evolutionary ancestors survive and interact with the immediate living world. Human beings have likely come to a general consensus about many emotional meanings.

This does not imply that there is only one way to feel or interpret feelings, but rather that it helps to know how an individuals' reactions compare with how most people would emotionally respond to a situation. Such knowledge helps define the general meaning of emotions in regard to relationships; for example, that happiness arises in harmonious environments, fear in response to threat, anger to goal blockage, and so on.

Cognitive information, in contrast, describes rules that have areas of application far beyond our immediate living world. Although the child's mathematical world may begin, in part, with counting fingers and toes, the world of mathematics refers to a sometimes imaginary but often useful numerical world of equations, arithmetical functions, and physically possible and impossible spaces, created both for their beauty and their application. Laws that draw on mathematics to describe gravity, acceleration, and the like, apply as much to rocks, stars, and other galaxies as they do to human beings.

Differences in expertise. This difference in domain of application has implications for the type of expertise in each area. Whereas in the general cognitive realm objects are studied, in the emotional realm, people are studied. This gives rise to the second difference, what we refer to as the issue of general versus expert convergence. Given that emotional information is biosocial (i.e., biological and learned), the expert becomes the expert to some extent by studying the group consensus, and becoming, as it were, more accurate about the group consensus than other individuals or small groups. For example, in the area of emotions, experts can reliably distinguish genuine from false, tense smiles (cf. Ekman, 1985). It is likely, however, that the general consensus can do this as well. The decisive contribution of expertise is probably to identify and distinguish between the two sorts of smiles more proficiently than the average person can, and to elucidate how people do this naturally. That is, the general group will be "messier" in identifying the consensus than experts would be. General–expert convergence occurs in the domain of emotions, however, because experts look for the correct answer by paying attention to the consensual information of the group.

It is easy to come up with instances, in teaching physics, for example, where the group consensus is simply wrong and the expert opinion is correct. In fact, physicists refer to *lay physics* to represent the sometimes incorrect but consensual notions that people hold about the physical qualities and motions of objects. A well-known example of this is the common misconception (at least among young children) that heavier objects fall faster than lighter ones. Here is a case where expert knowledge plainly trumps lay knowledge, as, since Galileo's famous experiments, it has been known that the two sorts of objects fall at the same rate. The difference here is that the expert in physics, for example, conducts experimental research in areas (i.e., the behavior of objects) of which the general person has no innate or pragmatic experience.

It is likely possible to come up with parallel instances of emotional expertise as the lay physics example. Suppose we ask "Why did Billy beat up Bobby?" Because Billy (a) was unhappy about himself, (b) felt very good about himself, or (c) was afraid. The lay answer is likely to be "a," on the basis of early psychological theories about bullying. More recent evidence, however, has suggested that "b" may be the correct answer, as bullies tend to have high self-esteem (e.g., Baumeister, 1997, pp. 149–153). Making such a claim, however, troubles emotions researchers more so, we suspect, than it would physicists. This is partly because discoveries in psychology are fraught with contention and have often been reversed as more knowledge is accumulated. This, of course, is also true of physics—after all, in the ancient Greek period, the expert answer to the problem of what sort of objects fell faster was that heavier objects fell faster. In addition, however, part of the problem relates to what we would call systematization of knowledge.

Differences in systematization and institutionalization of knowledge. The third difference between information (and hence, expertise) in the emotional and the cognitive domain is systematization and institutionalization of knowledge. By systematization, we refer to authoritative dictionaries, manuals, descriptions of operations, and other texts dealing with the subject matter from a consistent viewpoint. Emotional information has certainly been systematized to some degree, but there is less than universal agreement as to the systematizations. For example, one can trace various enumerations of the meanings of emotions through philosophy, psychology, and now, artificial intelligence programs that decipher emotions in stories. Still, these authoritative manuals are not well known, are only beginning to be taught in schools, and have not yet gained widespread cultural currency. In short, they are not entirely culturally sanctioned. Cultural reasons for this lag are described elsewhere (Mayer, Salovey, & Caruso, 2000a).

There has been a higher level of systematization of certain areas of cognitive information than areas of emotional information. Western schools, colleges, and universities focus on areas such as language, literature, and mathematics in ways that they simply do not focus on emotions. In part for that reason, more attention has been placed on creating canonical resource materials in literary, mathematical, and similar areas. With resources such as dictionaries and textbooks, fields such as language, history, and mathematics appear more fixed, certain, and objective than does the emotions area. Imagine, for a moment, that investigators in the field of emotions were required to establish a high school curriculum for teaching about the emotions and their meanings (Cohen, 1997, 2001; Elias et al., 1997). Committees would meet, commissions would be established, and ultimately, an authoritative expert body of knowledge would be produced. Biologists, chemists, and physicists have been doing this for years, with the tacit understanding that the body of expert knowledge will change over the years but that each iteration represents a further approximation of some ultimate truth. When this begins happening in emotions, it is likely that the divergence between experts and the general consensus will be easier to describe and detect.

Do the Tests Measure EI or Emotional Conformity?

If emotional information turns out to be the general consensus, and is little different from the expert consensus, then have we, as Roberts et al. (2001) wondered, created a test of conformity? In their words,

> A conformity construct is of real-world relevance, but it is highly misleading to label it as an intelligence, because it relates to person–environment fit rather than to any characteristic of the individual. Indeed, in some instances it is the nonconformist who should be deemed emotionally intelligent, for example, a writer or an artist who finds a new and original way of expressing an emotion. (Roberts et al., 2001, p. 227)

Let us begin by drawing a distinction between convergent thinking and conformist or conventional thinking. Convergent thinking is the capacity to pinpoint a correct answer or answers. Conventional thinking is the limitation of one's perspective. Guilford (1959) noted that most existing intelligence tests measure convergent thinking. For example, in reading comprehension, the correct answer is getting the same point as everybody else (or as the established experts have). In this sense, our test of EI is also convergent. But that does not mean that high scorers on the MEIS are conventional, any more than high scorers on the WAIS-III are necessarily conventional. High scorers typically can reason well beyond the designated answers they provide.

Our theory of EI states that emotional reasoning begins with perceiving emotions accurately. We do not see this as much different than saying that literary analysis begins with comprehending the basic content in stories or that spatial relations begins with being able to name shapes. No one argues that such simple identification of basic linguistic meanings or basic shapes is conformity. EI continues with abstract reasoning about emotions. This includes analyzing linguistic terms relevant to emotion, and analyzing alterations in emotional sequences that are likely to occur (e.g., that delay turns frustration into anger). This is full-fledged reasoning, different perhaps, but still arguably on a par conceptually with, say, understanding the proper order of an argument. If the examples of such emotional reasoning are, perhaps, a bit more limited than those found in linguistic or spatial reasoning, it is probably in large part due to the relative paucity of systematization in the emotions area. We expect more such systematization in the future, however, as people continue to come to grips with the importance of the area of emotion. Even with such systematization that presently exists, we have been able to come up with enough items to write two entirely different EI tests of hundreds of items, each of which works fairly well (i.e., the MEIS and MSCEIT series).

If we believe the test involves intelligence, then where does creativity enter in? Does it not depend on idiosyncratic emotional reactions? As Roberts et al. (2001) asked, should the nonconformist—a writer or an artist who finds a new and original way of expressing an emotion—be deemed emotionally intelligent? Absolutely, but note that discovering a new way of expressing an emotion doesn't necessarily involve inventing new emotional rules or having idiosyncratic emotional reactions. In music, for example, creativity may involve arranging musical tones in a new way to elicit old or new emotions and their blends. In fact, it is the very capacity of the writer or artist to portray relations among people and to describe things so as to come up with new blends of feelings or transitions among emotions, which in part marks the creative process. A composer rarely exercises creativity in writing music by playing consistently off key; a writer rarely creates deep, new feelings in a reader by failing to understand how a reader would react emotionally to the character in a similar situation.

Consider Averill and Thompson-Knowles' (1991) triad task of emotional creativity. In that task, participants were asked to create a story in which there are relationships that represent particular triads of emotions. In response to the emotions "serene, bewildered, and impulsive," one participant, who was rated quite highly on creativity, wrote the following:

> The clouds are few, the sky is clear. I'm at the top of the cliff. It's real peaceful up here. Suddenly, I want to jump. I don't know why, I just want to. Calmly, I look down at what would be my unquestioned doom. It looks peaceful; warm and friendly. But why, why do I want to dive into the hands of the grim reaper? What does this mean? I hesitate, then motion to jump, something strange pulls me back. It is the peacefulness of the cliff. I can't destroy the peacefulness. The wind feels like velvet against my skin as I slowly shake my head. Why? (p. 280)

This passage was rated high in novelty, authenticity, and creativity, yet what makes it work is that its anonymous author understood the meanings of emotion in a convergent sense. What was novel was the creation of a new situation to elicit the feelings.

Are the Tests Reliable?

Reliability at the Full-Scale, Branch, and Individual Task Level

The MEIS and our subsequent tests, the MSCEIT RV1.0 and V2.0, are all reliable tests at the full-scale level (where they all possess split-half reliabilities above $r = .90$). Our aforementioned manuscript (Mayer et al., 2001) reports MSCEIT V2.0 reliabilities at the branch level ranging from $r = .79$ to .91 using consensus scoring and $r = .77$ to .90 using expert scoring. The progression of tests from MEIS, studied by Roberts et al. (2001), to our newer tests, the MSCEIT RV1.1 and MSCEIT V2.0, also showed a gradual rise in reliability at the level of their individual tasks. Whereas using general consensus scoring on the 12 tasks of the MEIS yielded individual-task reliabilities ranging from alphas of .49 to .94 (Mayer et al., 1999), using the same scoring method for the MSCEIT V2.0 yields individual task alphas from $r = .64$ to .88 (Mayer et al., 2001 review).

Davies, Stankov, and Roberts (1998) raised concerns regarding reliability of an early 1990 ability test, which measured one aspect of EI. That test provided an empirical demonstration of the possibility of EI rather than a fully operationalized test (Mayer et al., 1990). Given our primary focus at that time, which was to demonstrate the existence of EI psychometrically, we were not too worried about the admittedly modest alpha of .63 of this rather constrained 1990 measure. To us, if EI existed, it is theoretically important, in part, because it brings a new perspective to the field's view of both emotion and intelligence. After all, almost any test's reliability can be enhanced simply by making it longer (e.g., the Spearman–Brown prophecy formula; Nunnally, 1978, pp. 210–211). That said, of course we wanted tests to be used by others in the research and applied worlds. For that reason, we worked hard to ensure that the MEIS and MSCEIT had full-scale reliabilities over .90, and we succeeded.

Roberts and his colleagues were presumably satisfied with the MEIS full-scale reliability of $r = .96$ that we reported in the 1999 article. Their focus, however, is almost exclusively, or entirely, on the reliability of the smallest components of the test, that is, individual subtests rather than on the branch scores and total test levels where reliabilities range from very good to excellent. Roberts et al. (2001) expected all the subcomponents of the test, even at the level of individual tasks, to be of uniformly high reliability. It is important to note that this has little to do with whether EI exists. Rather, it appears to be, for them, an issue of utility. Their perspective, as it is clarified in their article in this issue, appears to have been applied. They wanted to be assured that when people are tested, the scores those individuals obtain (i.e., are told about) at the individual task level are legitimate and accurate reflections of their ability. This is an obviously legitimate concern on their part. We recognize it to be of growing importance as tests such as the MSCEIT are prepared for general use.

Reliability or Accessibility?

Given that test reliability can be improved simply by lengthening the test, why did we not heed Davies et al.'s (1998) criticisms early on and make a test longer than the MEIS that was highly reliable even at the level of the individual task? We could have taken the MEIS (or the MSCEIT tests) and added items to them to do so. Instead, we chose to shorten our tests and build up reliability at the task level through careful item selection. The reason we did so was that investigators have begun to, and will need to, correlate the test with real-life criteria. A short, efficient test that provides reliable scores at the branch and total-test level, like MSCEIT V2.0, can stimulate research better and faster than a longer, more unwieldy and inefficient test can that has optimized reliability for every individual task. Through careful item selection, we have actually improved reliability at the task level (this was accomplished in part by dropping four tasks). To adopt Roberts et al.'s (2001) perspective for a moment, it is worth comparing the present MSCEIT V2.0 with the original WAIS (Wechsler, 1955) used until 1981. The WAIS was written after roughly 4 decades of experience with intelligence testing and replaced the earlier Wechsler–Bellevue Intelligence Test. The 11 WAIS subscale scores ranged in reliability from $r = .60$ to $r = .96$, not so different from the MSCEIT V2.0 (Matarazzo, 1972, p. 239).

Ongoing Issues

There are a few issues we are not yet prepared to address in this commentary because they are empirical questions that are yet to be resolved. To many people, the most important of these issues would probably concern whether the general and expert consensuses are uniform or cultural-bound in Western and non-Western societies. There are several additional issues that will chiefly be of interest to intelligence researchers. For example, why does emotional perception load most highly on factors of general EI? Here, our current work suggests that this finding may not hold up with our newer tests, and therefore we have not commented on it. Another item of chiefly psychometric interest concerns whether the average interrelation among EI tasks is lower than that for general cognitive IQ and, if so, why? This is an interesting question for which we have no answer at this time. We address concerns about the factor structure of the area in our aforementioned manuscript now under review for this journal. Finally, we continue to believe that developmental aspects of intelligence are a defining feature of them but have not commented on that because of limitations of space.

Summary and Conclusions

In this commentary, we have restated our conceptual approach to intelligence, emotions, and emotional intelligence. Roberts et al. (2001) discussed certain concerns about our theory and measurement of EI; to a large degree, those concerns centered on the low correlation between two methods of scoring the MEIS (general consensus and expert). In particular, Roberts et al.'s concerns raised the issue of whether, given the divergence of the scoring procedures, one could decide on correct answers for a test of EI. We are appreciative of the fact that Roberts, Zeidner and Matthews (2001) are engaged in this research and have raised these issues. Their work has encouraged us to present new data and to present a further consideration of the theoretical issues involved in scoring such tests. We have referred to important findings indicating that different scoring methods converge at the $r = .98$ level. In addition, we provided a plausible theoretical explanation of the basis on which correct answers to EI tests can be determined.

In addition, we considered some other concerns of Roberts et al. (2001), including test reliability. Although the reliability for the full MEIS is $r = .96$, it is lower for the 12 individual tasks. We have argued that the reliability issue raised by Roberts et al. is a limited problem and one that is addressed by our new scale, the MSCEIT V2.0, which is now available for research.

With such an assessment tool, it is now possible to ask not only whether EI exists, but whether it is important in various realms of our lives. We expect EI to be an important predictor of significant outcomes. What research does exist suggests that EI is likely to take its place alongside other important psychological variables as a predictor of various outcomes at school, home, and work. For example, higher EI may predict reduced levels of problem behavior such as drug use and interpersonal violence (Brackett, 2001; Formica, 1998; Mayer, Perkins, Caruso, & Salovey, 2001; Rubin, 1999; Salovey, Mayer, Caruso, & Lopez, in press; Trinidad & Johnson, in press).

The development and understanding of an intelligence requires a number of years of careful scrutiny and research. The most widely used cognitive scales of intelligence, the Wechsler Intelligence scales, are the product of 60 years of research. Moreover, that research itself was initiated after an earlier 40 years of work on the clinical assessment of intelligence. If the history of the study of intelligence is any guide, there is little question that there is much still to be learned about EI. The first 10 years of EI research have been frustrating but also immensely rewarding and full of promise. We look forward to continued research in the field and to learning more about what EI predicts.

References

Ackerman, P. L. (1997). Personality, self-concept, interests, and intelligence: Which concept doesn't fit? *Journal of Personality, 65,* 171–204.

Averill, J. R., & Thompson-Knowles. (1991). Emotional creativity. In K. T. Strongman (Ed.), *International review of studies on emotion* (Vol. 1, pp. 269–299). London: Wiley.

Baumeister, R. F. (1997). *Evil: Inside human violence and cruelty.* New York: Freeman.

Brackett, M. (2001). *Emotional intelligence and the life* *space.* Unpublished master's thesis, University of New Hampshire, Durham.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153–193.

Cohen, J. (Ed.). (1997). *Educating minds and hearts: Social emotional learning and the passage into adolescence.* New York: Teachers College Press.

Cohen, J. (Ed.). (2001). *Caring classrooms/intelligent schools: The social emotional education of young children.* New York: Teachers College Press.

Darwin, C. (1965). *Expression of the emotions in man and animals.* New York: Philosophical Library. (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.

Dyer, M. G. (1983). The role of affect in narratives. *Cognitive Science, 7,* 211–242.

Ekman, P. (1973). *Darwin and facial expression: A century of research in review.* New York: Academic Press.

Ekman, P. (1985). *Telling lies.* New York: Norton.

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., Kessler, R., Schwab-Stone, M. E., & Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators.* Alexandria, VA: Association for Supervision and Curriculum Development.

Formica, S. (1998). *Description of the socio-emotional life space: Life qualities and activities related to emotional intelligence.* Unpublished senior honors thesis, University of New Hampshire, Durham.

Frijda, N. H. (2000). The psychologist's point of view: In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 59–74). New York: Guilford Press.

Guilford, J. P. (1959). *Personality.* New York: McGraw-Hill.

Lazarus, R. (1991). *Emotions and adaptation.* New York: Oxford University Press.

Legree, P. I. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence, 21,* 247–266.

Matarazzo, J. D. (1972). *Wechsler's measurement and appraisal of adult intelligence* (5th ed.). New York: Oxford University Press.

Mayer, J. D. (2001). A field guide to emotional intelligence. In J. Ciarrochi, J. P. Forgas, & J. D. Mayer (Eds.), *Emotional intelligence and everyday life.* (pp. 3–24). New York: Psychology Press.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27,* 267–298.

Mayer, J. D., Caruso, D. R., & Salovey, P. (2000). Selecting a measure of emotional intelligence: The case for ability scales. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence.* (pp. 320–342). New York: Jossey-Bass.

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., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence, 17,* 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* (Vol. 5, pp. 43–75). Greenwich, CT: JAI Press.

Mayer, J. D., Perkins, D., Caruso, D. R., & Salovey, P. (2001). Emotional intelligence and giftedness. *Roeper Review, 23,* 131–137.

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., Salovey, P., & Caruso, D. R. (2000a). Emotional intelligence as zeitgeist, as personality, and as a standard intelligence. In R. Bar-On & J. D. A. Parker (Eds.), *Handbook of emotional intelligence* (pp. 92–117). New York: Jossey-Bass.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000b). Models of emotional intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 396–420). Cambridge, England: Cambridge University Press.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2001). *Emotional intelligence meets traditional standards for an intelligence again: Findings from the MSCEIT.* Manuscript submitted for publication.

Nunnally, J. C. (1978). *Psychometric theory.* New York: McGraw-Hill.

Psychological Corporation. (1997). *WAIS-III, WMS-III Technical Manual.* San Antonio, TX: Author.

Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1,* 196–231.

Rubin, M. M. (1999). *Emotional intelligence and its role in mitigating aggression: A correlational study of the relationship between emotional intelligence and aggression in urban adolescents.* Unpublished doctoral dissertation, Immaculata College, Immaculata, Pennsylvania.

Salovey, P., Bedell, B., Mayer, J. D., & Detweiler, J. B. (2000). Current directions in emotional intelligence research. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (2nd ed., pp. 504–520). New York: Guilford Press.

Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185–211.

Salovey, P., Mayer, J. D., & Caruso, D. R. (in press). The positive psychology of emotional intelligence. In S. J. Lopez & C. R. Snyder (Eds.), *The handbook of positive psychology.* New York: Oxford University Press.

Salovey, P., Mayer, J. D., Caruso, D. R., & Lopez, P. N. (in press). Measuring emotional intelligence as a set of abilities with the MSCEIT. In S. J. Lopez & C. R. Snyder (Eds.), *Handbook of positive psychology assessment.* Washington, DC: American Psychological Association.

Salovey, P., Woolery, A., & Mayer, J. D. (2001). Emotional intelligence: Conceptualization and measurement. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell handbook of social psychology: Interpersonal processes* (pp. 279–307). Malden, MA: Blackwell.

Solomon, R. C. (2000). The philosophy of emotions. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 3–15). New York: Guilford Press.

Sternberg, R. J. (1997). The concept of intelligence and its role in lifelong learning and success. *American Psychologist, 52,* 1030–1045.

Trinidad, D. R., & Johnson, C. A. (in press). The association between emotional intelligence and early adolescent tobacco and alcohol use. *Personality and Individual Differences.*

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale.* New York: Psychological Corporation.

Received June 25, 2001
Revision received July 2, 2001
Accepted July 6, 2001 ■

New Editors Appointed, 2003–2008

The Publications and Communications Board of the American Psychological Association announces the appointment of five new editors for 6-year terms beginning in 2003.

As of January 1, 2002, manuscripts should be directed as follows:

- For the *Journal of Applied Psychology,* submit manuscripts to **Sheldon Zedeck, PhD,** Department of Psychology, University of California, Berkeley, CA 94720-1650.

- For the *Journal of Educational Psychology,* submit manuscripts to **Karen R. Harris, EdD,** Department of Special Education, Benjamin Building, University of Maryland, College Park, MD 20742.

- For the *Journal of Consulting and Clinical Psychology,* submit manuscripts to **Lizette Peterson, PhD,** Department of Psychological Sciences, 210 McAlester Hall, University of Missouri—Columbia, Columbia, MO 65211.

- For the *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes,* submit manuscripts to **John F. Dovidio, PhD,** Department of Psychology, Colgate University, Hamilton, NY 13346.

- For *Psychological Bulletin,* submit manuscripts to **Harris M. Cooper, PhD,** Department of Psychological Sciences, 210 McAlester Hall, University of Missouri—Columbia, Columbia, MO 65211.

Manuscript submission patterns make the precise date of completion of the 2002 volumes uncertain. Current editors, Kevin R. Murphy, PhD, Michael Pressley, PhD, Philip C. Kendall, PhD, Chester A. Insko, PhD, and Nancy Eisenberg, PhD, respectively, will receive and consider manuscripts through December 31, 2001. Should 2002 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2003 volumes.