

New Directions and Alternative Approaches to the Measurement of Emotional Intelligence

Benjamin Orchard, Carolyn MacCann, Ralf Schulze, Gerald Matthews, Moshe Zeidner, and Richard D. Roberts

Imagine a test developer struggling with the complex task of creating an assessment tool to measure an exciting new construct desired by educational, industrial/organizational, and clinical psychologists alike. The test developer first turns to the peer-review literature for a definition to act as a basis for this test, but finds instead heated debate and much disagreement. Researchers are using many different approaches to measure this new construct. In addition, the scores associated with these different approaches seem to measure different entities. The dissension about which approach is the most useful or valid leaves the test developer confused over which approach might measure the “real” construct. The measurement of emotional intelligence (EI) is currently at just such a somewhat precarious stage.

Our aim in the current chapter is to evaluate the existing approaches for measuring EI and suggest some new ones, perhaps providing a clearer path for our apocryphal test developer. Currently, there are two basic varieties of measurement tools used to assess emotional intelligence: (1) the typical performance, self-report, scales; and (2) the maximum performance, ability, scales. Despite the same label being attached to these two types of scales, corresponding scores are only weakly correlated, suggesting that they may in fact assess two different constructs (Brackett, Rivers, Shiffman, Lerner, & Salovey, 2006; Goldenberg, Matheson, & Mantler, 2006; Zeidner, Shani-Zinovich, Matthews, & Roberts, 2005). Some well-known examples of self-report scales include the Assessing Emotions Scale (AES; Schutte et al., 1998) and the Emotional Quotient Inventory (EQ-*i*; Bar-On, 1997). Indeed, these exemplars are but a small subset of self-report instruments available for research and for operational use. The large volume of assessments using this methodology stands in stark contrast to maximum-performance scales, where there is only *one* such EI assessment currently available for operational use: the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, & Caruso, 2002), and



B. Orchard (✉)
Educational Testing Service, Princeton, NJ, USA
e-mail: borchard@ets.org

an earlier research-only instrument, the Multi-Factor Emotional Intelligence Scale (MEIS; see Mayer, Caruso, & Salovey, 1999).

In this chapter, we first outline the logic for evaluating validity evidence for EI tests, arguing that the self-report scales should not be labeled as if they could be used to validly assess emotional intelligence. We also evaluate the MSCEIT subtests using the same criteria, pointing out some strengths and weaknesses of this instrument. Finally, we suggest that alternative test development procedures (many from older traditions within emotions or social intelligence fields) may lead to promising new directions for the emerging sub-discipline of emotional intelligence.



Evaluating the Validity Evidence for an Emotional Intelligence Test

Given the label “emotional *intelligence*”, a non-expert entering the field might reasonably assume that EI relates to the processing and knowledge of emotional information. That is, EI is one of multiple factors of intelligence whose differentiation from existing intelligence constructs lies in the nature of the information processed: emotions. It is assumed that emotional information from an external source, that is, observing another person’s emotions, is also part of this domain, as well as one’s own observed and resultant emotions in a given situation.

The label *emotional intelligence* thus implies a number of criteria for evaluating the validity of an EI test (see Mayer et al., 1999; Roberts, Zeidner, & Matthews, 2007). **Firstly**, EI test scores should show substantial positive correlations to other tests of intelligence, demonstrating the positive manifold (that is, consistent positive correlations) found among intelligence tests. **Secondly**, EI assessments should not relate too strongly to any *one* of the specific types of intelligence. That is, correlations between EI and existing intelligence constructs should indicate that EI measures a type of intelligence, but not *exactly the same type of intelligence* that other tests already assess. **Thirdly**, EI scores should relate to indicators of emotional functioning (e.g., coping with stress, peer-evaluations of empathy), indicating that this *emotional* intelligence operates in the domain of the emotions. **Fourthly**, as EI is an intelligence rather than a personality trait, EI test scores should relate to personality in the same way that other tests of intelligence do (i.e., $r = .30$ or less; cf. Ackerman & Heggestad, 1997; Ashton, Lee, Vernon, & Jang, 2000). These four criteria refer to important evidence for the construct validity of EI measures (see also Mayer, Roberts, & Barsade, 2008).¹



Typical (Self-Report) Measures of Emotional Intelligence

It is quite clear from published research that self-report measures of EI do not meet Criterion 1 (substantial positive correlations with intelligence). It is also clear that they do not meet Criterion 4 (correlations with personality similar to

those of established intelligence measures). Indeed, as we shall demonstrate shortly, it seems that self-report tools assess aspects of personality rather than intelligence (see also, e.g., Mayer et al., 2008; O'Sullivan, 2007; Roberts et al., 2007; Scherer, 2007).

Correlations with Intelligence

Self-report measures of EI correlate only trivially with published measures of cognitive ability. For example, the EQ-*i* correlates $r = .08$ with both the General Adult Mental Ability Scale and the Wonderlic (Derksen, Kramer, & Katzko, 2002; Newsome, Day, & Catano, 2000). This trend is not isolated to the EQ-*i* but holds across all self-report measures of EI (see Barchard & Hakstian, 2004; Roberts, Schulze, & MacCann, 2008). This relationship is sufficiently low that even a substantial correction for attenuation would leave these findings well below desirable levels.

Correlations with Personality

Pérez, Petrides, and Furnham (2005; see also Tett, Fox, & Wang, 2005), evaluated a total of 21 self-report measures of EI, and found extremely high correlations (up to .87) with personality traits. The magnitude of these correlations indicates that self-report measures assess known aspects of personality. This finding clearly violates the fourth requirement: These measures should only be related to personality indicators to a similar degree as other measures of cognitive ability, which is not the case.

It is clear from this brief discussion of the main issues concerning self-report measures of EI that these instruments do not assess a type of intelligence, and thus have questionable construct validity. As measures of emotionality and emotional tendencies (rather than EI), the rating-scale methodology does have some obvious benefits for the test-developer or practitioner. After all, rating scales are: (a) reasonably easy to develop, (b) quick to administer, and (c) easily scored. However, labeling the resulting score as EI is misinformative and confusing, conflating an already complicated issue (see also Mayer et al., 2008). Roberts et al. (2007) suggest that research employing rating scales might still prove valuable, though it would benefit the scientific enterprise if it were couched as an attempt to more fully map the sub-factors of personality rather than anything to do with the intelligence domain.

Maximum Performance Measures: The MSCEIT

Assessments of EI should be analogous to cognitive ability tests, but with the content concerning emotions rather than acculturated knowledge (as in crystallized intelligence), abstract stimuli (as in fluid intelligence), visual patterns (as in

visualization abilities), or sounds (as in broad auditory intelligence). In the paragraphs that follow, we describe the MSCEIT (and MEIS) and evaluate the validity evidence against the four requirements introduced above.

The Four-Branch Model

The MEIS/MSCEIT conform to the hierarchical four-branch model developed by Mayer and Salovey (1997; Mayer et al., 2002), which are: (1) Emotion Perception, (2) Assimilating Emotions (into thoughts, plans, and actions), (3) Emotional Understanding, and (4) Emotion Management (of one's self and of others). Other researchers have expanded on this model with concepts and paradigms from the emotions, clinical psychology, or cognitive processing fields. For instance, Scherer (2007) argues that appraisal and communication are also potential bases for EI. Others have argued that the accurate perception of emotions or the ability to articulate emotions may also be related to EI (Davies, Stankov, & Roberts, 1998; Lane, 2000; Roberts et al., 2007). These views, however, are not inconsistent with the four-branch model; these components appear to share close conceptual parallels with the identifying and understanding branches.

Description of the MSCEIT

There are eight subtests in the MSCEIT, two for each of the four branches. The two Emotional Understanding tests employ a multiple-choice format (i.e., test-takers select the best of four options). In all of the six remaining tests, test-takers assign a rating to each option, rather than selecting the best option. For example, test-takers view a facial expression and then rate the presence of sadness, anger, happiness, and fear in a typical item from the Emotion Perception branch. In a typical Emotion Management item, test-takers might read about an emotional situation, and then rate the effectiveness of several strategies for managing such a situation. In order to avoid confusing this rating-based format with the self-report ratings, we refer to this type of item as a *rate-the-extent* scale. The rate-the-extent format is somewhat unusual for intelligence measurement procedures and may have implications for some research findings on the structure and correlates of the MSCEIT. The scoring of the MSCEIT is also unusual for a cognitive ability test, with responses weighted according to the judgment of an expert panel or large normative sample. Even so, the various major sub-scales of the MSCEIT demonstrate acceptable reliability, with scores of .93 and .91 for consensus and expert scoring respectively (Rivers, Brackett, Salovey, & Mayer, 2007). The four individual dimensions, however, have somewhat lower reliabilities, with studies finding reliabilities as low as .76.



Relationships with Cognitive Assessments

MSCEIT scores correlate at about $r = .18$ with fluid intelligence and about $r = .31$ with crystallized intelligence (Roberts et al., 2008). However, the strength of relationship to intelligence varies by branch. The strongest relationship with intelligence is that found between Emotional Understanding and crystallized intelligence ($r = .40$). All other branches show only small relationships to intelligence ($r = .03-.18$). Although these figures are based on a relatively small number of studies available from the peer-review literature, two conclusions can be drawn with some confidence. Firstly, if we accept that there should be a moderate or strong relationship between established measures of intelligence and EI, then the MEIS and the MSCEIT Emotional Understanding subscales are faring better than self-report measures. Secondly, the other branches (Emotion Identification, Facilitation, and Management) are not as strongly related to conventional measures of intelligence as Criterion 1 seems to demand of the construct.

One reason that Emotion Identification, Facilitation, and Management tests might not be substantially correlated with intelligence tests is the rate-the-extent format of test items. MacCann and Roberts (2008) tested this idea with a new research tool that assesses emotion management (the Situational Test of Emotion Management; see below) that may be administered in either multiple-choice or rate-the-extent format. Multiple-choice scores were more strongly related to Vocabulary scores than rate-the-extent scores ($r = .40$ compared to $r = .26$). Although Vocabulary is but one subfactor of intelligence, this study suggests that response format has an influence on scores, and could explain the stronger Understanding-Intelligence link (Roberts et al., 2008). In any case, with the response formats currently in place in the MSCEIT, only the Emotional Understanding branch clearly meets Criterion 1.

Relationships with Personality Assessments

Relationships of the MSCEIT to the Big Five dimensions of personality also differ by branch. In the Roberts et al. (2008) summary of recent publications, the only personality dimension to relate moderately (i.e., $r > .20$) with EI was Agreeableness. Agreeableness correlated at $r = .22$ with MSCEIT total scores and $r = .27$ with Emotion Management scores, low enough to indicate that the MSCEIT is not part of the personality domain. The small relationship to Openness often found for crystallized intelligence measures was not observed for EI. More generally, the conclusion is that MSCEIT scores, unlike that found for self-reports of EI, do not represent a facet (or facets) of personality, though the replicable relationship with Agreeableness is of some (and perhaps even considerable) theoretical interest.

Relationships with External Criteria Indicative of High Emotional Functioning

In addition to relationships with cognitive ability and personality, it is necessary to examine the relationship of EI with emotion-related criteria. The specification of such criteria is not entirely obvious, at least conceptually. In many cases, criteria involving an emotional component clearly have multiple causes, making for error variance in the criteria, which makes interpreting low correlations between EI and outcomes difficult. For example, relationship satisfaction may relate to EI, but may also have a number of other possible causes: financial security, social support, physical health, and lifestyle factors. Statistically controlling for all these variables is difficult if not impossible, at least in practice. Nevertheless, validity evidence for the emotional component of EI is important, and we summarize evidence of various kinds below, along with interpretations and caveats associated with these findings.

(1) *Coping and stress response.* A significant relationship between MSCEIT subscale scores and coping styles has been found by Gohm, Corser, and Dalsky (2005). Understanding and Management were related to behavioral disengagement ($r = -.25$ and $r = -.21$) and denial ($r = -.30$ and $r = -.21$). Management was also related to seeking emotional support ($r = .25$), seeking instrumental social support ($r = .23$), and to drug and alcohol disengagement ($r = -.16$). However, these were the only relationships among 13 possible coping scales, such that the conclusive statement “the MSCEIT predicts coping” might be considered an overstatement. In addition, a study by Brackett et al. (2006) found that male test-takers had total scores on the MSCEIT that predicted constructive/destructive coping responses, even after controlling for Big Five personality factors, psychological well-being, empathy, life satisfaction, and Verbal SAT scores (partial correlations ranged from $-.22$ to $-.33$). It is unclear why this finding was only significant for males, although it may be due to restriction of range among the females on both the MSCEIT and the constructive/destructive responses. Consistent with this idea, females scored significantly higher than males on the MSCEIT and constructive coping strategies, but lower on destructive strategies.

In one of the rare experimental studies of EI, Matthews et al. (2006) showed that EI as measured by the MSCEIT is significantly related to (pre-task) distress and worry stress states, even when controlling for the Big Five factors. Consistent with theoretical expectations, EI was also weakly, but significantly, related to an avoidance coping style under different stressful experimental conditions in this study. However, EI was not related to other coping styles (i.e., emotion-focused and task-focused coping) and, most importantly, EI did not predict task-induced stress when controlling for the Big Five factors. Hence, the evidence provided by this study does not support the notion that EI is a critical predictor of stress responses, as might be theoretically expected.

(2) *Relationship quality.* Couples where both partners were low on EI report lower relationship depth, lower support, poorer positive relationship quality, higher conflict, and greater negative relationship quality compared to couples where at least one member has high EI (Brackett, Warner, & Bosco, 2005). However, neither intelligence nor personality factors were controlled for in this study.

(3) *Social interaction.* Management subscale scores predicted self-reports of positive interactions with friends ($r = .31$), as well as friends' reports of positive and negative interactions and emotional support ($r = .33$; $-.30$; and $.26$; Lopes et al., 2004). Additionally, male scores on the MSCEIT predicted confederate and judge ratings of social behaviors after a social interaction, although female scores did not (Brackett et al., 2006). Again, this outcome may be due to the relatively high scores of females on both the predictors and the criteria.

(4) *Tobacco and alcohol use.* Scores on both the MEIS and the MSCEIT are negatively correlated with tobacco and alcohol use (Brackett, Mayer, & Warner, 2004; Trinidad & Johnson, 2002; Trinidad, Unger, Chou, & Johnson, 2005). The argument that these negative relationships constitute a form of validity evidence is not straightforward, as drug use is not strictly an emotional decision, even though there are emotional components. EI may act as a mediating variable, or may itself be mediated by other important variables (e.g., impulse control). There are a number of logical connections between EI and drug use that still need to be explicitly elucidated or empirically tested.

(5) *State anxiety.* MSCEIT scores also seem to be related to state anxiety, as the additional variance accounted for over intelligence and personality factors was estimated at 6% (Bastian, Burns, & Nettelbeck, 2005).

This body of evidence indicates that ability-based EI (as represented by the MSCEIT) does seem to relate to some important emotion-related outcomes. Notably, these correlations are also different for different branches, with Emotional Understanding and Emotion Management providing the greatest predictive utility. Given that Emotional Understanding also shows the most convincing validity evidence of the four branches, and that factor analyses conducted so far have found little support for the Assimilating Emotions branch (Ciarrochi, Chan, & Caputi, 2000; Mayer, Salovey, Caruso, & Sitarenios, 2003; Palmer, Gignac, Manocha, & Stough, 2005; Roberts, Zeidner, & Matthews, 2001), it conceivably makes sense to reconceptualize the four-branch model. In a revised model, greater weight might be given to Emotional Understanding, while Assimilating Emotions might be eliminated as a component of the model.

New Approaches to the Assessment of Emotional Intelligence

For a burgeoning research field such as EI to be defined by a single instrument with a single and somewhat unusual measurement method is a suboptimal state of affairs. Fortunately, new methods to assess EI are under development and the

work from many research areas steadily diffuses into the EI field. The assessment approaches we discuss in the passages that follow: (a) have been developed but are not commonly used in assessing EI, (b) are still in development, or (c) consist of concepts and methods that might potentially be leveraged to provide valid EI assessments. The list of all these different approaches to assessing EI includes information processing measures (such as inspection time and implicit association tests), as well as perceptual tests such as emotion recognition tasks, situational judgment tests (SJTs), and measurement paradigms from the conditional reasoning, principal-agent paradigm, and affective forecasting fields.

Information Processing Measures

Implicit Association Tests

The use of Implicit Association Tests (IATs) in detecting subtle biases is well documented – at the time of writing the current chapter, Greenwald, McGhee, and Schwartz’s (1998) article introducing the IAT method has been cited more than 800 times in the literature. In essence, the IAT paradigm assesses the relative strength of positive and negative associations test-takers have for two opposing ideas (e.g., the strength of positive and negative associations for obese people versus slim people might indicate an implicit evaluative bias in favor of slim people). Mauss, Evers, Wilhelm, and Gross (2006) demonstrate that the IAT may be used to assess the implicit biases towards emotional expression and disclosure (e.g., the words “emotional,” “disclose”) as compared to emotional regulation and control (e.g., the words “controlled,” “suppress”). Based on the premise that the people who regulate their emotions have implicit positive evaluations of emotion regulation, Mauss et al. (2006) suggest that their IAT (the Emotion Regulation–Implicit Association Test [ER–IAT]) assesses emotion regulation.

The ER–IAT demonstrated reasonable test-retest reliability ($r = .68$, $p < .001$, $N = 36$), but was unrelated to trait measures of emotion regulation. Clearly, both the small sample size and failure to provide convergent validity evidence render this study problematic. Indeed, the IAT measurement approach is not without its critics, as there is some disagreement about what the task actually measures (e.g., Fiedler & Bluemke, 2005; Mierke & Klauer, 2003). Nevertheless, it is one of the few available approaches that may have some potential to provide objective measures of emotion management. In addition, the approach can easily be adapted to assess other related constructs (e.g., empathy, emotion perception).

Emotion Recognition Assessments

The study of emotions in facial expressions is a far older research field than EI. Scientific research on facial expressions arguably stretches back more than a century, to Charles Darwin’s *The Expression of the Emotions in Man and Animals* (1872). In psychology, Paul Ekman’s seminal contributions to this

field date back to the 1950s. Ekman's corpus of work includes the specification of lawful rules linking muscular changes to facial expressions in the form of the Facial Action Coding System (FACS; Ekman & Friesen, 1978; Ekman & Rosenberg, 1997). The FACS was later used as the basis for several research tools assessing the recognition of facial expressions: (1) the Pictures of Facial Affect (POFA; Ekman & Friesen, 1976); (2) the Japanese and Caucasian Brief Affect Recognition Test (JACBART; Matsumoto et al., 2000); (3) the Micro-Expression Training Tool (METT; Ekman, 2004), and (4) the Subtle Expression Training Tool (SETT; Ekman, 2004). The vast research corpus in this domain could clearly be valuably applied to EI research. However, only in the last few years have some EI researchers begun to use these measures (e.g., Austin, 2005; Roberts et al., 2006).

In addition to these face-based assessments, there are currently assessments of emotion recognition ability for tone-of-voice such as the Vocal-I (e.g., Scherer, Banse, & Wallbott, 2001; Scherer, 2007) and the prosody measures from the Diagnostic Analysis of Nonverbal Accuracy (e.g., DANVA2-AP; Baum & Nowicki, 1998). The DANVA also includes subtests that assess the recognition of emotions from postures or gestures.

These assessments appear to correlate highly with each other when in the same modality, providing some evidence of convergent validity. For example, the DANVA2-AF and the JACBART correlate at $r = .80$ (Danforth, McIntire, & Schneider, 1997). However, Roberts et al. (2006) found that measures from different modalities (i.e., vocal and facial tests) were only weakly correlated. Specifically, the JACBART and Vocal-I correlated at $r = .17$. In addition, these measures do not seem to correlate with the Perceiving Emotions tests from the MSCEIT. Roberts et al. (2006) report a correlation of $r = .07$ between MSCEIT Faces and the JACBART. These findings suggest two testable hypotheses. Firstly, the ability to perceive emotions correctly may not be a single unitary ability, but may be specific to different modalities. Secondly, the MSCEIT method of rating the extent of emotion in a still photo of a face seems to capture a different skill from multiple-choice assessment of facial expression presented for only a fraction of a second (as in the JACBART). Whether the difference is due to the time limits for observation of the stimuli or due to the difference between multiple-choice versus rate-the-extent formats is not entirely clear at this juncture.

Emotional Intelligence Related to Inspection Time

Austin (e.g., 2005) has conducted studies relating self-reported, trait EI measures to an Inspection Time (IT) task, where respondents were given a speeded test and asked to discriminate between happy and neutral faces, sad and neutral faces, and two neutral symbols. IT is considered the minimal response time necessary to distinguish between two stimuli, and may be administered in a variety of different ways. Findings using this approach however, are rather mixed. For example, IT is weakly related to scores on a self-report EI measure ($r < .12$ for happy IT and sad IT), and an overall index of EI combining the various stimuli was found to be

weakly related to a general EI score from the Bar-On EQ-*i* ($r < .17$) (e.g., Austin, 2005; Austin & Saklofske, 2005; Stokes & Bors, 2001). What remains to be seen is if this IT task is substantially correlated with maximum performance measures of EI such as the MSCEIT. In going beyond the simple ability to distinguish between differing emotions, and considering how quickly an individual can accomplish this task, IT holds promise. The potential utility for such an application is rich, although it is necessary to tie this to real-world outcomes, as well as performance on maximum performance measures of EI.

Social Intelligence Measures as Emotion Perception Analogs²

The face, voice, and body language-recognition tasks discussed in the previous section represent the most direct analogs to emotion perception as conceptualized in the four-branch model (as assessed by the MSCEIT). However, there are several other assessments of nonverbal ability that appear conceptually related to the concept of EI and emotion perception. These include the Communication of Affect Receiving Ability Test (CARAT; e.g., Buck, 1984), the Profile of Nonverbal Sensitivity (PONS; Rosenthal, Archer, Hall, DiMatteo, & Rogers, 1979); the Interpersonal Perception Task (IPT; Costanzo & Archer, 1993); and the Child and Adolescent Social Perception Measure (CASP; Magill-Evans, Koning, Cameron-Sadava, & Manyk, 1995). Test-takers are presented with brief displays of people and need to discern such information as what stimuli the people might be looking at, what the relationship between people is (e.g., who is the boss, or the parent of a child), or what emotions the people are experiencing. The ability to correctly recognize emotions may form part of the social, interpersonal, or other kinds of nonverbal recognition that these tests assess.

In contrast to recognition measures, there are some existing assessments that attempt to tap explicit knowledge of social and/or emotional nonverbal phenomena. These include the Test of Nonverbal Cue Knowledge (TONCK; Rosip & Hall, 2004); and measures that assess test-takers' knowledge of deception cues, cues as to an individual's intelligence, and gender differences in social and expressive behaviors (Hall & Carter, 1999; Murphy, Hall, & Colvin, 2003; Vrij & Semin, 1996). Differences between explicit and implicit knowledge may be particularly important in EI, where facts and skills are rarely explicitly taught, but instead generally learned in a tacit or implicit way.

Situational Judgment Tests

Situational judgment tests (SJTs) are a type of test where individuals are presented with a situation (vignette) and then select either the most appropriate response or their typical response out of a list of possible choices. The method for constructing such a test is simple enough: (a) Items are generated (usually by critical incident interviews with subject-matter-experts); (b) responses are

generated (usually by a second group of subject-matter experts); and then (c) a scoring key is developed by a further group of experts (McDaniel & Nguyen, 2001). Several variations on this classic three-step procedure exist, such that virtually any test with situations as the item stimuli is commonly described as an SJT. Earlier situational measures of EI generally did not follow the three-step methodology. It is only in the last five years that situation-based tests of social or emotional intelligence have been informed by the SJT literature.

Early Precursors to SJT Measures of EI

Existing measures of EI resembling SJTs include the Levels of Emotional Awareness Scale (LEAS; Lane, Quinlan, Schwartz, Walker, & Zeitlin, 1990), the Emotional Accuracy Research Scale (EARS; Mayer & Geher, 1996), the MEIS Stories test (Mayer et al., 1999), and the MEIS and MSCEIT management tests. If we consider social intelligence as part of the EI domain, then Chapin's (1942) Social Insight Test might also be considered as a (very) early situational measure of EI. Each of these tests presents a situation as the basis for an item, but there is no indication that the classic three-step approach to developing SJTs was taken. For the MEIS and LEAS, the test authors generated the situations and scoring key. For the MSCEIT Management tests, an expert panel was used for scoring. For the EARS, items and responses were generated from abridged transcriptions of interview data, and scored according to the interviewees' endorsements (i.e., target scoring). More detail and example items from the LEAS, EARS, MEIS Stories test, and Social Insight test are given below:

1. *Levels of Emotional Awareness Scale* (Lane et al., 1990). Test-takers give open-ended responses to situations such as the following: "*You and your best friend are in the same line of work. There is a prize given annually to the best performance of the year. The two of you work hard to win the prize. One night the winner is announced: your friend. How would you feel? How would your friend feel?*"

Responses are scored from 0 to 5 in terms of the sophistication of the emotional detail involved, in line with a cognitive-developmental model of emotional experience. Variation in LEAS scores may be due to verbal ability as well as emotional awareness, as it requires sophisticated verbal expression of this emotional awareness.

2. *Emotional Accuracy Research Scale* (Mayer & Geher, 1996). Each item consists of several situations derived from the same individual (these situations were generated by interviewing the individual about their emotional experiences). For example: "*My best friend's father died this weekend. He had diabetes for a long time and as he got older his health grew worse and worse. I went to his funeral on Monday. Many of my friends from high school were also there because we all wanted to be there for our friend and because we all knew and liked her father. It made me realize how lucky I am to have younger healthy parents when I saw my friend standing there crying. Just watching her huge family come pouring into the synagogue also made me sad.*"

Test-takers are required to select the most appropriate word from twelve forced-choice pairs (e.g., mad – delighted; stomping feet – happy for another), where the correct alternative was whatever the target endorsed.

3. *MEIS Stories Test* (Mayer et al., 1999). Each item consists of a story of 2–3 sentences describing events happening to a fictitious person. For each story, test-takers rate the intensity of seven different emotions the protagonist experienced. Ratings were given on a five-point scale from “definitely present” to “definitely NOT present.” For example: *“I was at home, and I got this upsetting phone call from a woman I work with. She is turning into a very aggressive and extremely demanding persona and was downright rude on the phone. She made unreasonable demands and was very angry. I think she’s jealous and upset that she isn’t involved in this successful project that I’m part of. Then, I had dinner with a friend who is really together and focused. I am questioning the work I am doing and wishing that I could be more even and focused like my friend.”*

4. *Social Insight Test* (Chapin, 1942). Each item describes a scenario consisting of social information about what the protagonist does. The test-taker must select the most plausible explanation for the protagonists’ behavior from four possible alternatives (where the best alternative is determined by expert judgment). For example: *“Mr. Asher, when told that an acquaintance had purchased a new automobile, was heard to criticize him very strongly for spending so much money for a car when he probably could not afford one. Not long after this incident, Mr. Asher himself bought an expensive new automobile. About the same time he placed another mortgage on his house. Why did Mr. Asher criticize his acquaintance for an act he afterwards performed himself? (a) Because he probably had ‘money left to him’ upon the death of a near relative. (b) Criticism of his acquaintance got rid of an ‘uneasy feeling’ about something he contemplated doing himself.”*

These early measures point to the application of *emotional situations* as useful stimuli when assessing EI. However, none of these measures (with the possible exception of the EARS) applied an SJT approach to test development. Neither the situations nor the possible responses were based on real emotional experiences of an appropriate sample (as is the case in the classic SJT approach, where situations and responses are generated via interview data). In the last five years, a number of research teams have developed similar measures of emotion management and related abilities, but with a methodological basis that more clearly follows the SJT method (thus allowing greater ecological validity – that is, the test situation more closely mimics reality).

SJT Measures Directly Targeted to Assess Social and Emotional Intelligence

There are several emerging research groups, across the globe, that are developing social and emotional intelligence tests after the SJT tradition. A selection of such efforts is described in what follows.

1. *Social Intelligence Test – Magdeburg (SIM)*. Based on a newly developed and empirically tested facet model of social intelligence (Weis & Süß, 2005, 2007),

a multiple-construct, multimodal measure of social intelligence has been developed by a research team in Magdeburg, Germany: the Social Intelligence Test – Magdeburg (SIM; see Seidel, 2007; Süß, Seidel, & Weis, 2008). In this assessment battery, social understanding, social memory, and social perception are assessed using textual stimuli (e.g., emails, letters, diary entries), audio stimuli (e.g., recordings of phone-calls), photos (e.g., people showing gestures relevant in conversations), and video-based stimuli (e.g., people conversing in the kitchen as they cook a meal). The stimuli were selected on the basis of a taxonomy of social situations (see Süß et al., 2008) and culled from actual footage of volunteers (the protagonists) who were recorded for long periods of time. After viewing the stimuli that experts (academic psychologists) selected as test items, the protagonists answered detailed questions about the social and emotional content of those materials, which were used as the basis for scoring. That is, the SIM is an ecologically valid assessment of social intelligence based on target scoring. Structural analyses provide support for the underlying theoretical facet structure and social intelligence as a non-redundant construct of intelligence.

2. *Situational Test of Emotional Understanding (STEM)*. MacCann and Roberts (2008) developed SJTs assessing components of EI using accepted methodology underlying the SJT approach. The Situational Test of Emotion Management (STEM) was developed using interview data to specify situations, written responses from a second sample to specify response options, and expert panels to specify scoring keys. The STEM assesses fear-, anger- and sadness-management and can be administered in two formats: (1) multiple-choice (where test-takers select the best response to the situation); and (2) rate-the-extent (where test-takers rate the effectiveness of each response option). MacCann and Roberts (2008) showed that multiple-choice and rate-the-extent tests show different patterns of correlations: Multiple-choice scores correlate more highly with vocabulary and less highly with personality dimensions, and rate-the-extent scores correlate more highly with personality and less highly with vocabulary. An example item from the STEM (MacCann & Roberts, 2008) follows: “Clayton has been overseas for a long time and returns to visit his family. So much has changed that Clayton feels left out. What action would be the most effective for Clayton? Option 1: Nothing, it will sort itself out soon enough. Option 2: Tell his family he feels left out. Option 3: Spend time listening and getting involved again. Option 4: Reflect that relationships can change over time.

3. *Intrapersonal and Interpersonal Abilities*. Freudenthaler and Neubauer (2005) developed two multiple-choice measures of emotion management: (1) self-management (intrapersonal abilities) and (2) the management of others (interpersonal abilities). Responses were generated by distilling open-ended responses to various scenarios, and scored according to a panel of experts (psychologists and psychotherapists). Freudenthaler and Neubauer (2007) compared the correlates of their measures under “would do” (i.e., what the test-taker would do in that situation) versus “should do” (what the test-taker thought the best response was) instructions. Compared to the “should do” instruction, the “would do” condition showed stronger correlations with

personality, and weaker correlations with intelligence. An example item from the Interpersonal Abilities Test (Freudenthaler & Neubauer, 2005) follows: *“Your father is very scared of a complicated operation which is absolutely necessary. Option (1 of 4): To calm him down I advise him to talk to the doctor once more.”*

Research on these three SJT-based measures of EI shows the importance of examining the effects of different response formats, instructions, and test modalities. It is clear from this research that such characteristics of test items change the constructs assessed. This multi-method focus is a feature of emerging research using SJTs and other methodologies. While the systematization of test development practices afforded by the SJT approach appears useful, SJTs are not free from scoring problems. It is still difficult to be certain that any SJT is going to have a response set that can be unambiguously scored by experts: There are simply not any sufficiently developed models to allow for this. Despite this limitation, SJTs have seen widespread use in industrial/organizational psychology for purposes of selection and promotion.

Future Directions for SJT Measures of EI

There are some clear concerns for text-based SJTs: Both the ecological validity and dependence on text comprehension are less than ideal. Weis and Süß's (2007) work with different modalities in social intelligence highlights this point to some extent – text-based items were more difficult than video- or audio-based items, and correlated more strongly with cognitive abilities. What is still needed is a rationally-developed video-based SJT assessing EI. Although Weis and Süß's tests can be considered to be SJTs, the culling of incidents from real (social) life meant that less control could be exercised over the content dimensions of the test. Furthermore, the content emphasized in their tests was social relations; where emotions may be important but do not necessarily function as the main aspect to determine correct responses. The development of a video-based SJT assessing emotion management, using MacCann and Roberts' (2008) STEM as a starting point, and the classic three-step SJT approach as a methodology, is one of the ongoing projects of the current group of collaborators. Using scripts developed on the basis of interview data, and professional actors and film crew to act out the scenarios allows content specification and control as well as ecological validity and independence from reading ability.

With regard to concerns about expert scoring, MacCann and Roberts (2008) also developed a test of emotional understanding (the Situational Test of Emotional Understanding, or STEU) that used information from Roseman's (1984) appraisal theory as the basis for scoring test items according to explicit and verbalizable standards rather than expert opinions. Although reliability of this test was low in a non-university sample, correlations with other expert-scored EI tests indicate that such a method has potential to be included among these new assessment approaches.



Assessing Emotional Intelligence Through Implicit Beliefs: The Conditional Reasoning Paradigm

James and collaborators have assessed aggression, achievement motivation, and other traits using the “Conditional Reasoning” paradigm (James, 1998; James, McIntyre, Glisson, Bowler, & Mitchell, 2004; James et al., 2005). The basis behind conditional reasoning is that test-takers who hold a particular implicit belief about the world (e.g., “Everyone is out to get me”) will interpret an ambiguous situation differently from people who do not hold this belief (e.g., “A person who bumps into me is being deliberately cruel, rather than accidentally clumsy”). Items consisting of such ambiguous situations might then discriminate between people who do and who do not hold such sets of implicit beliefs (known as justification mechanisms). Research so far has indicated that this is a valid way to assess aggression and achievement motivation. If people with low EI also hold a particular set of implicit beliefs about the world, then this measurement paradigm might usefully be extended to EI.

Interestingly, conditional reasoning bears some resemblance to both IAT measures (because of the targeting of implicit beliefs and ideas) and the SJT methodology (because of its reliance on vignettes and response options). Unlike a SJT, however, the conditional reasoning tasks rely on having two logically plausible response options, only one of which would be chosen by people holding the implicit belief of interest. In the example below (James et al., 2005), options B and D are both logical, but the logic of option D requires that the test-taker believes in the “powerful others” bias (that people will always inflict harm on those less powerful than themselves): “*The old saying, ‘an eye for an eye,’ means that if someone hurts you, then you should hurt that person back. If you are hit, then you should hit back. If someone burns your house, then you should burn that person’s house. Which of the following is the biggest problem with the ‘eye for an eye’ plan? Option A: It tells people to ‘turn the other cheek’ (illogical). Option B: It offers no way to settle a conflict in a friendly manner (logical, no aggression). Option C: It can be used only at certain times of the year (illogical). Option D: People have to wait until they are attacked before they can strike (logical, aggression).*”

Transferring this methodology to EI assessment would require the identification of implicit biases relating to high or low levels of EI. For example, individuals who are poor at emotional expression may believe that the expression of negative emotions is shameful, is a burden to other people, is a sign of weakness, or is a characteristic of children rather than adults. Test items catering to these biases might be developed according to the Conditional Reasoning paradigm. Developing tests using this method would require some conceptual and theoretical work at the outset, which might prove useful for advancing the EI field quite apart from test development per se.

An Emotion-Based Adaptation of the Principal-Agent Paradigm

The principal-agent paradigm (PAP) assesses the ability of an agent to evaluate the preferences of a principal, based on known values that the principal holds for different features of the decision event. To take a trivially simple example, if the agent knows that the principal likes ice-cream a lot, but does not like cheesecake much, the agent will be able to predict the principal's dessert choice and place an order for the principal at a restaurant (i.e., act on his or her behalf). The PAP originated in the economics literature, to assess how well an agent can learn how much value the principal attaches to different attributes of a set of objects. The agent observes several examples of the principal's choices between exemplars of the set of objects in complex situations, and then must decide what the principal's preferences would be in a new complex situation (Huber, Ariely, & Fischer, 2002).

In transferring this methodology to psychological test development, we can consider the "agent" to be the test-taker, and the "principal" to be the protagonist in a vignette test item. An example of a complex situation might be selecting which type of island cruise vacation to go on. For simplicity's sake, let us assume that there are five relevant attributes of the cruise, each with three levels of quality (poor, fair, and good). These aspects might be: (1) *Cost* (\$3000 = poor, \$2000 = fair, \$1000 = good); (2) *Accommodation* (poor, fair, good); (3) *Likelihood of major storm* (90% = poor, 60% = fair, 30% = good); (4) *Shipboard entertainment quality* (poor, fair, good); and (5) *Cruise length* (3 days = poor, 6 days = fair, 9 days = good). A principal who values saving money more than any of the other attributes might choose the \$1000 package even if the other four attributes were poor. In contrast, a principal for whom money is less important in decision making might select the most expensive cruise, even if the other four attributes were only marginally better for the \$3000 versus the cheaper cruises. That is, the *values of the principal* would affect the *relative weight of each attribute* in making a choice. This is the key to the PAP: The agent must account for the principal's value structure in predicting the choices the principal would make.

Applying the PAP to EI research may provide an assessment of understanding emotions. Just like the agent in economic applications needs a good understanding of the principal's value or attitudinal structure to make appropriate choices on his or her behalf, the agent in an EI setting would require a high level of emotional understanding (or empathy) to comprehend the principal's emotional functioning. More specifically, in order to correctly predict what the principal would feel and do in a certain situation (including emotion-based decision making), the agent would need to learn or understand the principals' event-emotion, emotion-emotion, and emotion-behavior contingencies. That is, the agent needs to acquire knowledge and make valid inferences about: (a) emotions that typically follow certain events or event classes, (b) a potential chain of emotional reactions,

and (c) the action that most likely will result from the existence of a certain emotional state. In this context, the emotional reactions to events, the sequences of emotions, and the subsequent behavioral consequences may be either conscious and deliberate or automatic reactions of the principal. The process components and their interactions compose what is called the principal's emotional functioning or emotional structure.

Depending on the circumstances under which emotional knowledge acquisition takes place, the ability to perceive emotions may be an influential factor for assessing understanding emotions with the PAP. In addition, understanding the emotions of a particular principal may be enhanced if general knowledge about emotion contingencies is available and accessible (e.g., knowledge consistent with appraisal theories), as well as applicable to the emotional structure of a particular principal. The potential influences of emotion perception and emotion knowledge on EI measurement procedures are not unique for PAP measurement procedures but apply to most other measurement approaches of understanding emotions as well (e.g., the SJT approach).

When using the PAP in EI research, the decision making and prediction tasks would need to be framed appropriately with different emotions instead of values and utilities as the content structure to be acquired and used by the agent. For example, instead of a decision on which vacation package to choose, an emotion-based PAP (or EPAP) might ask which potential partner would be dated, which approach to an unjust accusation from a co-worker would be taken, or which method of expressing condolences to a friend would be made, based on the principal's emotional "preferences" and automatic emotional reactions. Examples for such preferences or reactions might include the need for confrontation, security, adventure, affection, or intellectual stimulation; or the tendency towards joy, sadness, frustration, jealousy, or anger.

In an implementation of an EPAP, the agent/test-taker would be presented with a profile of the principal's typical emotional reactions. This presentation can take many forms, from observation of the principal's behavior to explicit descriptions (e.g., verbal or graphical) and definitions of the contingencies to be learned. The behavior that is described can: (a) be overt or covert, (b) include event-emotion, emotion-emotion, as well as emotion-behavior contingencies, and (c) be video-based, computer-animated, text-based, etc. The more explicit this step, the less influence can be exerted by emotion perception components at this stage.

In a subsequent step, the acquired emotional structure can be assessed with learning-type tasks to estimate the test-taker's capability to learn this type of emotional knowledge. For example, after receiving information about the principal's emotional structure, the test-taker is presented with a vignette along with the principal's potential reactions. The vignette and the potential reactions would allow direct application (i.e., without any transfer) of the information gathered in the previous step. The agent's task is to rate the

strength of the principal's emotional reactions in the situation. Emotional reactions may include mood states, behaviors indicative of certain emotions, and the like.

Next, a new situation would be described that is emotion-laden. Several possible reactions (e.g., behavioral choices) are offered. The reactions represent different results from different emotional structures, only one of which most likely follows from the structure to be learned in the previous step. In this prediction-type task, the agent's job is to predict the response that the principal would most likely show. The responses offered cannot be directly derived from the knowledge about the emotional structure but require inferences and transfer to the new situation. As is usually done as part of the PAP, the agent is asked to act on behalf of the principal, as if the agent would actually be the principal. This highlights the fact that a core component of this approach to assess understanding emotions is the ability to be empathetic.

Theoretically this method provides more information than the standard instructions from a SJT (asking what the test taker *would* do or what the person in the vignette *should* feel) as it both recognizes that individuals with different emotional structures prefer different courses of action. In addition, this method allows for construction of theoretically "correct" answers, although care would be needed to avoid multiple viable answers.

The EPAP approach to measuring EI is a far more accurate simulation of social and emotional knowledge applied outside of a test room than many other approaches, as it acknowledges that different people have different emotional reactions, and that learning these reactions is an important skill. However, there are several concerns with this approach, first of which is that it has yet to be developed into a usable measure of EI, which places a heavy burden of development on any researcher interested in testing the validity of this approach. Another concern is that it may be difficult for individuals to learn any but the simplest emotional response contingency if only short vignettes are used.

Affective Forecasting and Emotional Intelligence

Affective forecasting (AF) is the ability to accurately forecast one's emotional state at some point in the future, typically in response to a specific emotion-inducing event (for a review of the AF literature, see Hsee & Hastie, 2006). While the research on AF as a relevant measure of EI is limited, high-EI individuals should be able to better predict their future emotions, both because of increased understanding and increased control of their emotions. The AF paradigm shares some similarities with the EPAP: In both cases the test-taker must make predictions based on event-emotion contingencies. In the case of AF, however, the principal and the agent are the same individual.

Dunn, Brackett, Ashton-James, Schneiderman, and Salovey (2007) recently published an article showing a relation between errors in AF and MSCEIT scores: $\beta = -.22$ for total scores and $\beta = -.31$ for Management scores (i.e., smaller errors in forecasting are associated with higher levels of EI, and particularly Management). Dunn et al. argue that AF constitutes a stable trait, as AF accuracy scores on two different tasks (a presidential election and a term paper) correlated at .28 (cf. however, Orchard, 2005). If it can be demonstrated that AF accuracy is indeed a stable individual differences variable (or ability), then it should be a useful tool in assessing EI, either as a part of the assessment process or in validating a measure of EI as a dependent variable. In either case, this is an area of research that deserves further attention.

Conclusion

In the light of the potential of this last group of measures, let us reconsider the test-developer struggling with the task of creating a valid assessment tool to assess EI. Certainly, the plethora of possibilities has not simplified her task, since there are now a number of ways to move forward. Although there is still no one clear path to follow, there are at least two clear principles. Firstly, the self-report methodology does not result in a measure of emotional intelligence. Secondly, no single test measures all aspects of EI: Each individual scale assesses one sub-construct (e.g., recognition, understanding or management of emotions, empathy, or emotional expressiveness). As a corollary to this second principle, we suggest that the test developer would need to select the development paradigm most appropriate for the construct of interest. For example, inspection time measures seem to be best for assessing simple information processing-type constructs such as emotion recognition, while the SJT paradigm might best be used to assess emotion management. Given the dependence of the EPAP on reading another person's emotions and values, this might best be used to assess understanding emotions and empathy. In any case, developing a test that purports to measure "EI" as a general factor makes little sense, given that different branches or sub-constructs show different relationships to different criteria.

In comparing the strengths and weaknesses of the different approaches, some generalizations can be made. Advantages of the information processing and emotion recognition tasks include easy-to-understand instructions, low verbal load (making this nuisance factor less problematic), and an obvious (and uncontroversial) scoring key. However, these simple information-processing tools do not offer the rich context, ecological validity, or cognitive complexity of the situation-based assessments (i.e., SJT, EPAP, AF, and Conditional Reasoning paradigms).

Overall, the potential research questions suggested by this chapter could easily occupy the time of numerous researchers for years to come. For example,

research is clearly needed on each specific measure, particularly the accumulation of validity evidence. Over time, this would allow for the development of even more accurate and valuable assessments of EI, and a taxonomic model that maps respective subcomponents of EI to outcomes that are valued by society. The results of such a cumulative program of research would silence the critics of EI research and provide great insight into individual differences in emotions and how they play out in every day life.

Acknowledgments This research was supported in part by U.S. Army Research Institute (ARI) Contract W91WAW-07-C-0025 to the Educational Testing Service (ETS). We thank Patrick Kyllonen, Don Powers, Matthew Ventura, and Dan Eignor for comments on an earlier draft of this manuscript, and Mary Lucas, Jennifer Minsky, and Cris Valkyria for supporting the preparation of this chapter. All statements expressed in this article are the authors' and do not necessarily reflect the official opinions or policies of the U.S. government, ARI, or ETS.

Notes

1. Resolving the efficacy of these four criteria is no trivial undertaking, as correlational evidence needs, for example, to take into account corrections for statistical artifacts and may be sample dependent. Meta-analyses aimed at these issues are appearing in the literature (see e.g., Roberts et al., 2008), as are studies where both corrected and uncorrected correlation coefficients are reported.
2. Emotional intelligence shares much conceptual overlap with the "social intelligence" domain, and we treat emotional and social intelligence as largely overlapping constructs in the following account. Indeed, the term *social intelligence*, when traced to its likely progenitor John Dewey (1909; see Landy, 2006), does share close conceptual parallels with the definition of EI: "Ultimate moral motives and forces are nothing more or less than *social intelligence* (italics in the original) – the power of observing and comprehending social situations – and *social power* (italics in the original) – trained capacities of control – at work in the service of social interests and aims" (p. 43). Ultimately, social and emotional intelligence will need to be disentangled, but for present purposes we are agnostic on this point.

References

- Ackerman, P. L., & Heggestad, E. D. (1997). Intelligence, personality and interests: Evidence for overlapping traits. *Psychological Bulletin*, 121, 219–245.
- Ashton, M. C., Lee, K., Vernon, P. A., & Jang, K. L. (2000). Fluid intelligence, crystallized intelligence, and the openness/intellect factor. *Journal of Research in Personality*, 34, 198–207.
- Austin, E. J. (2005). Emotional intelligence and emotional information processing. *Personality and Individual Differences*, 39, 403–414.
- Austin, E. J., & Saklofske, D. H. (2005). Far too many intelligences? On the communalities and differences between social, practical, and emotional intelligences. In R. Schulze & R. D. Roberts (Eds.), *International handbook of emotional intelligence* (pp. 107–128). Cambridge: Hogrefe & Huber.

- Bar-On, R. (1997). *Bar-On Emotional Quotient Inventory (EQ-i): Technical manual*. Toronto: Multi-Health Systems.
- Barchard, K. A., & Hakstian, R. A. (2004). The nature and measurement of emotional intelligence abilities: Basic dimensions and their relationships with other cognitive ability and personality variables. *Educational and Psychological Measurement*, 64, 437–462.
- Bastian, V. A., Burns, N. R., & Nettelbeck, T. (2005). Emotional intelligence predicts life skills, but not as well as personality and cognitive abilities. *Personality and Individual Differences*, 39, 1135–1145.
- Baum, K. M., & Nowicki, S. (1998). Perception of emotion: Measuring decoding accuracy of adult prosodic cues varying in intensity. *Journal of Nonverbal Behavior*, 22, 89–108.
- Brackett, M. A., Mayer, J. D., & Warner, R. M. (2004). Emotional intelligence and its relation to everyday behavior. *Personality and Individual Differences*, 36, 1387–1402.
- Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N., & Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology*, 91, 780–795.
- Brackett, M. A., Warner, R. M., & Bosco, J. S. (2005). Emotional intelligence and relationship quality among couples. *Personal Relationships*, 12, 197–212.
- Buck, R. (1984). *The communication of emotion*. New York: Guilford.
- Chapin, F. S. (1942). Preliminary standardization of a social insight scale. *American Sociological Review*, 7, 214–228.
- Ciarrochi, J. V., Chan, A. Y. C., & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences*, 28, 539–561.
- Costanzo, M., & Archer, D. (1993). *The Interpersonal Perception Task-15 (IPT-15): Videotape and technical manual*. Berkley, CA: University of California Extension Media Center.
- Danforth, M. M., McIntire, K. A., & Schneider, H. G. (1997, February). Measuring cue perception: An assessment of the reliability and validity of three instruments. Poster presented at the meetings of the Southeastern Psychological Association, Atlanta, GA.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. Chicago, IL: University of Chicago Press.
- 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.
- Derksen, J., Kramer, I., & Katzko, M. (2002). Does a self-report measure for emotional intelligence assess something different than general intelligence? *Personality and Individual Differences*, 32, 37–48.
- Dewey, J. (1909). *Moral principles in education*. New York: Houghton Mifflin.
- Dunn, E. W., Brackett, M. A., Ashton-James, C., Schneiderman, E., & Salovey, P. (2007). On emotionally intelligent time travel: Individual difference in affective forecasting ability. *Personality and Social Psychology Bulletin*, 33, 85–93.
- Ekman, P. (2004). *Emotions revealed*. New York: Times Books.
- Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Journal of Environmental Psychology and Nonverbal Behavior*, 1, 56–75.
- Ekman, P., & Friesen, W. V. (1978). *The facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.
- Ekman, P., & Rosenberg, E. L. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system*. New York: Oxford University Press.
- Fiedler, K., & Bluemke, M. (2005). Faking the IAT: Aided and unaided response control on the implicit association test. *Basic and Applied Social Psychology*, 27, 307–316.
- Freudenthaler, H. H., & Neubauer, A. C. (2005). Emotional intelligence: The convergent and discriminant validities of intra- and inter-personal abilities. *Personality and Individual Differences*, 39, 569–589.

- Freudenthaler, H. H., & Neubauer, A. C. (2007). Measuring emotional management abilities: Further evidence of the importance to distinguish between typical and maximum performance. *Personality and Individual Differences*, 42, 1561–1572.
- Gohm, C. L., Corser, G. C., & Dalsky, D. J. (2005). Emotional intelligence under stress: Useful, unnecessary, or irrelevant? *Personality and Individual Differences*, 39, 1017–1028.
- Goldenberg, I., Matheson, K., & Mantler, J. (2006). The assessment of emotional intelligence: A comparison of performance-based and self-report methodologies. *Journal of Personality Assessment*, 86, 33–45.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology*, 77, 350–359.
- Hsee, C. K., & Hastie, R. (2006). Decision and experience: Why don't we choose what makes us happy? *TRENDS in Cognitive Science*, 10, 31–37.
- Huber, J., Ariely, D., & Fischer, G. (2002). Expressing preferences in a principal-agent task: A comparison of choice, rating, and matching. *Organizational Behavior and Human Decision Processes*, 87, 66–90.
- James, L. R. (1998). Measurement of personality via conditional reasoning. *Organizational Research Methods*, 1, 131–163.
- James, L. R., McIntyre, M. D., Glisson, C. A., Bowler, J. L., & Mitchell, T. R. (2004). The conditional reasoning measurement system for aggression: An overview. *Human Performance*, 17, 271–295.
- James, L. R., McIntyre, M. D., Glisson, C. A., Green, P. D., Patton, T. W., LeBreton, J. M., et al. (2005). A conditional reasoning measure for aggression. *Organizational Research Methods*, 8, 69–99.
- Landy, F. J. (2006). The long, frustrating, and fruitless search for social intelligence: A cautionary tale. In K. R. Murphy (Ed.), *A critique of emotional intelligence: What are the problems and how can they be fixed?* (pp. 81–123). Mahwah, NJ: Erlbaum.
- Lane, R. D. (2000). Levels of emotional awareness: Neurological, psychological, and social perspectives. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 171–191). San Francisco: Jossey-Bass.
- Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The levels of emotional awareness scale: A cognitive-development measure of emotion. *Journal of Personality Assessment*, 55, 124–134.
- Lopes, P. N., Brackett, M. A., Nezlek, J. B., Schütz, A., Sellin, I., & Salovey, P. (2004). Emotional intelligence and social interaction. *Personality and Social Psychology Bulletin*, 30, 1018–1034.
- MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, 8, 540–551.
- Magill-Evans, J., Koning, C., Cameron-Sadava, A., & Manyk, K. (1995). The child and adolescent social perception measure. *Journal of Nonverbal Behavior*, 19, 151–169.
- Matsumoto, D., LeRoux, J., Wilson, C., Raroque, J., Ekman, P., Yizarry, N., et al. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior*, 24, 179–209.
- Matthews, G., Emo, A. K., Funke, G., Zeidner, M., Roberts, R. D., Costa, P. T., Jr., et al. (2006). Emotional intelligence, personality, and task-induced stress. *Journal of Experimental Psychology: Applied*, 12, 96–107.
- Mauss, I. B., Evers, C., Wilhelm, F. H., & Gross, J. J. (2006). How to bite your tongue without blowing your top: Implicit evaluation of emotion regulation predicts affective responding to anger provocation. *Personality and Social Psychology Bulletin*, 32, 589–602.

- Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27, 267–298.
- Mayer, J. D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence*, 22, 89–113.
- Mayer, J. D., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, 59, 507–536.
- Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and EI: Educational implications* (pp. 3–34). New York: Basic Books.
- Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) user's manual*. Toronto: Multi-Health Systems.
- Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3, 97–105.
- McDaniel, M. A., & Nguyen, N. T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9, 103–113.
- Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the implicit association test. *Journal of Personality and Social Psychology*, 85, 1180–1192.
- Murphy, N. A., Hall, J. A., & Colvin, C. R. (2003). Accurate intelligence assessments in social interactions: Mediators and gender effects. *Journal of Personality*, 71, 465–493.
- Newsome, S., Day, A. L., & Catano, V. M. (2000). Assessing the predictive validity of emotional intelligence. *Personality and Individual Differences*, 29, 1005–1016.
- Orchard, B. M. (2005). *The effects of individual differences on affective forecasting*. Unpublished Masters Thesis, Clemson University, South Carolina.
- O'Sullivan, M. (2007). Trolling for trout, trawling for tuna: The methodological morass in measuring emotional intelligence. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns* (pp. 258–287). New York: Oxford University Press.
- Palmer, B., Gignac, G., Manocha, R., & Stough, C. (2005). A psychometric evaluation of the Mayer-Salovey-Caruso Emotional Intelligence Test version 2.0. *Intelligence*, 33, 285–305.
- Pérez, J. C., Petrides, K. V., & Furnham, A. (2005). Measuring trait emotional intelligence. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 181–202). Cambridge, MA: Hogrefe & Huber.
- Rivers, S. E., Brackett, M. A., Salovey, P., & Mayer, J. D. (2007). Measuring emotional intelligence as a set of mental abilities. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns* (pp. 258–287). New York: Oxford University Press.
- Roberts, R. D., Schulze, R., & MacCann, C. (2008). The measurement of emotional intelligence: A decade of progress? In G. Boyle, G. Matthews, & D. Saklofske (Eds.), *The Sage handbook of personality theory and assessment* (pp. 461–482). New York: Sage.
- Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J., & Maul, A. (2006). Exploring the validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) with established emotions measures. *Emotion*, 6, 663–669.
- Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusion. *Emotion*, 1, 196–231.
- Roberts, R. D., Zeidner, M., & Matthews, G. (2007). Emotional intelligence: Knowns and unknowns. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns* (pp. 419–474). New York: Oxford University Press.
- Roseman, I. J. (1984). Cognitive determinants of emotion: A structural theory. *Review of Personality and Social Psychology*, 5, 11–36.
- Rosenthal, R., Archer, D., Hall, J. A., DiMatteo, M. R., & Rogers, P. L. (1979). Measuring sensitivity to nonverbal communication: The PONS test. In A. Wolfgang (Ed.), *Nonverbal behavior: Applications and cultural implications*. New York: Academic Press.

- Rosip, J. C., & Hall, J. A. (2004). Knowledge of nonverbal cues, gender, and nonverbal decoding accuracy. *Journal of Nonverbal Behavior*, 28, 267–286.
- Scherer, K. R. (2007). Componential emotion theory can inform models of emotional competence. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence: Knowns and unknowns*. (pp. 101–126). New York: Oxford University Press.
- Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross Cultural Psychology*, 32, 76–92.
- Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper, J. T., Golden, C. J., et al. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences*, 25, 167–177.
- Seidel, K. (2007). *Social intelligence and auditory intelligence – useful constructs?* Unpublished Doctoral Thesis, Otto-von-Guericke-University, Magdeburg, Germany. Retrieval from: <http://diglib.uni-magdeburg.de/Dissertationen/2007/kriseidel.pdf>
- Stokes, T. L., & Bors, D. A. (2001). The development of a same-different inspection time paradigm and the effects of practice. *Intelligence*, 29, 247–261.
- Süß, H.-M., Seidel, K., & Weis, S. (2008). Neue Wege zur leistungs-basierten Erfassung sozialer Intelligenz und erste Befunde [New ways to performance-based assessment of social intelligence and new results]. In W. Sarges & D. Scheffer (Eds.), *Innovationen in der Eignungsdiagnostik* [Innovations in personnel selection]. Göttingen, Germany: Hogrefe.
- Tett, R. P., Fox, K. E., & Wang, A. (2005). Development and validation of a self-report measure of emotional intelligence as a multidimensional trait domain. *Personality and Social Psychology Bulletin*, 31, 859–888.
- Trinidad, D. R., & Johnson, C. A. (2002). The association between emotional intelligence and early adolescent tobacco and alcohol use. *Personality and Individual Differences*, 32, 95–105.
- Trinidad, D. R., Unger, J. B., Chou, C.-P., & Johnson, C. (2005). Emotional intelligence and acculturation to the United States: Interactions on the perceived social consequences of smoking in early adolescents. *Substance Use and Misuse*, 40, 1697–1706.
- Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of Nonverbal Behavior*, 20, 65–80.
- Weis, S., & Süß, H.-M. (2005). Social intelligence – A review and critical discussion of measurement concepts. In R. Schulze & R. D. Roberts (Eds.), *Emotional intelligence: An international handbook* (pp. 204–230). Cambridge, MA: Hogrefe & Huber.
- Weis, S., & Süß, H.-M. (2007). Reviving the search for social intelligence – A multitrait-multimethod study of its structure and construct validity. *Personality and Individual Differences*, 42, 3–14.
- Zeidner, M., Shani-Zinovich, I., Matthews, G., & Roberts, R. D. (2005). Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure. *Intelligence*, 33, 369–391.