



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Measurement of ability emotional intelligence: Results for two new tests

**Citation for published version:**

Austin, EJ 2010, 'Measurement of ability emotional intelligence: Results for two new tests' British Journal of Psychology, vol. 101, no. 3, pp. 563-578. DOI: 10.1348/000712609X474370

**Digital Object Identifier (DOI):**

[10.1348/000712609X474370](https://doi.org/10.1348/000712609X474370)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

British Journal of Psychology

**Publisher Rights Statement:**

© Austin, E. J. (2010). Measurement of ability emotional intelligence: Results for two new tests. British Journal of Psychology, 101(3), 563-578doi: 10.1348/000712609X474370

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



Measurement of ability emotional intelligence (EI): Results for two new tests

Elizabeth J. Austin

Department of Psychology, University of Edinburgh, UK

Correspondence to: Dr. E. J. Austin, Department of Psychology, School of  
Philosophy, Psychology and Language Sciences, University of Edinburgh, 7 George  
Square, Edinburgh EH8 9JZ, UK. [Elizabeth.Austin@ed.ac.uk](mailto:Elizabeth.Austin@ed.ac.uk) Phone +44 131 651  
1305. Fax +44 131 650 3461

Acknowledgements

Michael Scullin provided valuable assistance in the development of the Blends task.  
Copies of the STEM and STEU were kindly provided by Dr. Carolyn MacCann. This  
research was funded by the British Academy and the Leverhulme Trust.

## Measurement of ability emotional intelligence (EI): Results for two new tests

### Abstract

Emotional intelligence (EI) has attracted considerable interest amongst both individual differences researchers and those in other areas of psychology who are interested in how EI relates to criteria such as well-being and career success. Both trait (self-report) and ability EI measures have been developed; the focus of this paper is on ability EI. The associations of two new ability EI tests with psychometric intelligence, emotion perception and the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) were examined. The new EI tests were the Situational Test of Emotion Management (STEM) and the Situational Test of Emotional Understanding (STEU) (MacCann & Roberts, 2008). Only the STEU and the MSCEIT Understanding Emotions branch were significantly correlated with psychometric intelligence, suggesting that only understanding emotions can be regarded as a candidate new intelligence component. These understanding emotions tests were also positively correlated with emotion perception tests, and STEM and STEU scores were positively correlated with MSCEIT total score and most branch scores. Neither the STEM nor the STEU were significantly correlated with trait EI tests, confirming the distinctness of trait and ability EI. Taking the present results as a starting point, approaches to the development of new ability EI tests and models of EI are suggested.

**Keywords:** Emotional intelligence, emotion perception, validity

The relatively new construct of emotional intelligence (EI) has been a focus of much research activity and general interest. This can be attributed to the appeal of examining the ways in which people differ in their emotion-related capabilities, particularly if such differences can be linked to important real-world criteria. Because of the diversity of EI-related outcomes, EI research is carried out in a wide range of psychology sub-disciplines beyond the area of individual differences. Findings in one area where EI has been identified as potentially salient, psychological well-being, are discussed in more detail below. Other active areas within EI research include the study of associations with job performance and leadership (e.g. Barbuto & Burbach, 2006; Carmelli & Josman, 2006; van Rooy & Viswesvaran, 2004), academic success (e.g. Parker, Summerfeldt, Hogan, & Majeski, 2004), social competence (e.g. Mavroveli, Petrides, Rieffe, & Bakker, 2007) and quality of relationships (e.g. Smith, Heaven, & Ciarrochi, 2008; Zeidner & Kaluda, 2008).

#### *Measurement of EI and evidence for validity*

As EI research has developed it has diverged into two subfields, trait and ability EI (Petrides & Furnham, 2003; Petrides, Furnham, & Mavroveli, 2007). The ability EI perspective takes as a starting point the presumption (discussed in more detail in the next section) that EI is a cognitive ability which is not measured by standard intelligence tests and which relates to reasoning and problem solving in the emotional domain. By contrast, the trait EI perspective regards EI as being located within the personality domain <sup>1</sup>.

There is now a well-established consensus that the two versions of EI are distinct and do not measure the same construct. The measurement methods for the two forms of EI differ. Ability EI tests resemble standard intelligence tests, whilst trait EI is

measured via self-report; correlations between trait and ability EI test scores have consistently been found to be low. Example correlations found in recent studies are .04 (Bastian, Burns, & Nettelbeck, 2005), .21, .18 (Bracket & Mayer, 2003), .34 (O'Connor & Little, 2003). In addition, the correlation patterns with intelligence and personality differ for trait and ability EI. Trait measures show medium to large correlations with the major five-factor model personality dimensions, and are generally uncorrelated with intelligence. Ability measures show the reverse pattern of positive correlations with intelligence test scores and low correlations with personality.

Correlations between the MSCEIT and measures of general ability have been reported in recent studies to be in the range .25 to .32 (Bastian et al., 2005; O'Connor & Little, 2003; Zeidner, Shani-Zinovich, Matthews, & Roberts, 2005), with the corresponding range for trait EI being -.21 to .15. A meta-analysis covering the earlier ability EI MEIS test and trait EI reported a true score correlation of .33 for ability EI and .09 for trait EI (van Rooy & Viswesvaran, 2004). There is also evidence for stronger associations of ability EI with crystallised than with fluid ability (e.g. Farrelly & Austin, 2007; MacCann, Roberts, Matthews, & Zeidner, 2004; Mayer, Caruso, & Salovey, 1999; Roberts, Zeidner, & Matthews, 2001). Typical correlations of EI measures with personality traits are exemplified by the results reported for the Big Five by Austin, Farrelly, Black, and Moore (2007). These were, for trait EI, Neuroticism -.56, Extraversion .43, Openness .12, Agreeableness .37, Conscientiousness .24, and for the MSCEIT -.07, .03, .11, .21, -.01. Both these correlational findings and the measurement method for trait EI, which involves self-reports of typical performance rather than problem-solving (maximal) performance, indicate that trait EI tests do not measure a form of intelligence, and any claim that

they do so is mistaken. The viewpoint that trait EI is appropriately placed in the personality domain is supported by a recent factor-analytic study (Petrides, Pita, & Kokkinaki, 2007) which showed trait EI emerging as a separate lower-order factor within established personality models.

There is evidence for the criterion validity of both trait and ability EI measures. Taking the area of health and well-being as an example, the broad theoretical prediction is that EI should be positively associated with measures of psychological health such as life satisfaction and negatively associated with indicators of psychological distress such as stress and depression. The reason for expecting such associations is that intrapersonal components of EI (such as the ability to perceive one's own emotions and regulate mood) should assist with stress management and adaptive coping, with interpersonal components of EI also being expected to play a role in well-being by enabling high EI individuals to form better social support networks. Both trait and ability EI have been found to correlate as expected with psychological well-being measures, for example positively with life satisfaction and negatively with anxiety, depression and loneliness (e.g. Bastian et al., 2005; Ciarrochi, Chan, & Caputi, 2000; Saklofske, Austin, & Minski, 2003), whilst associations between EI and adaptive coping have also been found (Bastian et al., 2005; Saklofske, Austin, Galloway, & Davidson, 2007). There is evidence for the incremental validity of EI in accounting for additional variance in health and well-being and other criteria when intelligence and personality are controlled for, but the additional percentage variance accounted for is generally 5% or less (Zeidner, Roberts, & Matthews, 2008).

#### *Issues in ability EI research*

The main focus of this paper is ability EI. The claim that this construct represents

a new form of intelligence not currently assessed by standard intelligence test batteries is a controversial and disputed one (e.g. Petrides et al., 2007; Zeidner et al., 2008). The associations found between ability EI and intelligence mentioned above can be interpreted as evidence for EI as a new form of intelligence which fits into the positive manifold of cognitive abilities (Carroll, 1993), although a more detailed examination of the findings to date shows mixed results, with not all EI components being consistently associated with intelligence test scores, and with the most robust findings relating only to tests of emotional understanding (e.g. Farrelly & Austin, 2007; Roberts et al., 2001; Zeidner et al., 2005). It is also the case that establishing that a test is correlated with intelligence tests is not logically equivalent to establishing that the test itself measures a form of intelligence, with additional validity evidence also being necessary (for example, for EI, evidence of associations with objective behavioural measures of emotional capability). Some specific concerns that have been raised in connection with the methods used to score ability EI tests, which tend to undermine rather than support the claim that EI is an intelligence, are discussed below.

Current research on ability EI is dominated by the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT, Mayer, Salovey, & Caruso, 2002), with results also available for its predecessor the MEIS (Mayer et al, 1999). These tests are both based on a theoretical model which divides EI into four branches: perceiving, using, understanding and managing emotions.

Whilst there a body of evidence supporting the criterion validity of the MSCEIT (see Mayer, Roberts, & Barsade, 2008 for a recent summary), there has been criticism of the use of consensus or expert scoring for this test (e.g. Matthews, Zeidner, & Roberts, 2007). The two forms of scoring are highly consistent, correlating in the

range .93-.98 (Mayer et al., 2002), but either method results in the maximum score on an item being obtained when the respondent chooses the option previously endorsed by the majority of either a standardisation sample or a group of experts. This scoring method does not correspond with that used in standard intelligence tests, where test items have well-defined correct answers. The structural validity of the MSCEIT has also been questioned, as several studies indicate a lack of fit with its theoretical factor structure (Keele & Bell, 2008; Palmer, Gignac, Manoca, & Stough, 2005; Rossen, Kranzler, & Algina, 2008).

In addition to specific issues which have been raised by critics of the MSCEIT, an important contrast between the situation for ability EI measurement and that in other areas of individual differences research is the current lack of alternatives to this test. Within the fields of both intelligence and personality researchers can select from a wide range of tests. This is also the case for trait EI, with a number of independently-developed tests being available, for example the TEIQue (Petrides & Furnham, 2003), the SSRI, (Schutte et al., 1998) and the EQ-i (Bar-On, 1997). This means that for trait EI, as for intelligence and personality, psychometric properties such as reliability and validity can be compared between tests, as can the performance of tests developed using different theoretical perspectives.

By contrast to the situation with trait EI, there is currently no alternative broad-bandwidth ability EI measure (i.e. one which assesses all the EI components specified in the Mayer-Salovey-Caruso or any competing model) similar in scope to the MSCEIT. In a recent survey of the current status of EI research, Roberts, Zeidner, and Matthews (2007) emphasised the importance of the development of a range of tests for the assessment of ability EI in order to clarify issues related to both its theory and measurement. Although alternative broad-bandwidth measures of ability EI are



currently lacking, progress has been made in the development of new tests targeting EI components, two of which are examined in the present work.

The Situational Test of Emotion Management (STEM) and Situational Test of Emotional Understanding (STEU) were developed by MacCann and Roberts (2008), who also presented validity evidence for these tests (for example positive correlations with an intelligence test, divergent validity with respect to personality and, for the STEU, positive association with a MEIS subtest). Although targeting two EI components also assessed by branches of the MSCEIT, they differ from it both in the manner in which the test items were developed and, importantly, in availability to researchers and transparency of test scoring criteria. The MSCEIT is a commercial test with scoring performed by a test company, rather than the scoring key being made available to researchers. The STEM and STEU, and their scoring keys, are freely available. This means that (in contrast to the MSCEIT) access to these tests is not limited by considerations of cost, and work on developing and improving them (e.g. creating and evaluating new items) can take place within the research community.

The content of the items of the STEU was derived from Roseman's (2001) appraisal-based emotion model, in which the emotion felt by a person is derived from features of their appraisal of the emotion-generating situation (for example relief is associated with the perception that an unpleasant situation has stopped or been averted). The test covers the emotions sadness, pride, relief, joy, regret, gratitude, distress, hope, contempt, surprise, frustration, anger, fear and dislike. It has 42 multiple-choice items which test respondents' knowledge of which emotion is most likely to be felt in a range of situations. The situations are described either in a decontextualised manner or in the context of a concrete example from work or personal life, with an equal number of decontextualised, personal and work items.

Scoring is based on the theoretical model used to construct the items meaning that, within the framework of this model, veridical scoring is possible, for example the response “relief” to the item asking what a person would feel if an annoying neighbour moved away is scored correct. This veridical scoring system was used in the present study, with each STEU item scored as either correct or incorrect. The STEM was developed using the situational judgement test method (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001). The test items were created by extracting scenarios in which emotion management would be required from semi-structured interviews with 50 individuals. A variety of scenarios covering the emotions sadness, anger and fear across a range of situation types (e.g. for anger situations involving arguments, being impeded in goal striving and unfairness), and using personal and work contexts were generated. Potential response options to these were then generated by a second group of 99 individuals; four of these responses were selected for each item and these were then assessed by experts to obtain scoring weights. The test comprises 44 multiple-choice items in which a scenario is presented (for example a close friend moving overseas permanently); the respondent has to choose among the four response options for the most effective action for the person experiencing that situation.

### *The present study*

This paper extends the information available on the STEM and STEU by examining their associations with the MSCEIT and with performance on emotion perception tests. Emotion perception is a key criterion variable which needs to be examined when validating EI tests; individuals high in EI would be expected to display superior emotion perception abilities such as the identification of facial expressions of emotion. In addition to being considered a component of EI, emotion

perception is regarded as underpinning higher-level capabilities such as understanding and managing emotions, since accurate perception of emotions is necessary for the exercise of these higher-level abilities (Mayer, Salovey, Caruso, & Sitarenios, 2001). Thus scores on ability tests assessing all components of EI are expected to be positively associated with performance on emotion perception tasks. In the present study two tests of emotion perception were used, one of which was not time-limited, whilst the other was an inspection time (IT) task in which a limited amount of time was available to process an image of a face and identify its expression. This task has been used in previous work on associations between EI and emotion perception (Farrelly & Austin, 2007). Two intelligence tests, a vocabulary test and a test involving completing a number of letter series by deriving the logical rule underlying each of them were also included, and the divergent validity of the STEM and STEU with respect to trait EI was examined by also including two trait EI measures.

On the basis of the existing results for the STEM, STEU and MSCEIT, and the theoretical alignment of the STEM and STEU with the MSCEIT Managing and Understanding Emotions branches, the following hypotheses were formulated:

H1. STEM and STEU scores will be positively and significantly correlated with MSCEIT branch and total scores. Because ability EI components are expected (like intelligence battery sub-tests) to form a general factor, significant correlations with all MSCEIT branches would be expected, but the STEM would be expected to correlate most strongly with the MSCEIT Managing Emotions branch and the STEU most strongly with the MSCEIT Understanding Emotions branch.

H2. STEM, STEU and MSCEIT scores will be positively correlated with scores on

emotion perception tasks. This follows from the above discussion of emotion perception ability as a criterion variable in validating EI tests.

H3. STEM, STEU and MSCEIT scores will be positively correlated with intelligence test scores. In the light of previous findings reviewed above, stronger correlations with vocabulary test performance (a crystallised ability measure) than with series completion would be expected.

H4. The STEM and STEU will display divergent validity with respect to trait EI, showing weak or zero associations with trait EI scores. This expectation follows from findings, summarised above, showing that trait and ability EI are distinct constructs.

## Method

### *Participants*

The participants were 339 undergraduate students (238 females, 101 males) attending a UK university. The mean age of the group was 21.96 years, standard deviation 4.22 years.

### *Materials*

In addition to the STEM and STEU, which are described in the introduction, the following measures were included:

*Cognitive ability.* Two tests from the Gf/Gc Quickie Test Battery (Stankov, 1997) were used. This battery is appropriate for and has been previously used with university student groups (e.g. Farrelly & Austin, 2007; MacCann et al., 2004). The Letter Series test is a series completion task with 15 letter items which can be solved by applying logical rules. The Vocabulary test contains 18 multiple-choice items. The

Vocabulary test was untimed; the Series test was administered with a time limit of four minutes.

*Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) v2.0.* (Mayer et al., 2002). This 141-item measure provides four branch scores (Perceiving, Using, Understanding and Managing Emotions) and an overall EI score; the two area scores (Experiential, Strategic) also available for this test as composites of branch scores are not considered here. Scores are provided by the test company, Multi-Health Systems. The consensus scoring option and standardised scores were used.

*Face Blends.* This computer task required participants to identify the emotion expressed by a face which was made up of a blend of two pictures of the same person expressing two different emotions. Stimuli were a blend of two of the three emotions of sadness, anger and fear with the proportion of a particular emotion in the blend (e.g. anger) being 10%, 20%, 30%, 40%, 60%, 70%, 80% or 90%, together with unblended faces showing each of the three emotions. The blends were derived from one female and one male exemplar from the Ekman pictures of facial affect (Ekman & Friesen, 1976) and were obtained from the stimulus set supplied with the Facial Expressions of Emotion – Stimuli and Tests (FEEST; Young, Perrett, Calder, Sprengelmeyer, & Ekman, 2002). All eight blends were used for each of the three emotion pairs for each exemplar, giving a total of 48 blended emotion items, which combined with the three unblended emotion pictures for each exemplar gave a total of 54 items. The correct answer for each stimulus was the majority emotion. The image size on the screen was 9 cm x 14 cm. At the start of the task participants were told that each face would “mostly” express anger, fear or sadness and were instructed to respond as quickly as possible without sacrificing accuracy. The presentation of each face was preceded by a fixation cross (+) in the middle of the screen displayed for 500

ms; the face was then displayed until the participant responded by pressing 7 (anger), 8 (fear) or 9 (sadness) on the numeric keypad. Stimuli were presented in a random order subject to the stipulation that the same emotion type would be repeated no more than three times in a row; the ordering was the same for all participants. Each response was followed by a blank screen for 500 ms and then the fixation cross for the next trial. The series of 54 experimental trials was preceded by six practice trials in which one male and one female exemplar of the three pure emotions (using pictures of different individuals from those used in the experimental trials) was used, with participants being given feedback on their performance; no feedback was given in the subsequent experimental trials.

*Sad faces inspection time (IT) task* (Austin, 2004, 2005). This task involved stimuli (two male and two female examples with sad and neutral expressions, giving a total of eight stimuli) taken from the Ekman and Friesen (1976) stimulus set. The same mask, a neutral female face (different from any of the stimuli), was used in all trials. Stimuli were presented on a computer screen; the on-screen size for the stimuli and mask was 55 x 85 mm. After the appearance of the mask, which was displayed for 500 ms, participants were instructed to press a key to indicate whether they thought that the stimulus face was sad or neutral. The task began with a practice session with each stimulus being presented twice, once at a duration of 500 ms and once at a duration of 800 ms; stimulus type /duration combinations were presented in a random order. Participants were able to repeat the practice session as many times as they wished. The data collection phase involved each stimulus being presented once at each of the following durations: 25, 32, 44, 57, 63, 75, 88, 100, 150, 250, 350, 400 ms, with stimulus type/duration combinations again presented in a random order. Thus each participant performed eight trials at each duration and 96 trials in total.

The score for this task was the total number of stimuli identified correctly.

*Trait EI.* The trait EI tests were the 30-item short form of the TEIQue (TEIQue-SF; Petrides & Furnham, 2006) and a subset of 22 items from the modified version of the Schutte et al. (1998) EI scale (SSRI) described by Austin, Saklofske, Huang, and McKenney (2004). These were the items which loaded on this scale's two strongest factors.

### *Procedure*

An initial group of participants was recruited to examine the associations of the STEM and the STEU with other measures. Group 1 included 104 participants who completed the STEM, STEU, Vocabulary, Series, SSRI and the IT task; 53 of these participants also completed the TEIQue. A further 100 participants completed the STEU, SSRI, TEIQue and the IT task. Group 2 comprised 135 participants in a separate study who completed the STEM, STEU, MSCEIT, Vocabulary, Series and the Blends task. The recruitment procedure was identical for each study, via a website advertising part-time jobs and opportunities for research participation which could be accessed by all students within the university; re-participation was not permitted. Participants were tested individually in a quiet room; Group 2 participants were given a link to access the MSCEIT on the web and completed it prior to attending their individual session. The STEM, Series and Vocabulary tests were completed on paper; the remaining tests were in computerised format. Some results on reaction times to the TEIQue, SSRI and STEU items have been presented elsewhere [citation removed in accordance with journal policy on self-identification].

### Results

Examination of the test scores showed that there were no significant differences in these between the groups, so data for the two groups were combined for the analyses

below. The total number of participants who completed each test and descriptive statistics are shown in Table 1. The internal reliabilities of the STEM and STEU were .67, .48, whilst for the MSCEIT branches internal reliabilities were Perceiving .86, Using .58, Understanding .66, Managing .66, and the total score internal reliability was .90. Internal reliabilities for the TEIQue-SF and SSRI were .81, .87 respectively. Plots of the scores for the Vocabulary and Series tests indicated that these were normally distributed with no ceiling effects, confirming that these tests were of reasonable difficulty for students. Examination of sex differences showed that, after correction for multiple comparisons, the only significant differences were for the STEM and the trait EI SSRI scale, with females scoring higher on both of these ( $t(237) = 4.53$ , Cohen's  $d = .59$ ,  $p < .001$ ;  $t(201) = 2.77$ ,  $d = .39$ ,  $p = .006$ ).

Table 1 near here

#### *Associations of the STEM and STEU with MSCEIT scores*

Table 2 shows correlations amongst the STEM, STEU, MSCEIT total score and MSCEIT branch scores. STEM and STEU scores were positively and significantly correlated and also showed significant associations with MSCEIT total score and some branch scores, with the STEM being significantly correlated with the MSCEIT Using, Understanding and Managing branches and the STEU with the Using and Understanding branches. As expected, the STEU showed a higher correlation with the Understanding branch than with the other MSCEIT branches ( $p$  values for correlation difference, .012 for Perceiving, .035 for Using and .015 for Managing, all  $N = 135$ ). By contrast, the correlations of the STEM with the Using, Understanding and Managing branches were not significantly different from one another. Partial



correlations controlling for scores on the intelligence tests are also shown. It can be seen that the differences between the full and partial correlations were generally small, suggesting that the associations amongst these tests are not accounted for by intelligence. None of these differences was significant.

*Associations of ability EI tests with emotion perception tasks*

Table 3 shows the correlations of the STEM, STEU and MSCEIT with performance on the Blends task, and associations of the STEM and STEU with performance on the Sad Faces IT task. Here the STEU and MSCEIT Understanding but not the STEM or other MSCEIT branches were significantly correlated with emotion task performance. For the IT tasks, in order to examine how ability EI scores were related to speed of information processing, correlations were also examined for summed scores on the shortest (25-53ms), intermediate (64-100ms) and longest (150-400ms) durations. STEM score was not significantly correlated with any of these scores, but STEU score was significantly correlated with scores at both intermediate and long durations ( $r = .24, p = .001, r = .21, p = .006$ ). Thus higher STEU scores were not associated with improved performance when very rapid processing of the stimulus was required and participants were performing at or near chance levels (proportion correct .53-.67), but high scorers were advantaged as the stimulus duration increased beyond 53ms.

*Associations of ability EI tests with intelligence test scores*

The associations of the STEM, STEU and MSCEIT with Vocabulary and Series test scores are also shown in Table 3. Both Vocabulary and Series showed significant positive correlations with MSCEIT Understanding, and Vocabulary also with the STEU; Series was negatively correlated with MSCEIT Perceiving.

Tables 2 and 3 near here

### *Associations of the STEM and STEU with trait EI*

The correlations of the STEU with the TEIQue and SSRI were .03 ( $N = 144$ ), -.04 ( $N = 195$ ); the corresponding correlations for the STEM were .12 ( $N = 53$ ), .13 ( $N = 103$ ). None of these correlations were significant.

### *Reliability issues - additional analyses*

A concern with the above findings is the low internal reliability of the STEU. Comparing the internal reliabilities of the STEM, STEU and MSCEIT branches, the internal reliability of the STEM (although lower than desirable) falls within the range of values for the MSCEIT branches, whilst the STEU is outside this range. Given that the internal reliabilities of the STEU and to a lesser degree the STEM were found to be low by standard psychometric criteria, item analysis was used to identify a subset of more highly intercorrelated items within each test, with the item whose removal would raise the internal reliability of the test the most being removed at each iteration until no further reliability gain could be obtained. This procedure resulted in the retention of 29 STEM and 30 STEU items<sup>2</sup>. The internal reliability of the shortened STEM was .73, which is within the psychometrically acceptable range. The internal reliability of the shortened STEU was .58, which is more comparable to the values for some of the MSCEIT branches, but still undesirably low. Examination of the correlations of the shortened and more reliable versions of the STEM and STEU with the MSCEIT and emotion perception tasks showed that the correlation pattern was similar to that found with the full-length scales and with no significant differences between correlations calculated using the short and full-length tests.

Discussion

The results described above provide useful information about the psychometric properties of two new measures of ability EI, the STEM and the STEU, extending the results of MacCann and Roberts (2008) by examining the associations of these tests with the MSCEIT, emotion perception tasks, and trait EI. It was found that the STEM and STEU were significantly correlated with MSCEIT total and branch scores (in line with H1), although not with the Perceiving Emotions branch or, in the case of the STEU, with the Managing Emotions branch. At the more detailed level, evidence was found that the STEU was, as expected, most strongly associated with the MSCEIT Understanding Emotions branch, but the STEM did not show a significantly stronger association with the Managing Emotions branch compared to its other significant associations with MSCEIT branches. This suggests that there is convergence between the STEU and its corresponding MSCEIT branch, but this is not apparent for the STEM in these data. It should be noted that the less than optimal reliabilities for the STEM, STEU and some MSCEIT branches place an upper bound on their correlations with other measures, so the current results do not provide an unequivocal verdict on whether the STEM and STEU measure the same constructs as MSCEIT Managing and Understanding respectively. The correlations amongst the STEM, STEU and MSCEIT branch and full-scale scores did not change significantly when intelligence test scores were partialled out. This indicates that the overlap amongst these tests is not accounted for by intelligence, and in particular by the verbal skills required to comprehend and successfully complete these mostly highly verbal tests (Wilhelm, 2005).

The expected associations with emotion perception tasks (H2) were found for the STEU and MSCEIT Understanding but not for the STEM and the other MSCEIT branches. In terms of criterion validity, this indicates that the tests of emotional

understanding show the appropriate properties. The trend in STEU correlation as a function of the stimulus duration for the IT task showed that high scorers were only advantaged at intermediate and long stimulus durations, but did not show any performance advantage for very brief stimulus durations, suggesting that performance was enhanced for conscious rather than pre-conscious processing of emotional stimuli. Interestingly, the MSCEIT Perceiving branch was found not to correlate with emotion perception task performance; this result, and similar findings in other studies (Farrelly & Austin, 2007; Roberts et al., 2006) suggest that the Perceiving branch is not a measure of emotion perception as commonly understood. The lack of convergent validity evidence between MSCEIT Perceiving and other emotion perception tests clearly requires further investigation. An issue that needs to be considered in future work is that the MSCEIT Perceiving items are different in format from the veridically scored items (requiring the identification of a single emotion) used in the present study and in those cited above, instead requiring an estimate of the extent to which a face or a scene conveys each of a number of emotions.

Only the STEU and MSCEIT Understanding were significantly positively correlated with intelligence (H3). The expected differential patterning of correlations with the vocabulary and series completion test was seen only for the STEU. As above, in the discussion of correlations amongst the EI tests, the question arises of whether the correlations of ability EI tests with intelligence tests (where found) are accounted for by common verbal content. It is certainly the case that the STEM, STEU and MSCEIT (with the exception of the pictorial Perceiving branch items) all require a high level of verbal comprehension, but only MSCEIT Understanding and the STEU were significantly correlated with intelligence, whilst MSCEIT Using and Managing, and the STEM (which also use verbally-based items) did not correlate with

intelligence. This suggests that the observed correlations between the emotional understanding tests and intelligence are not solely accounted for by verbal ability. The issue of why only the tests of emotional understanding were positively correlated with intelligence is an interesting one. Within the framework of their hierarchical model of EI Mayer et al. (2001, p235) argue that understanding emotions is the “most cognitively saturated” of the four EI branches, with the strongest association to abstract reasoning and emotional information-processing; the current findings are consistent with this viewpoint.

The distinctiveness of the STEM and STEU from trait EI (H4) was confirmed, in line with previous findings on associations between trait and ability EI tests (Bastian et al., 2005), with neither test being significantly correlated with the trait EI measures.

A more general issue is whether ability EI represents a new form of intelligence. The current results are in line with previous findings (Farrelly & Austin, 2007; Roberts et al., 2001; Zeidner et al., 2005) which indicate that, within the framework of the Mayer-Salovey-Caruso ability EI model, only measures of understanding emotions show consistent evidence of positive correlations with standard intelligence tests. This suggests that there is a case for understanding emotions being regarded as a candidate intelligence component. For the MSCEIT this issue seems likely to remain disputed due to the use of consensus or expert scoring for this test. As discussed in the introduction, this contrasts markedly with the veridical scoring of standard intelligence tests, with correct answers being derivable on the basis of objective knowledge or logical reasoning, and is a major component of criticisms of EI as a putative new intelligence. In this context, the veridical scoring of the STEU is a particularly interesting development, although it should be noted that the definition of the correct responses to the STEU items operates within the framework of a

specific model of emotional appraisal (Roseman, 2001), rather than the correct answers being objectively verifiable.

In addition to issues with scoring at the scale level, it should be noted that the individual items of the STEM, STEU and MSCEIT differ from intelligence test items in making use of hypothetical scenarios; examples of these were given earlier for the STEM and STEU, and the item content of the MSCEIT Using, Understanding and Managing scales is also based on this approach. In an intelligence test item the respondent has to solve an actual problem, whereas the typical format for an ability EI test item is to be asked what you/another *should* do in a given situation, or what the best action *would* be. The items are thus concerned with the respondent's view of how to identify and deal with information and challenges in the emotional domain, but do not directly assess the capabilities (e.g. managing emotions) encapsulated in the hypothetical scenarios presented in the items.

Given the sparse correlation patterns of ability EI with intelligence tests found in this and other studies, the similar magnitudes of test/test correlations with and without intelligence scores controlled for, the lack of veridically scored tests (with the STEU scoring procedure only satisfying a weak version of veridical scoring) and the problematic features of the item formats currently used in ability EI tests, a strong case cannot currently be made for EI as a new intelligence. The results for tests of emotional understanding are suggestive and should be followed up in future work. The correlational results are inconsistent with EI being a component of the positive manifold of intelligence tests, although this finding should be further tested with a multidimensional intelligence test battery, which would provide a more robust representation of the *g*-factor than was attainable with the two tests used in the present study. The lack of conventional intelligence properties for other EI components does

not of course preclude the continued study of their associations with criterion variables such as well-being and success which can be theoretically and empirically linked to emotion-related capabilities.

The development of the STEM and STEU is undoubtedly positive in terms of the availability of these tests of ability EI and their scoring keys to researchers, and the results on these tests presented here and by MacCann and Roberts (2008) provide evidence for their validity. The internal reliabilities of these tests were however found to be lower than desirable and, in the case of the STEU, an attempt to improve the test's reliability using item analysis was not particularly successful. Comparing the internal reliabilities for the STEM and STEU in this study (.67, .48) with previously reported values in the two studies of MacCann and Roberts (2008) (STEM .68, .61 STEU .71, .43) and for the MSCEIT Managing and Understanding branches in the present study (.66 for both), shows a reasonably consistent picture for the STEM, both between studies and in comparison to MSCEIT Managing in the same sample. The STEU internal reliability values in two out of the three studies are low, and it fares less well than MSCEIT Understanding in the same sample. These comparisons suggest that more work is needed on developing items for the STEU.

A feature of the current generation of ability EI measures (which may be an indicator of a more fundamental issue underlying the reliability problems discussed above) is the focus on tests which assess EI components such as understanding or managing emotions across a range of emotions; this is the case for the MSCEIT, STEM and STEU. This design feature assumes that there is no significant intra-individual variability in the processing of specific emotions. It however seems likely that, for example, some people may be good at understanding and managing joy and elation but less capable with regard to sadness and distress, or vice versa. For multi-

emotion EI tests, the existence of a variety of such profiles would mean that the test was not unidimensional and would reduce item-item correlations and hence test reliability. Given the central status of at least partially separable positive emotion/reward sensitivity and negative emotion/punishment sensitivity systems in current personality models (e.g. Gomez & Cooper, 2008), and the evidence for the existence of distinct neural systems underlying different emotions (Panksepp, 1998), the creation and study both of separate ability EI tests covering positive and negative emotions, and of tests specialised to single basic emotions would seem to be appropriate. (This issue also highlights a limitation of the present study – only negative emotions were used in the emotion perception tests, so the issue of differential performance on positive and negative emotion perception tasks could not be examined.) The creation and use of ability EI tests specialised to positive and negative emotions and to single emotions would allow the empirical examination of the correlations between tests relating to the processing of different emotions, leading both to a better understanding of reliability issues in ability EI assessment and the possibility of examining whether a general factor underlying performance in a given type of test (e.g. of emotional understanding) across different emotions exists. The latter would appear to be a key issue in the study of ability EI which has been neglected to date. In the absence of data, a general point can be made in this context, which is that the existence of individual profiles of relative strengths and weaknesses in relation to different emotions would not necessarily undermine the idea of an “emotional g-factor” as envisaged in current work of ability EI. A possible finding is one analogous to that for intelligence, where population correlations amongst different types of test (e.g. verbal, mathematical) form a positive manifold, but varying individual profiles (for example showing high verbal ability and relatively



lower mathematical ability) are found. If however scores on EI tests relating to different emotions were uncorrelated or only very weakly correlated at the population level, this would indicate the non-existence of an underlying general factor.

In considering the development of a wider range of ability EI measures, another possible direction which should be explored is that of tests which assess rapid (not necessarily conscious) processing of emotional information, as in the emotional IT task used in the present study, and in tests of social perception which require processing of subtle non-verbal cues. This approach contrasts with that adopted in current ability EI tests (reflecting the dominant ability EI model) which rely on the respondent making use of explicit conscious knowledge of emotions in response to test items. The development of models of EI which relate to implicit as well as explicit emotional knowledge, and of test formats which are appropriate to measure the latter is desirable. Related to this is a need for tests which assess the speed or efficiency with which novel emotional information is processed. These attributes are likely to be more closely related to a fluid component of EI than to the crystallised emotion knowledge which is the focus of current EI research.

Another key issue in ability EI measurement is the question of how ability EI should be modelled. The STEM, STEU and MSCEIT are all based on the Mayer-Salovey-Caruso four-branch model of EI, which currently dominates research and debate on ability EI. This leaves open the question of whether this model is the best or most complete one for ability EI. Examination of its structure indicates that some aspects of EI are not being fully assessed by the MSCEIT. Examples are lower-level facets currently not assessed separately within the four-branch model such as empathy, stress management and emotional expressivity. In addition, when formulating hypotheses about the associations of EI with real-world outcomes, a

separation into the superordinate dimensions of inter- and intrapersonal EI is often theoretically useful, corresponding to two distinct ways in which emotional abilities can be deployed. Thus expanding the current model used both upwards to these two broad domains, and downwards to more specific facets of EI is clearly desirable.

## Notes

<sup>1</sup>Mixed models of EI, not specifically considered here, include both ability and personality components. In the present work the definitions of trait and ability EI are aligned to the method of measurement used in their operationalization (Petrides & Furnham, 2003), so self-report measures associated with an underlying mixed model are considered to be trait EI measures.

<sup>2</sup>The items were: STEM 4, 6, 7,8, 9,10,11,13,16,19,21,22,23,24,26,27, 28, 29,31,32,34, 35, 36,37,40,41,42, 43, 44; STEU 2,3,5,6,8,9,10, 11, 12, 13, 15,16,17,19,20, 21, 22,24,25,26, 27, 29,31,32,33, 38, 39,40,41,42

<sup>3</sup>Tables 2 and 3 present correlations relating to the study hypotheses. The full correlation matrix amongst all study variables is available from the author.

## References

- Austin, E. J. (2004). An investigation of the relationship between trait emotional intelligence and emotional task performance. *Personality and Individual Differences*, 36, 1855-1864.
- Austin, E. J. (2005). Emotional intelligence and emotional information processing. *Personality and Individual Differences*, 39, 403-414.
- Austin, E. J., Farrelly, D., Black, C., & Moore, H. (2007). Emotional intelligence, Machiavellianism and emotional manipulation: Does EI have a dark side? *Personality and Individual Differences*, 43, 179-189.
- Austin, E. J., Saklofske, D. H., Huang, S. H. S., & McKenney, D. (2004). Measurement of emotional intelligence: testing and cross-validating a modified version of Schutte *et al*'s (1998) measure. *Personality and Individual Differences*, 36, 555-562.
- Bar-On, R. (1997). *Bar-On Emotional Quotient Inventory (EQ-i): technical manual*. Toronto, Canada: Multi-Health Systems.
- 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.
- Barbuto, J. E., & Burbach, M. E. (2006). The emotional intelligence of transformational leaders: A field study of elected officials. *Journal of Social Psychology*, 146, 51-64.
- Brackett, M., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin*, 29, 1147-1158.

- Carmeli, A., & Josman, Z. E. (2006). The relationship among emotional intelligence, task performance, and organizational citizenship behaviors. *Human Performance*, 19, 403-419.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. Cambridge, MA: Cambridge University Press.
- Ciarrochi, J. V., Chan, A. Y., & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences*, 28, 539-561.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Farrelly, D., & Austin, E. J. (2007). Ability EI as an intelligence? Associations of the MSCEIT with performance on emotion processing and social tasks and with cognitive ability. *Cognition and Emotion*, 21, 1043-1063.
- Gomez, R., & Cooper, C. (2008). Reinforcement sensitivity theory and mood induction studies. In P. J. Corr (Ed.), *The reinforcement sensitivity theory of personality* (pp 291-316). New York: Cambridge University Press.
- Keele, S. M., & Bell, R. C. (2008). The factorial validity of emotional intelligence: An unresolved issue. *Personality and Individual Differences*, 44, 487-500.
- MacCann, C., & Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, 8, 540-551.
- MacCann, C., Roberts, R. D., Matthews, G., & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based emotional intelligence (EI) tests. *Personality and Individual Differences*, 36, 645-662.

- Matthews, G., Zeidner, M., & Roberts, R. D. (2007). Emotional intelligence. Consensus, controversies, and questions. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence. Knowns and unknowns* (pp. 3-46). New York: Oxford University Press.
- Mavroveli, S., Petrides, K. V., Rieffe, C., & Bakker, F. (2007). Trait emotional intelligence, psychological well-being and peer-rated social competence in adolescence. *British Journal of Developmental Psychology*, 25, 263-275.
- Mayer, J. D., Caruso, D.R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 27, 267-298.
- 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., & Caruso, D. R. (2002). *MSCEIT user's manual*. Toronto: Multi-Health Systems.
- Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios (2001). Emotional intelligence as a standard intelligence. *Emotion*, 1, 232-242.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgement tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-740.
- O'Connor, R. M., & Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: Self-report versus ability-based measures. *Personality and Individual Differences*, 35, 1893-1902.
- Palmer, B. R., 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.

Panksepp, J. (1998). *Affective neuroscience*. New York: Oxford University Press.

Parker, J. D. A., Summerfeldt, L. J., Hogan, M. J., & Majeski, S. (2004). Emotional intelligence and academic success: Examining the transition from high school to university. *Personality and Individual Differences*, 36, 163-172.

Petrides, K. V., & Furnham, A. (2003). Trait emotional intelligence: Behavioural validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality*, 17, 39-57.

Petrides, K. V., & Furnham, A. (2006). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology*, 36, 552-569.

Petrides, K. V., Furnham, A., & Mavroveli, S., (2007). Trait emotional intelligence. In G. Matthews, M. Zeidner, & R. D. Roberts (Eds.), *The science of emotional intelligence. Knowns and unknowns* (pp. 151-166). New York: Oxford University Press.

Petrides, K. V., Pita, R., & Kokkinaki, F. (2007). The location of trait emotional intelligence in personality factor space. *British Journal of Psychology*, 98, 273-289.

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 conclusions. *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. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K. R. Scherer & A. Schorr (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 68-91). New York: Oxford University Press.

Rossen, E., Kranzler, J. H., & Algina, J. (2008). Confirmatory factor analysis of the Mayer-Salovey-Caruso Emotional Intelligence Test V 2.0 (MSCEIT). *Personality and Individual Differences*, 44, 1258-1269.

Saklofske, D. H., Austin, E. J., Galloway, J., & Davidson, K. (2007). Individual difference correlates of health-related behaviours: preliminary evidence for links between emotional intelligence and coping. *Personality and Individual Differences*, 42, 491-502.

Saklofske, D. H., Austin, E. J., & Minski, P. S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences*, 34, 707-721.

Schutte, N. S., Malouff, J. M., Hall, L. E., Haggerty, D. J., Cooper J. T., Golden, C. J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence, *Personality and Individual Differences*, 25, 167-177.

Smith, L., Heaven, P. C. L., & Ciarrochi, J. (2008). Trait emotional intelligence, conflict communication patterns, and relationship satisfaction. *Personality and Individual Differences*, 44, 1314-1325.

Stankov, L. (1997). *Gf-Gc Quickie test battery*. Sydney, Australia: School of Psychology.



- Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior*, 65, 71-95.
- Wilhelm, O. (2005). Measures of emotional intelligence: Practice and standards. In R. Schulze & R. D. Roberts (Eds.), *The International Handbook of Emotional Intelligence* (pp 131-154). Göttingen, Germany: Hogrefe.
- Young, A., Perrett, D., Calder, A., Sprengelmeyer, R., & Ekman, P. (2002). *Facial Expressions of Emotion – Stimuli and Tests (FEEST)*. Bury St. Edmunds: Thames Valley Test Company.
- Zeidner, M., & Kaluda, I. (2008). Romantic love. What's emotional intelligence (EI) got to do with it? *Personality and Individual Differences*, 44, 1684-1695.
- Zeidner, M., Roberts, R. D., & Matthews, G. (2008). The science of emotional intelligence. Current consensus and controversies. *European Psychologist*, 13, 64-78
- 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.

Table 1. Descriptive statistics

	<i>N</i>	Mean	Standard deviation
Vocabulary	228	10.77	2.64
Series	223	10.94	2.03
STEM	239	109.45	7.97
STEU	331	28.46	3.75
MSCEIT Perceiving	135	97.70	13.04
Emotions			
MSCEIT Using Emotions	135	96.08	11.37
MSCEIT Understanding	135	100.21	10.06
Emotions			
MSCEIT Managing	135	93.68	7.62
Emotions			
MSCEIT total	135	96.32	10.35
Blends	135	48.99	3.40
Sad IT task	191	73.55	6.94
TEIQue	152	53.02	4.39
SSRI	203	87.07	9.60

Mean error rates for the emotion perception tasks were Blends 9.6%, Sad IT 23.4%.

Table 2<sup>3</sup>. Correlations amongst ability EI measures

	STEM	STEU	Perceive	Use	Understand
STEM	.29*** (.29, .29,.28)				
STEU		.16 (.17, .17.,19)			
Perceive	.13 (.13, .13, .12)		.40*** (.43, .43.,43)		
Use	.25** (.22, .23,.24)	.21* (.22, .19,.21)		.32*** (.29, .29,.30)	
Understand	.40*** (.38, .40,.39)	.44*** (.35,.36,.33)	.28** (.29, .34,.35)		
Manage	.30*** (.31,.30,.30)	.17 (.20, .15,.17)	.31*** (.31, .31,.31)	.20* (.21, .19,.19)	.17* (.21, .20,.21)
MSCEIT	.36*** (.35, .35,.35)	.33*** (.32, .30,.30)	-	-	-

$N = 135$  for correlations with MSCEIT scores,  $N$  range 223-238 for other correlations.

Part-whole correlations of MSCEIT branch scores with total score are omitted.

Values in brackets are correlations with intelligence test scores partialled (in the order Vocabulary, Series, both). The MSCEIT branches Perceiving, Using, Understanding, Managing Emotions are abbreviated to Perceive, Use, Understand, Manage. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

Table 3. Correlations of ability EI tests with emotion perception and intelligence test scores.

	Blends	Sad IT	Vocabulary	Series
STEM	.07 (135)	.13 (91)	.12 (228)	.07 (223)
STEU	.30*** (135)	.21** (183)	.32*** (227)	.10 (223)
Perceive	.08 (135)	-	-.08 (127)	-.19* (123)
Use	.00 (135)	-	.02 (127)	.01 (123)
Understand	.18* (135)	-	.23* (127)	.25** (123)
Manage	.17 (135)	-	-.03 (127)	-.09 (123)
MSCEIT total	.14 (135)	-	.02 (127)	-.06 (123)

The Sad IT task was scored as total number of correct identifications; this task was not used in the MSCEIT study. Sample size for each correlation is in brackets. The MSCEIT branches Perceiving, Using, Understanding, Managing Emotions are abbreviated to Perceive, Use, Understand, Manage. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .