Psychological Topics 20 (2011), 3, 393-412

Original scientific paper – UDC – 159.942.972

Emotional Intelligence as Assessed by Situational Judgment and Emotion Recognition Tests: Building the Nomological Net

Carolyn MacCann

University of Sydney, Australia

Nicola Pearce

University of Cardiff;
University of Sydney, Australia

Richard D. Roberts

Educational Testing Service,
Princeton, US

Abstract

Recent research on emotion recognition ability (ERA) suggests that the capacity to process emotional information may differ for disparate emotions. However, little research has examined whether this findings holds for emotional understanding and emotion management, as well as emotion recognition. Moreover, little research has examined whether the abilities to recognize emotions, understand emotions, and manage emotions form a distinct emotional intelligence (EI) construct that is independent from traditional cognitive ability factors. The current study addressed these issues. Participants (*N*=118) completed two ERA measures, two situational judgment tests assessing emotional understanding and emotion management, and three cognitive ability tests. Exploratory and confirmatory factor analyses of both the understanding and management item parcels showed that a three-factor model relating to fear, sadness, and anger content was a better fit than a one-factor model, supporting an emotion-specific view of EI. In addition, an EI factor composed of emotion recognition, emotional understanding, and emotion management was distinct from a cognitive ability factor composed of a matrices task, general knowledge test, and reading comprehension task. Results are discussed in terms of their potential implications for theory and practice, as well as the integration of EI research with known models of cognitive ability.

Keywords: emotional intelligence (EI), emotion recognition ability (ERA), Situational Test of Emotion Management (STEM), Situational Test of Emotional Understanding (STEU), factor analysis

✉ Carolyn MacCann, School of Psychology, University of Sydney, Sydney, NSW, 2007, Australia. E-mail: *carolyn.maccann@sydney.edu.au*

The authors would like to thank the following individuals for assistance with the preparation of this paper: Denis Bratko, Jeremy Burrus, Beryl Hesketh, Patrick Kyllonen, Mary Lucas, John Sabitini, and Vladimir Taksic.

This study has two primary goals that both aim to investigate the nature of emotional intelligence (EI). First, we examine whether EI is emotion-specific, with distinct components relating to the processing of fear, anger, and sadness-related stimuli. Previous studies have examined this issue for the lower branches of EI (i.e., perception/recognition of emotions; Schlegel, Grandjean, & Scherer, 2012). The current study builds on this research by attempting to replicate such findings for the higher branches of EI (i.e., understanding and managing emotions). Second, we will determine whether different ways of conceptualizing and measuring EI can collectively form one coherent EI construct that is independent from known cognitive abilities. We focus on two distinct measurement paradigms in EI: (a) tasks assessing the higher branches of EI through text-based situational judgments (e.g., Freudenthaler & Neubauer, 2007; MacCann & Roberts, 2008); and (b) tasks assessing emotion recognition ability (ERA). There is fast-growing acceptance of ERA tests as an alternative way to conceptualise and measure EI (e.g., Austin, 2004; Bänziger, Grandjean, & Scherer, 2009; Roberts et al., 2006). Determining whether ERA and text-based EI measures measure the same construct, and whether this construct can be subsumed under existing theories of intelligence can therefore help inform EI research through integration with intelligence and emotions research. These issues are outlined in further detail below.

Emotional Intelligence: Definition and Background

Briefly, EI is generally described as four related sets of abilities, ranging in complexity from relatively simple emotion processing skills (emotion perception) through to more complex and socially contextualized abilities such as understanding the relationship between different emotions, time courses and situations (emotional understanding) and the ability to manage or regulate the experience or generation of emotions (emotion management) (e.g., Mayer, Roberts, & Barsade, 2008; Mayer, Salovey, Caruso, & Sitarenios, 2001). These four branches of EI are thought to have a roughly hierarchical order, where abilities in the higher branches depend on abilities in the lower branches. For example, the ability to manage one’s emotions requires firstly that one recognize emotions.

To date, this four-branch hierarchical model of EI has been operationalized primarily with the Mayer-Salovey-Caruso Emotional Intelligence Test battery (MSCEIT; Mayer et al., 2001). The MSCEIT contains two tasks for each of the four branches: Emotion Perception, Emotion Facilitation of Thought, Emotional Understanding, and Emotion Management. Collectively, the two highest branches of EI (emotional understanding and emotion management) are known as Strategic EI, and the two lower branches are known as Experiential EI. Several alternative assessments of EI have been developed in recent years (e.g., Bänziger et al., 2009; Freudenthaler & Neubauer, 2007; MacCann & Roberts, 2008; Mayer & Geher, 1996). However, none of the studies examining alternatives to the MSCEIT cover both the Experiential and Strategic areas of EI. One of the goals of the current research is to integrate two different measurement paradigms which are gaining acceptance as useful measures of EI: Situational judgment tests as assessments of Strategic EI, and Emotion Recognition Ability (ERA) tasks as assessments of emotion perception. If both of these types of tasks (which use different measurement methods and focus on different branches) formed a single factor then this would constitute compelling evidence for a higher-order emotional intelligence construct that is not simply a method factor.

The Specificity of Emotions and Emotional Intelligence

O'Sullivan and Ekman (2004) suggested that the processing involved in emotional tasks might be different for different emotions. Recent research on emotion recognition ability supports this view, at least for the lower branches such as emotion perception. Age deficits in emotion recognition differentially affect different emotions, with the strongest deficits noted for sadness (Mill, Allik, Realo, & Valk, 2009). Suzuki, Hoshino, and Shigemasu (2010) demonstrate that recognition of positive emotions may be a distinct process from recognition of negative emotions. Schlegel et al.'s research (2012) suggests that emotion recognition may differ for distinct emotion families. For example, recognition of irritation and anger may be distinct from recognition of sadness and despair, which may be distinct from recognition of joy and happiness. That is, particular patterns of strengths and deficits in EI may relate to the particular kind of emotion processed.

Although findings of emotion-specificity are beginning to emerge for the lower branches of EI, there has been little research on this issue for emotional understanding and management (i.e., Strategic EI). To the authors' knowledge, no one has yet addressed the question of whether Strategic EI might be different for different types of emotions, and indeed, whether this might account for the finding that these tests are often factorially complex and not highly internally consistent. The current research will address this imbalance by examining whether distinct factors relating to fear, anger, and sadness underlie participants' responses to situational judgment tests of emotional understanding and emotion management. Specifically, the current study will use modified versions of the Situational Test of Emotional Understanding (STEU) and the Situational Test of Emotion Management (STEM; see MacCann & Roberts, 2008). Variations of the instruments were developed in order that each item from both tests was clearly related to fear, anger, or sadness.

In the current study, item parcels representing three different emotions (fear, anger, and sadness) were created from both the STEM and STEU to investigate whether these parcels load on distinct emotion-specific factors or on one general factor. This direct comparison of three-factor emotion specific models between the STEM and the STEU was made possible by creating additional STEU items representing sadness, anger, and fear, as a model of the original fourteen emotions covered by the STEU would have been too complex to examine empirically. For both the STEU and the STEM, nine item parcels (with three parcels for each of the emotions: anger, fear, and sadness) were used as indicators in a series of structural analyses. In Model 1, anger, fear, and sadness indicators form three separate factors. Model 2 is nested within Model 1, where all nine indicators load on the same factor, such that the increase in model fit for positing emotion specificity can be statistically evaluated. Based on suggestions of emotion specificity in the emotions and EI literature, it is expected that emotion specific models of emotional understanding and emotion management will provide a better fit to the data than one-factor general models.

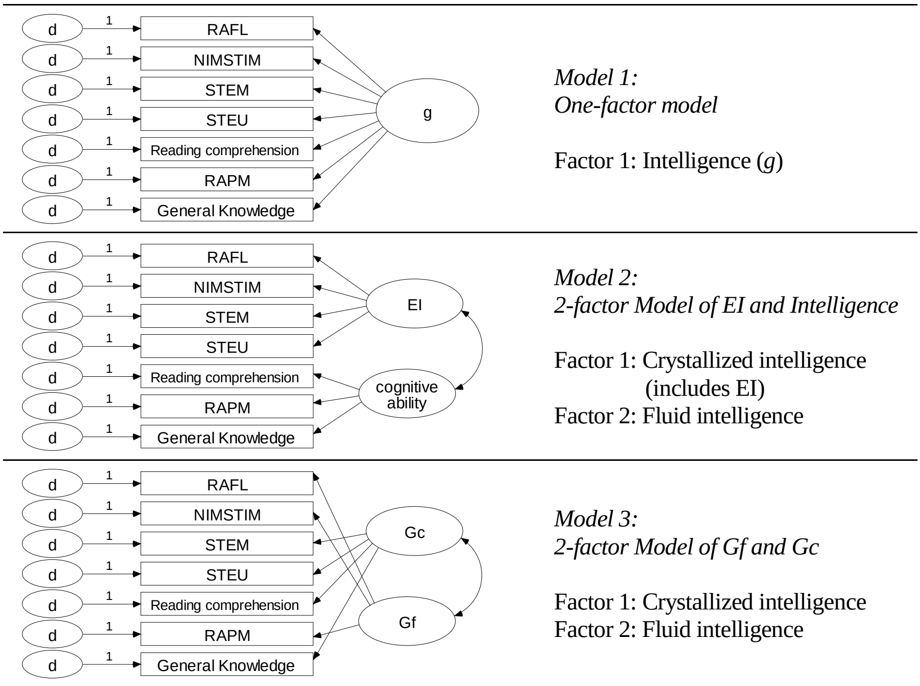
Is there a Higher-Order EI Factor Distinct from Methods?

Research on emotional abilities is increasingly turning to emotion recognition ability tests as a viable way to assess individual differences in the ability to recognize or perceive emotions (Austin, 2004; Roberts et al., 2006). These tests come from the emotions research tradition rather than EI field and assess how well people can recognise emotions in different stimuli. They are scored by standards based on emotions research rather than by consensus, and are conceptually linked to the Emotion Perception branch of the four branch hierarchical model of EI, although operationalised somewhat differently to the MEIS and MSCEIT Perception tests (Roberts et al., 2006).

If EI is indeed a distinct type of intelligence, then these ERA tests should relate strongly to other components and measures of EI (i.e., Strategic EI) and this relationship should be stronger than the relationship between EI and other intelligences. Roberts et al. (2006) found small to moderate relationships between Strategic EI (as measured by the MSCEIT) and two ERA measures: the JACBART (which assesses recognition of emotions in facial expressions) and the Vocal-I (which assesses recognition of emotions in vocal expressions; Scherer, Banse, & Wallbott, 2001). However, in a factor analysis ERA tasks loaded on fluid (Gf) and crystallized (Gc) factors rather than Experiential or Strategic EI factors defined by MSCEIT tasks. Both ERA tasks loaded on Gf and the Vocal-I also loaded on Gc. This evidence contradicts the idea that EI tasks will collectively form a distinct factor separable from other cognitive abilities. Roberts et al. (2006) suggests that ERA is relatively distinct from Experiential EI (despite a strong conceptual relationship), and instead relates more strongly to Gf and Gc. In conjunction with results from MacCann (2010), where Strategic EI was strongly related to Gc, this suggests that ERA and EI may form parts of Gf and Gc respectively, rather than both forming a single EI construct.

Based on the results of this research, the current study will readdress the issue of whether ERA and EI form a single construct by comparing three separate models (shown in Figure 1) with exploratory and confirmatory factor analysis. Model 1 is a one-factor model in which Gf, Gc, Strategic EI, and two ERA markers all load on general intelligence (*g*). Model 2 distinguishes between EI and other forms of cognitive ability (a general EI versus *g* model). Model 3 investigates whether ERA and EI form parts of Gf and Gc respectively by loading ERA markers onto a Gf factor and Strategic EI on a Gc factor (with the Vocal-I cross-loading, as was the case in the Roberts et al. [2006] study). ERA will be measured in both vocal expressions (with the Vocal-I, as in the Roberts et al. [2006] investigation) and facial expressions (with a measure derived from the Nim Stim set of facial expressions)[1].

Figure 1. *The Structure of Three Competing Models of Intelligence (see text for details)*



[1] Development of the MacBrain Face Stimulus Set was overseen by Nim Tottenham and supported by the John D. and Catherine T. MacArthur Foundation Research Network on Early Experience and Brain Development. Please contact Nim Tottenham at tott0006@tc.umn.edu for more information concerning the stimulus set.

Summary of Hypotheses

Hypothesis 1: An emotion-specific model of the STEU items consisting of anger-understanding, fear-understanding, and sadness-understanding will provide a better fit to the data than a one-factor model.

Hypothesis 2: An emotion-specific model of the STEM items consisting of anger-management, fear-management, and sadness-management will provide a better fit to the data than a one-factor model.

Hypothesis 3: It is expected that ERA and Strategic EI will form a distinct EI factor which is separable from traditional cognitive ability. Specifically, a two-factor model (EI and cognitive ability as separate factors) will fit the data better than a one-factor model (where all indicators load on the one cognitive ability factor). This would indicate that EI is distinct from other cognitive abilities. In exploratory research, a third model (where Strategic EI forms part of Gc and ERA forms part of Gf) will also be examined to the to see if a Gf/Gc model can partition the variance in test scores in a meaningful way.

Method

Participants

Participants were 118 first year undergraduate psychology students from the University of Sydney (68 female) participating in this study for course credit. Ages ranged from 18 to 32 with a median of 19 (*M*=19.3, *SD*=1.8). In total, 20 participants were non-native English speakers.

Materials

Emotional Intelligence (EI) Tests

Situational Test of Emotional Understanding (STEU). Participants were given 40 items to complete at their own pace. Of these, 25 items were selected from MacCann and Roberts (2008) for their high item-total correlation (including 9 that assessed sadness-, anger-, and fear-understanding). An additional 15 items were developed to allow sufficient items to assess fear, anger, and sadness content. Thus, 8 items related to sadness, 8 to anger, and 8 to fear. This additional development was undertaken so that an examination of emotion specificity was possible. In this study, the STEU is scored dichotomously according to standards derived from Roseman’s (2001) appraisal theory (for more information on scoring, see MacCann & Roberts, 2008).

Situational Test of Emotion Management (STEM). Participants were given 24 items from the STEM to answer in ratings-based format (8 items related to sadness, 8 to anger, and 8 to fear)[2]. Items were selected from MacCann and Roberts (2008) for their high item-total correlations. In this study, the STEM was scored by standardised expert distance (i.e., the same scoring rubric used for the ratings-based STEM by MacCann and Roberts (2008), as this was the most reliable).

Emotion Recognition Ability (ERA) Tests

Recognition of Affect in a Foreign Language (Vocal-I). The Vocal-I (Scherer et al., 2001) consists of 30 items assessing recognition of emotion in tone-of-voice. In each item, participants listen to a nonsense phrase given twice and then need to select which of five emotions the voice expresses (joy, sadness, anger, fear, or no emotion). The task was computerised and participants were each given a set of headphones. The Vocal-I has been used frequently in cross-cultural research on emotion (Scherer et al., 2001). Internal consistency reliability has been reported at .51 (Roberts et al., 2006).

Nim Stim Faces Test. The Nim Stim Faces Test was constructed specifically for use in the current study, using stimuli from the Nim Stim stimulus set (Tottenham, 2004). Stimuli used in the task were 48 emotion-expressing faces of African American, Caucasian, and Asian appearance (12 for each emotion of sadness, anger, fear, and joy). Each emotional expression was presented with a forward and backwards mask of the same face with a neutral expression. The forward mask was presented for one second before the stimulus face and the backwards mask was left onscreen until participants selected a response from four possible options (sadness, anger, fear, and joy). After a response was selected, the task moved immediately to presentation of the next forward mask. In line with findings from Matsumoto et al. (2000), half of the stimuli faces were presented for 200 ms and half for 100 ms. However, these presentation times were probably too long, as there was 100% accuracy on 6 items (5 of which were joy items) and greater than 95% on 15 more items (7 of which were joy items). These 21 items were excluded before computing total scores, since they would contribute very little variability to this test. The final score was thus derived from 27 items, none of which pertained to joy-recognition.

[2] Due to a computer error, responses to one of the sadness-management items were not recorded, effectively making this a 23-item test.

Intelligence Tests

Reading Comprehension Test. Participants were given a subset of 43 items from Part II of the Co-operative Reading Comprehension Test, Form Y (ACER, 1978) with an 8 minute time limit. Participants were required to read excerpts of text and then answer multiple-choice questions that assessed their understanding of these excerpts. Reading comprehension is a measure of crystallized intelligence (Gc) under Carroll (1993) and a measure of reading and writing (Grw) under McGrew (1997). In the current study, reading comprehension is expected to load on intelligence and Gc factors in structural analyses of these test scores.

Raven's Advanced Progressive Matrices. Participants were given a subset of 20 items from Raven's Advanced Progressive Matrices (Raven, Raven, & Court, 1998) to complete within an 8 minute time limit. Raven's Progressive Matrices is considered a pure measure of fluid intelligence (Gf; Carroll, 1993), although the task also involves some spatial abilities (Gv; e.g., Colom, Escorial, & Rebollo, 2004). In each item, a three by three grid of designs follows a pattern both across and down the grid and participants must select from eight possible responses the pattern that goes in the lower right corner to complete both patterns.

Knowledge. Participants were given set 1 (28 items) from the Knowledge test from the Intelligence Structure Test (I-S-T, Amthauer, Brocke, Leipmann, & Beauducel, 2000). Items assess general knowledge in several domains, and participants must select the best answer from among five possible options. General knowledge is known to be a marker of crystallized intelligence (Gc; Carroll, 1993).

Procedure

Participants volunteered to participate based on a brief description of the experiment. After reading a two-page description of the study, participants signed a consent form indicating their willingness to participate in the experiment. After signing consent forms, participants completed all tests on PC computers. Test order was counter-balanced across four conditions, to ensure that fatigue or boredom did not contribute possible confounds. This study was approved by the Sydney University Human Research Ethics Committee.

Results

Reliability and Descriptive Statistics

Reliability and descriptive statistics (proportions of accurate answers) are shown in Table 1, along with comparisons of male and female performance. The General Knowledge test, STEU, and STEM were appropriately reliable for research purposes, but the reliability of the Vocal-I and the Nim Stim Faces was marginal. Even with the 21 Nim Stim items that showed greater than 95% accuracy excluded, internal consistency reliability was marginal: .56. The Vocal-I had a reliability of .43, and 8 items were removed to improve reliability (using half the sample to select items and the other half to calculate reliability). Group means, however, suggested that many of the scores from the battery of tests were construct valid. Thus, there was a significant male advantage on the Knowledge test, consistent with previous studies (e.g., Ackerman, Bowen, Beier, & Kanfer, 2001; Lynn, Irwing, & Cammock, 2001). There was also a significant female advantage on the STEM, consistent with previous studies (e.g., MacCann, 2010).

Table 1. *Reliability, Descriptive Statistics, and Gender Differences for all Measures*

	All (*N*=118)			Males (*N*=50)		Females (*N*=68)		Gender differences (*t*)
	α	*M*	*SD*	*M*	*SD*	*M*	*SD*	
Reading comprehension	-	.26	.11	.27	.11	.25	.12	0.21
Raven's Matrices	-	.54	.17	.56	.15	.53	.18	0.16
General Knowledge	.64	.58	.13	.65	.11	.54	.12	0.87**
STEU	.70	.71	.11	.70	.11	.72	.11	0.21
STEM	.74	.72	.20	.67	.19	.76	.20	0.44*
Vocal-I	.43	.67	.10	.66	.09	.69	.09	0.32
Vocal-I (short)	.56	.70	.12	.68	.12	.71	.12	0.25
Faces (short)	.67	.88	.11	.87	.11	.88	.10	0.05

Note. Vocal-I = Recognition of Affect in a Foreign Language (all 30 items); Vocal-I (short) = 22 most reliable items from the Vocal-I; Faces (short) = 27 NIMSTIM Faces items with difficulties <.95. Effect size for gender is calculated as Hedge's *g*, with negative values indicating higher scores for females.

*p<.05, **p<.01.

Emotion Specificity of the STEU and STEM

To check whether Understanding and Management constitute interpretable constructs for different emotions, structural analysis of STEU and STEM item parcels were conducted separately for the STEU and STEM. Each analysis compared a three-factor model (where sadness-, anger-, and fear-related items formed three separate factors) to a one-factor model (where sadness-, anger-, and fear-related items all loaded on the same factor).

Descriptive statistics and correlations between the items parcels are provided in Table 2 so that the interested reader can explore the structural analysis further. Parameters for the STEU are shown on the left and lower left triangle and parameters for the STEM are shown on the right and upper right triangle. It can be seen that the correlations are stronger for the STEM than the STEU.

Table 2. *Descriptive Statistics and Correlations between Nine STEU Item Parcels (Left and Lower Left) and Nine STEM Item Parcels (Italicised in Right and Upper Right)*

		Correlations									STEM	
		1	2	3	4	5	6	7	8	9	*M*	*SD*
1. Anger subset 1			*.16*	*.21**	*.18*	*.19**	*.19**	*.10*	*.17*	*.06*	*3.89*	*0.78*
2. Anger subset 2		.17		*.34***	*.08*	*.22**	*.26***	*.17*	*.22**	*.31***	*4.01*	*0.80*
3. Anger subset 3		-.02	.11		*.14*	*.26***	*.27***	*.17*	*.23**	*.36***	*2.47*	*0.85*
4. Fear subset 1		.06	.14	.05		*.36***	*.41***	*.24**	*.19**	*.04*	*4.09*	*0.99*
5. Fear subset 2		.12	.04	.02	.03		*.48***	*.33***	*.33***	*.36***	*3.88*	*1.11*
6. Fear subset 3		.06	.20	.05	.27**	.00		*.19**	*.29***	*.21**	*2.60*	*0.94*
7. Sadness subset 1		.20*	.27*	.12	.17	.15	.04		*.28***	*.24**	*3.96*	*0.80*
8. Sadness subset 2		-.03	.10	.06	.06	.16	.14	.14		*.29***	*2.53*	*0.96*
9. Sadness subset 3		.12	.18	.07	.14	.13	.10	.29**	.29**		*2.69*	*0.86*
STEU	*M*	1.33	2.31	1.65	1.93	2.07	1.49	2.62	1.49	1.89		
	SD	0.86	0.80	0.48	0.68	0.69	0.48	0.55	0.52	0.36		

Note. STEM = Situational Test of Emotion Management ; STEU = Situational Test of Emotional Understanding; Missing data deleted list-wise, N=109.

$^{*}p$<.05, $^{**}p$<.01.

Emotion specificity of the STEU: Structural Analysis. Factor loadings, factor correlations, and fit indices for the structural analysis of STEU item parcels are shown in Table 3. For the analysis of the STEU, the three-factor model fit significantly better than the one-factor model, showing a significant reduction in χ^2 values and a large decrease in the AIC indices. Fit indices were acceptable for the three-factor model (i.e., RMSEA < .06 and both CFI and GFI > .95). However, several of the paths in this model were not significant, and the model overfit the data (χ^2<df). Thus, evidence for emotion-specificity in the Understanding construct is present, but because of overfit especially needs to be interpreted with some degree of caution. Fear and Anger factors were more strongly correlated with each other (r=.74) than with Sadness (r=.58 and r=.47, respectively).

Emotion specificity of the STEM: Structural Analysis. Table 3 also shows the factor loadings, factor correlations, and fit indices for one-factor and three-factor models of the STEM item parcels. An emotion-specific model for Management fit the data better than a general Management factor, according to both χ^2 and AIC differences. However, fit indices were good for the one-factor model as well as the three-factor model (RMSEA < .06 and both CFI > .95). Although the three-factor model was overfit, all paths in the model were significant for the STEM analysis, such that the three-factor model appears to be the best fit to the data. Anger, Fear, and Sadness factors were highly inter-correlated (all correlations were greater than .60) indicating considerable overlap between these constructs.

Table 3. *Factor Loadings, Factor Correlations, and Fit Indices for the Structural Analysis of Strategic EI Emotion-Specific Item Parcels*

	Structural Analysis of the STEU				Structural Analysis of the STEM			
	One-factor (df=27)	Three-factor (df=24)			One-factor (df=27)	Three-factor (df=24)		
		F1 Anger	F2 Fear	F3 Sadness		F1 Anger	F2 Fear	F3 Sadness
1. Anger 1	.30**	.28$^{\#}$			.30**	.33**		
2. Anger 2	.38**	.54**			.42**	.54**		
3. Anger 3	.43**	.19$^{\#}$			.47**	.62**		
4. Fear 1	.60**		.56**		.44**		.55**	
5. Fear 2	.54**		.35**		.50**		.55**	
6. Fear 3	.55**		.56**		.48**		.46**	
7. Sadness 1	.46**			.52**	.45**			.50**
8. Sadness 2	.21$^{\#}$			.12$^{\#}$	.70**			.74**
9. Sadness 3	.59**			.48$^{\#}$	.63**			.67**
Correlations								
Fear (F2)		.74				.79		
Sadness (F3)		.58	.47			.61	.76	
Fit Indices								
χ^2	47.51		16.89		33.07		20.39	
$\Delta\chi^2$			30.63*				12.68**	
GFI	.92		.97		.94		.96	
CFI	.84		1.00		.96		1.00	
RMSEA	.08		.00		.04		.00	
AIC	83.51		58.89		69.07		62.39	

Note. STEU = Situational Test of Emotional Understanding; STEM = Situational Test of Emotion Management; GFI = Goodness-of-fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error Approximation; AIC = Akaike Information Criterion; All RMSEAs are not significantly greater than .05.

$^{\#}$These paths not significant; $^{*}p$<.05, $^{**}p$<.01.

Relationships between Strategic EI, ERA, Gf, and Gc

Correlations between intelligence (Gf and Gc indicators), Strategic EI, and ERA are shown in Table 4. All correlations between intelligence, EI, and ERA are positive, indicating that EI and ERA are measuring intelligence to some extent. STEU and STEM scores were moderately related, consistent with results from MacCann and Roberts (2008).

Table 4. *Correlations between Cognitive Ability (Gf and Gc), Strategic EI, and Emotion Recognition Ability Measures*

	1	2	3	4	5	6
1. Reading comprehension						
2. Raven's Matrices	$.33^{**}$					
3. General Knowledge	$.43^{**}$	$.35^{**}$				
4. STEU	$.35^{**}$	$.35^{**}$	$.29^{**}$			
5. STEM	.17	$.20^{**}$	.16	$.43^{**}$		
6. Vocal-I	$.30^{**}$	.13	.06	$.31^{**}$	$.27^{**}$	
7. Faces	.13	$.28^{**}$	.07	$.20^{**}$	.13	.03

Note. STEU = Situational Test of Emotional Understanding; STEM = Situational Test of Emotion Management; Vocal-I = Recognition of Affect in a Foreign Language; Faces = Nim Stim Faces Test; Missing data was deleted list-wise, N=109.

$^{**}p<.01$.

The ERA measures were unrelated to each other. Even examining ERA separately for each emotion (i.e., sadness-recognition, fear-recognition, or anger-recognition in both faces and voices) produced no significant relationships between Faces and Vocal-I. This is similar to Roberts et al. (2006) findings that the Vocal-I was only weakly related to a different facial expression recognition task (r=.22). Correlates of the emotion-specific factors of the STEU and STEM were not markedly different for the different factors, except for the Nim Stim Faces' relationship to the STEU. Only the sadness factor of the STEU related to the Nim Stim Faces Test. However, given that 90 possible correlations were examined, this could easily be a Type I error rather than a genuine difference, particularly given the low reliability of the Nim Stim Faces Test.

Locating EI in the Structure of Intelligence: Exploratory and Confirmatory Factor Analysis

Exploratory factor analysis was run on the three cognitive ability, two Strategic EI, and two ERA tests to inform the confirmatory factor analyses planned and presented in Table 5. Although three eigenvalues were greater than 1, a three-factor solution resulted in a singlet for reading comprehension, so a two-factor solution was instead extracted (using Maximum Likelihood and oblique rotation, since factors were assumed to be correlated). Factor loadings are shown in Table 5. The pattern of loadings suggests that the two factors (correlated at *r*=.52) resemble an EI factor composed of STEU, STEM, and Vocal-I scores, and an intelligence factor composed of Reading Comprehension, Raven's Matrices, and Knowledge Test scores. Faces did not load saliently on either factor, although it did have a stronger loading on EI than it did on the putative intelligence factor.

Table 5. *Exploratory and Confirmatory Factor Analysis of Cognitive Ability, Strategic EI and ERA Markers*

	Exploratory Factor Analysis (*df*=8)		Confirmatory Factor Analyses (CFA)				
			Model 1 (*df*=14)	Model 2 (*df*=13)		Model 3 (*df*=13)	
			g	*g*	EI	Gc	Gf
Reading Comprehension	.13	**.55**	.58*	.65*		.59*	
Raven’s Matrices	.17	**.45**	.50*	.57*			.80*
General Knowledge	-.13	**.77**	.55*	.61*		.50*	
STEU	**.65**	.16	.68*		.81*	.69*	
STEM	**.61**	-.06	.47*		.53*	.47*	
Vocal-I	**.46**	-.03	.39*		.40*	.40*	
Faces	.19	.11	.28*		.25*		.35*
Fit Indices							
CFI			.89		.98		.91
GFI			.94		.96		.94
RMSEA			.09		.03		.08
AIC			53.18		44.70		51.76
χ^2	10.52		25.18		14.70		21.76

Note. Factor loadings > .30 are in boldface. g = General Intelligence; EI = Emotional Intelligence; Gc = Crystallized Intelligence; Gf = Fluid Intelligence; STEU = Situational Test of Emotional Understanding; STEM = Situational Test of Emotion Management; Vocal-I = Recognition of Affect in a Foreign Language; Faces = Nim Stim Faces Test; CFI = Comparative Fit Index; GFI = Goodness-of-fit Index; RMSEA = Root Mean Square Error Approximation; AIC = Akaike Information Criterion; Missing data was deleted list-wise, *N*=109.

*p<.05.

This solution was checked with confirmatory factor analysis, and two other competing structures (described in Figure 1) were also examined. In Model 1, all 7 indicators load on one factor; whereas Model 2 distinguishes between an EI factor (STEU, STEM, Vocal-I, and Faces) and an Intelligence factor (Raven’s Matrices, Reading Comprehension and Knowledge) (correlated at *r*=.66). Model 3 delineates the indicators into Gf (Raven’s Matrices, Nim Stim Faces, and Vocal-I) and Gc (Reading Comprehension, Knowledge, Vocal-I, STEU, and STEM) (correlated at *r*=.66). Discrepancy matrices and modification indices suggested that the Vocal-I loaded only on Gc and not on Gf, so the Gf cross-loading was excluded. Factor loadings and inter-correlations as well as selected indices of model fit are shown in Table 5 for all three models.

Fit indices are only good for Model 2, with CFI and GFI > .95 and RMSEA < .05. Fit indices for Model 1 were unacceptable, and indices for Model 3 were marginally acceptable (RMSEA not significantly greater than .05, and both CFI and GFI > .90 but not > .95). Model 2 (EI versus intelligence) fit the data significantly better than Model 1 (the one-factor model; $\Delta\chi^2(1)=10.49$, $p<.001$; ΔAIC=8.49), indicating a distinction between EI and other cognitive abilities. Model 3 (Gf and Gc factors) did *not* show a significant improvement in fit over a one-factor model ($\Delta\chi^2(1)=3.42$, $p>.05$; ΔAIC=1.42). In conjunction with the exploratory factor analysis supporting Model 2, these analyses suggest that Strategic EI and ERA abilities may cohere to form an EI construct that is distinguishable from other forms of intelligence. However, the correlation between EI and cognitive ability factors was quite high (*r*=.66), echoing the high correlation between Strategic EI and Gc in MacCann (2010).

Separate analyses of ERA indicators (Vocal-I and Faces) with intelligence and Strategic EI indicators (STEU and STEM) with intelligence proved informative here. In a one-factor model of Strategic EI and intelligence, modification indices and discrepancy matrices suggested that STEU and STEM scores might form a separate factor (i.e., factors were Strategic EI and intelligence, correlated at *r*=.59). In a one-factor model of ERA and intelligence, modification indices and discrepancy matrices suggested that Raven's Matrices and Faces should form a separate factor (i.e., factors were Gf and Gc, correlated at *r*=.53). That is, ERA scores fit into the Gf/Gc theory of intelligence (with auditory judgments of tone-of-voice loading with Gc and the visual inspection time task of facial expression recognition loading with Gf) but Strategic EI scores formed a different factor to intelligence.

Discussion

Results from this study highlight the ways that scores on emotion-related tests might segment differently according to the specific emotions dealt with as well as the processes involved in such tests. Relationships to Gf and Gc seem to differ depending on the processes involved in the tasks, with inspection time paradigms (as in facial expression recognition tasks) relating to Gf, and auditory and knowledge-based processes relating to Gc. These issues are further discussed below.

Emotion Specificity of the STEU and STEM

The possibility that different constructs within EI might differ for different emotions was supported by analyses of the STEU and STEM. For both tests, emotion-specific models positing different factors for anger, fear, and sadness provided a better fit to the data than one Understanding or Management factor (although emotion-specific factors were strongly correlated with each other in both cases). There are three possible corollaries that follow from this emotion specificity. First, tests of EI need to include a range of different emotions in their item content for tests to be representative of the entire content domain of EI. Second, there may be differential prediction of criteria for different emotions, which has implications for applications in selection as well as program evaluation. Third, different training programs may be differentially effective for different emotions, implying that intervention designs might need to be tailored specifically to the underlying deficit (e.g., fear-, anger- or sadness-management). These points are elaborated below.

Emotion Specificity: Implications for Test Development

One important form of validity evidence for a psychological test is that its content be both relevant and representative of the domain of interest (AERA, APA, & NCME, 1999). If the processing of different emotions such as fear, anger, and sadness constitute distinctively different aspects of EI, then one of the requirements for content representation is that tests include content representative of each of these emotions. For example, an emotion management test in which 90% of the content relates to dealing with anger or conflict would *not* be representative of emotion management, but only of the sub-construct of anger management. Thus, when developing EI instruments, a range of different emotions should be included in the content, otherwise the tests may not be representative of the EI domain.

Emotion Specificity: Implications for Prediction of Criteria

If understanding or managing emotions is specific to different types of emotions, then these more specific constructs may turn out to have greater predictive power and utility than the omnibus constructs of emotional understanding and management. For example, outcomes relating to rumination and anxiety might feasibly be predicted by fear-management but not by anger-management or sadness management. Similarly, conduct disorders may relate to anger-management but not to fear- or sadness-management. Not only should emotion-specific EI scores be matched to emotion-specific criteria, but these may even differ across the four branches of EI. For example, fear-management but not fear-understanding may predict job performance in jobs involving high elements of risk. This study did not include measures of well being, life success, or mental health, which different emotions might differentially predict. Future research might examine whether the ability to understand or manage particular emotions can predict different varieties of useful life outcomes.

Emotion Specificity: Implications for Interventions and Training

A further implication of this emotion specificity is the need for targeted interventions. For example, if an individual has deficits in anger-management but not fear- or sadness-management then the most effective intervention might focus on anger. This might involve strategies such as re-training misattributions of intentionality, or developing communication methods that are assertive rather than aggressive. Such strategies are effective for anger management (e.g., Acton & During, 1992) but are arguably not as relevant for fear or sadness management.

Relationships among Cognitive Tasks (Intelligence and EI)

There was a very high degree of relationship between STEU and STEM indicating that they are clearly measure the same construct. Both STEU and STEM also related to the ERA measures, forming an EI factor with these assessments that was distinct from "non-emotional" intelligence. It was surprising however, that Vocal-I and Nim Stim Faces did *not* relate to each other. Part of this may be a problem with the lack of variability in the Nim Stim Faces but, equally, emotion recognition ability may be quite distinct for different modalities. Extensions of the Gf/Gc model of intelligence to different sensory modalities support this interpretation: visual, auditory, kinaesthetic, and olfactory abilities (Gv, Ga, Gk, and Go) form factors distinct from the memory or reasoning processes involved in these tasks (see e.g., Danthiir, Roberts, Pallier, & Stankov, 2001; Horn & Stankov, 1982; Stankov & Horn, 1980; Stankov, Seizova-Cajic, & Roberts, 2001). In addition, Roberts et al. (2006) found that a visual ERA measure (a facial expression recognition task based on the inspection time paradigm) was unrelated to the MSCEIT's Faces test. Although stimuli in the tasks were virtually identical, different abilities were required when an inspection time paradigm was used.

If the *emotional content* of ERA items predominates the processes used to perform the tasks (rather than the *sensory modality* or *method of presentation*) then an *emotion recognition ability* may be posited. However, if other factors such as sensory modality or method of presentation predominate, then tasks form part of other known abilities: the emotional content is only a surface characteristic of the task. From the evidence so far, the latter appears to be the case for the ERA measures. The auditory task (Vocal-I) was more strongly related to Gc (as Ga tasks generally are – Horn and Stankov [1982] found a Ga/Gc factor correlation of r=.54) and the inspection time task (Nim Stim Faces) was more strongly related to Gf (as inspection time measures tend to be; see Grudnik & Kranzler, 2001) than either were to each other. However, when Strategic EI was added to the mix, the shared variance due to emotional content was sufficient for emotional intelligence to form a factor separate to "non-emotional" intelligence, although factors were highly correlated (r=.73).

The strong correlation between EI and intelligence echoes the strong correlation between Strategic EI and Gc obtained in MacCann (2010). With correlations in this range, the distinctiveness of EI from intelligence needs to be evaluated functionally as well as structurally. Based on Horn and Noll (1997), MacCann (2010) suggested grounds for determining EI's distinction from intelligence might include age trends, neurological processes, correlates, and group differences. There is evidence for one of these conditions in the current study (i.e., different gender differences in the EI measures to that found for the cognitive ability indicators). Women were superior at emotion management whereas men were superior at general knowledge (as expected from prior findings, Ackerman et al., 2001; Ciarrochi, Chan, & Caputi, 2000; Day & Carroll, 2004; Mayer, Caruso, & Salovey, 2000).

Limitations of This Study

The samples used for structural equation modelling in this study were rather small for such a purpose (N=109). For this reason, further research with larger samples is needed to confirm results showing the emotion specificity of Understanding and Management, as well as the distinctiveness of EI from other mental abilities. In addition, the emotion-specific models showed quite high correlations between factors, and some evidence that models overfit the data, such that accepting the structure at the one-factor level would not be unreasonable. Further clarification of this structural issue with a larger sample would be ideal.

Further limitations of this study relate to the markers of ERA and of intelligence. The Nim Stim was an experimental measure created for the purpose of this study, and clearly suffered from ceiling effects. The Nim Stim was developed in preference to using the existing JACBART in order to examine differences in presentation latency, as well as to include a more limited range of emotions (i.e., to exclude contempt). In retrospect, it is clear that using a well-developed existing measure such as the JACBART would have been better practice. Two of the intelligence markers (Raven's Matrices and Reading Comprehension) were given within strict time limits, meaning that speed as well as accuracy would have contributed to participants' scores. It is possible that this speed factor may have contributed to the results of the confirmatory factor analysis (i.e., that the separation of cognitive ability from EI may in fact have been the separation of speeded tests from non-speeded tests). Replication of these results using un-timed tests would appear useful.

General Conclusions

Results from this study inform research on EI in two important ways. Firstly, results suggest that the component abilities of EI may be distinct for different types of emotions. The implications of this emotion-specificity are that EI tests need to include content that represent a range of emotions, and that there is a further possible need to match both interventions and potential criteria to particular emotional content. Second, ERA assessments and situational measures of Strategic EI *do* seem to be parts of the same overall construct of EI, although they may involve different cognitive processes. Thus, results from the current study show that the structure of EI is both more simple and more complex than previous research might suggest: More simple in that different varieties of assessment all measure the same EI factor, but more complex in that such a structure may be meaningfully delineated into different parts based on the emotions involved.

References

ACER (1978). *Co-operative Reading Comprehension Test: Form Y*. Hawthorn, Victoria, Australia: ACER.

Ackerman, P.L., Bowen, K.R., Beier, M.E., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology, 93*, 797-825.

Acton, R.G., & During, S.M. (1992). Preliminary results of aggression management training for aggressive parents. *Journal of Interpersonal Violence*, *7*, 410-417.

AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington DC: AERA.

Amthauer, R., Brocke, B., Leipmann, D., & Beauducel, A. (2000). *Intelligenz-struktur-test*. Gottingen: Hogrefe.

Austin, E.J. (2004). An investigation of the relationship between trait emotional intelligence and emotional task performance. *Personality and Individual Differences, 36*, 1855-1864.

Bänziger, T., Grandjean, D., & Scherer, K.R. (2009). Emotion recognition from expressions in face, voice, and body. The Multimodal Emotion Recognition Test (MERT). *Emotion, 9*, 691-704.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.

Ciarrochi, J.V., Chan, A.Y.C., & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality & Individual Differences, 28*, 539-561.

Colom, R., Escorial, S., & Rebollo, I. (2004). Sex differences on the progressive matrices are influenced by sex differences on spatial ability. *Personality & Individual Differences, 37,* 1289-1293.

Danthiir, V., Roberts, R.D., Pallier, G., & Stankov, L. (2001). What the nose knows: Olfaction and cognitive abilities. *Intelligence, 29,* 337-361.

Day, A.L., & Carroll, S.A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviours. *Personality & Individual Differences, 36,* 1443-1458.

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.

Grudnik, J.L., & Kranzler, J.H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence, 29,* 523-535.

Horn, J.L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-91). New York: Guilford Press.

Horn, J.L., & Stankov, L. (1982). Auditory and visual factors of intelligence. *Intelligence, 6*, 165-185.

Lynn, R., Irwing, P., & Cammock, T. (2001). Sex differences in general knowledge. *Intelligence, 30*, 27-39.

MacCann, C. (2010). Further examination of emotional intelligence as a standard intelligence: A latent variable analysis of fluid intelligence, crystallized intelligence, and emotional intelligence. *Personality and Individual Differences*, *49,* 490-496.

MacCann, C., & Roberts, R.D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion, 8*, 540-551.

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., … Goh, A. (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.

Mayer, J.D., & Geher, G. (1996). Emotional intelligence and the identification of emotion. *Intelligence, 22*, 89-114.

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. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*, 267-298.

Mayer, J.D., Salovey, P., Caruso, D.R., & Sitarenios, G. (2001). Emotional intelligence as a standard intelligence. *Emotion, 1*, 232-242.

McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 151-179). New York: Guilford Press.

Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion, 9*, 619-630.

O'Sullivan, M., & Ekman, P. (2004). Facial expression recognition and emotional intelligence. In G. Geher (Ed.), *Measuring emotional intelligence* (pp. 89-109). New York: Nova Science Publishing.

Raven, J., Raven, J.C., & Court, J.H. (1998). *Raven manual: Sec 4. Advanced progressive matrices*. Oxford: Oxford Psychologists Press.

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.

Roseman, I.J. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K.R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 68-91). 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.

Schlegel, K., Grandjean, D., & Scherer, K. (2012). Emotion recognition: Unidimensional ability or set of modality- and emotion-specific skills. Submitted to *Personality and Individual Differences*.

Stankov, L., & Horn, J.L. (1980). Human abilities revealed through auditory tests. *Journal of Educational Psychology, 72*, 21-44.

Stankov, L., Seizova-Cajic, T., & Roberts, R.D. (2001). Tactile and kinesthetic perceptual processes within the taxonomy of human cognitive abilities. *Intelligence, 29*, 1-29.

Suzuki, A., Hoshino, T., & Shigemasu, K. (2010). Happiness is unique: A latent structure of emotion recognition traits revealed by statistical model comparison. *Personality and Individual Differences, 48*, 196-201.

Tottenham, N. (2004). *Nim Stim face stimulus set*. Retrieved from: Research network on early experience and brain development website: http://www.macbrain.org/faces.

Received: December 25, 2011