

# **New Approaches to Measuring Emotional Intelligence: Exploring Methodological Issues with Two New Assessment Tools**

Carolyn Elizabeth MacCann

Doctor of Philosophy in Psychology

University of Sydney, 2006

## Abstract

New scoring and test construction methods for emotional intelligence (EI) are suggested as alternatives for current practice, where most tests are scored by group judgment and are in ratings-based format. Both the ratings-based format and the proportion-based scores resulting from group judgments may act as method effects, obscuring relationships between EI tests, and between EI and intelligence. In addition, scoring based on standards rather than group judgments add clarity to the meaning of test scores. For these reasons, two new measures of emotional intelligence (EI) are constructed: (1) the Situational Test of Emotional Understanding (STEU); and (2) the Situational Test of Emotion Management (STEM). Following test construction, validity evidence is collected from four multi-variate studies. The STEU's items and a standards-based scoring system are developed according to empirically derived appraisal theory concerning the structure of emotion [Roseman, 2001]. The STEM is developed as a Situational Judgment Test (SJT) with situations representing sadness, fear and anger in work life and personal life settings. Two qualitative studies form the basis for the STEM's item development: (1) content analysis of responses to semi-structured interviews with 31 psychology undergraduates and 19 community volunteers; and (2) content analysis of free responses to targeted vignettes created from these semi-structured interviews ( $N = 99$ ). The STEM may be scored according to two expert panels of emotions researchers, psychologists, therapists and life coaches ( $N = 12$  and  $N = 6$ ).

In the first multi-variate study ( $N = 207$  psychology undergraduates), both STEU and STEM scores relate strongly to vocabulary test scores and moderately to Agreeableness but no other dimension from the five-factor model of personality. STEU scores predict psychology grade and an emotionally-oriented thinking style after controlling vocabulary and personality test scores ( $\Delta R^2 = .08$  and  $.06$  respectively). STEM scores did not predict academic achievement but did predict emotionally-oriented thinking and life satisfaction ( $\Delta R^2 = .07$  and  $.05$  for emotionally-oriented thinking and  $.04$  for life satisfaction). In the second multi-variate study, STEU scores predict lower levels of state anxiety,

and STEM scores predict lower levels of state anxiety, depression, and stress among 149 community volunteers from Sydney, Australia. In the third multi-variate study ( $N = 181$  psychology undergraduates), Strategic EI, fluid intelligence (Gf) and crystallized intelligence (Gc) were each measured with three indicators, allowing these constructs to be assessed at the latent variable level. Nested structural equation models show that Strategic EI and Gc form separate latent factors ( $\Delta\chi^2(1) = 12.44, p < .001$ ). However, these factors relate very strongly ( $r = .73$ ), indicating that Strategic EI may be a primary mental ability underlying Gc. In this study, STEM scores relate to emotionally-oriented thinking but not loneliness, life satisfaction or state stress, and STEU scores do not relate to any of these. STEM scores are significantly and meaningfully higher for females ( $d = .80$ ), irrespective of gender differences in verbal ability or personality, or whether expert scores are derived from male or female experts. The fourth multi-variate study ( $N = 118$  psychology undergraduates) distinguishes an EI latent factor (indicated by scores on the STEU, STEM and two emotion recognition ability measures) from a general cognitive ability factor (indicated by three intelligence measures;  $\Delta\chi^2(1) = 10.49, p < .001$ ), although again cognitive ability and EI factors were strongly related ( $r = .66$ ). Again, STEM scores were significantly higher for females ( $d = .44$ ) and both STEU and STEM relate to Agreeableness but not to any other dimension from the five-factor model of personality. Taken together, results suggest that: (1) STEU and STEM scores are reasonably reliable and valid tests of EI; (2) EI tests assess slightly different constructs to existing measures of Gc, but more likely form a new primary mental ability within Gc than an entirely separate construct; and (3) the female superiority for EI tests may prove useful for addressing adverse impact in applied settings (e.g., selection for employment, promotion or educational opportunities), particularly given that many current assessment tools result in a male advantage.

## Notes

Data in the current thesis are derived from two qualitative studies and four multi-variate quantitative studies involving participants from first year and third year undergraduate psychology courses as well as community volunteers. For each of the six studies, approval was obtained from the Sydney University Human Research Ethics Committee (HREC) prior to commencement of the study. Parts of Chapter 1 were published in revised form as a review article in the *International Journal of Organizational Analysis* (MacCann, Matthews, Zeidner, Roberts, 2003), and as a book chapter in *Measuring Emotional Intelligence: Common Ground and Controversy*, edited by Glenn Geher (MacCann, Matthews, Zeidner, Roberts, 2004). Although three academics contributed substantially to these two publications, this thesis is solely my own work (subject to the suggestions and influences of my two primary supervisors, Professor Beryl Hesketh, and Dr Richard Roberts).

## Acknowledgements

This research was supported by the Australian Postgraduate Award, and by three PRGs (Post-Graduate Research Grants) given internally by the University of Sydney. There are a number of people I want to thank for their academic and emotional support over the last four years. The first of these is Professor Beryl Hesketh, who took over my primary supervision in 2003, and found time for me despite her incredibly hectic schedule. Her dedication, professionalism, and determination saw me through this (epic) journey. I would also like to thank Dr Richard Roberts (originally my original primary supervisor, now an associate), whose inspiration, creativity, and mentoring were invaluable. I also want to thank my parents, Dr Bob MacCann, and Mrs Ann MacCann, for their many hours of emotional support and proof reading that must have been unbelievably tedious. Thank-you also to Adam Pringle for the years of support and understanding while I studied full-time. Thank-you also to the academics who have advised and helped me along the way: Barbara Griffin, David Grayson, Sabina Kleitman, Sophie Ellwood, Gerry Pallier, Damian Birney, Tony Grant, and Margaret Charles. Thank you to my close friends: Michelle Barlow, Michelle Dawson, Vanessa Lupton, Ian and Pat Pringle. Thank you also to my fellow PhD students for advice, support and even (occasionally) fun: Lisa Karlov, David Bowman, Jane Rouse, Katherine Ryan, Veneta Bastian, Blossom Stephan, Pooja Sawricker, Justine Gatt, and Carlo Capponechia.

## Table of Contents

Title Page: New Approaches to Measuring Emotional Intelligence .....	1
Abstract .....	2
Notes.....	4
Acknowledgements .....	5
Table of Contents .....	6
Preface.....	8
Chapter 1: Literature Review .....	11
1. Outline of Chapter 1 .....	12
2. Background to EI: History, Models, and Concepts .....	12
3. Evaluation of Self-Report EI .....	15
4. Performance-Based EI: Methodological Issues.....	18
5. Dealing with Methodological Issues.....	26
6. Validity Issues in Performance-Based EI .....	29
7. Goals and Aims.....	38
8. Concluding Comments for Chapter 1 .....	40
Chapter 2: Test Development of the STEU and STEM .....	42
Part I of Chapter 2 .....	43
Part I of Chapter 2: Test Development of the STEU.....	43
1. Appraisal Theories of Emotion.....	43
2. Using Roseman’s Model to Generate Items for the STEU.....	46
3. Final Description of the STEU .....	47
Part II of Chapter 2: Test Development of the STEM.....	48
1. Study 1: Generation of Situations.....	49
2. Study 2: Response Option Generation.....	63
3. Study 3: Expert Judgment of the Correctness of Situations .....	71
4. Study 4: Development of a Standards-Based Scoring Key for the STEM .....	76
Part III of Chapter 2: Chapter Summary .....	79
1. Overview of Test Development of the STEU and STEM .....	79
2. Limitations of Test Development of the STEU and STEM.....	80
3. Use of the STEU and STEM in Subsequent Chapters.....	81

Chapter 3: Methodological Issues and Test Validation of the STEU and STEM .....	82
1. Introduction to Chapter 3 .....	82
2. Method .....	92
3. Results .....	98
4. Discussion .....	117
Chapter 4: Does EI Predict Psychological Distress? .....	125
1. Introduction to Chapter 4 .....	125
2. Method .....	130
3. Results .....	134
4. Discussion .....	146
Chapter 5: Distinguishing EI from Existing Constructs: A Latent Variable Analysis .....	151
1. Introduction to Chapter 5 .....	151
2. Method .....	163
3. Results .....	171
4. Discussion .....	194
Chapter 6: Do Strategic EI and Emotion Recognition Ability Form a Unified EI Construct? .....	203
1. Introduction to Chapter 6 .....	203
2. Method .....	212
3. Results .....	217
4. Discussion .....	230
Chapter 7: General Discussion .....	237
1. Standards-Based Scoring in Emotional Intelligence .....	238
2. Consensus Scoring and Ratings-based Response Formats as “Method Effects” .....	243
3. Validity of Performance-Based EI .....	254
4. Validity of Trait EI .....	260
5. Alternative Directions for the SJT Instrument Construction .....	261
6. Concluding Comments .....	263
Reference List .....	264

## Preface

This thesis has two primary foci: (1) to critique and potentially advance test development methodologies for socio-emotional constructs such as emotional intelligence (EI); and (2) to examine the validity of EI as a construct, particularly the strategic component of EI. Chapter 1 provides a detailed background on current research in EI, describing the major trait-based and performance-based theoretical models of EI. Current scoring mechanisms in performance-based EI are reviewed in detail, with two possible method effects implicated: a response-format method effect, and a consensus-scoring method effect. New standards-based approaches to scoring are suggested for the development of two new tests of EI: the Situational Test of Emotional Understanding (STEU) and the Situational Test of Emotion Management (STEM). Existing evidence on the convergent, divergent and criterion-related validity of current performance-based EI tests is summarised, both to provide a context for this thesis' aims to validate the STEU and STEM as instruments, as well as to examine the current validity evidence for the EI construct.

Following on from this introductory exposition, Chapter 2 describes the development of the STEU and the STEM measures of Strategic EI. The STEU is developed according to Roseman's (2001) empirically-derived appraisal theory of the structure of emotions. This theory is also used to develop verbalisable standards to score the STEU, as an alternative to expert or normative judgment. The STEM is developed according to the Situational Judgment Test (SJT) paradigm, and two possible scoring systems are created: (1) a standards-based system based on a summary of empirical findings in coping research; and (2) an expert-based scoring system derived from the judgments of EI researchers, counsellors, counselling students and life coaches. As well as two potential scoring systems, two different response formats for the STEM are possible: (1) a multiple-choice format; and (2) a ratings-based format, where each option is rated for correctness (rather than the single best option chosen, as is the case in multiple-choice tests). Chapter 2 also examines the validity of expert



judgment, assessing whether the experts scored more highly on tests than novices, and whether gender differences in EI scores are due to the gender composition of the expert sample.

Following Chapter 2's detailed description of the STEU and STEM's test construction, Chapter 3 outlines a scoring and validation study of these measures. This chapter examines the effect of response format by comparing the multiple-choice and ratings-based versions of the STEM in terms of their reliability and validity. Validity evidence for the STEU and STEM is evaluated according to their relationship to each other and to a third general measure of EI, their distinction from vocabulary and personality, and their incremental prediction of alexithymia, personality, life satisfaction, psychological distress and academic achievement when vocabulary and personality are controlled.

Although test validity was examined from several perspectives in Chapter 3, Chapter 4 addresses the need for validity evidence to be obtained from samples other than first year undergraduate psychology students. In Chapter 4, the reliability of the STEU and STEM is examined in a sample of community volunteers from central Sydney, along with the concurrent validity of the tests with respect to state measures of anxiety, depression and stress, and some very brief demographic indicators of social well-being.

Chapter 5 examines the relationship between EI and intelligence when both are measured as latent variables – i.e., crystallized intelligence ( $G_c$ ), fluid intelligence ( $G_f$ ) and EI are each assessed with multiple indicators (rather than with a single test score, as in Chapters 3 and 4). In addition, all EI tests are administered in multiple-choice format to control for response format. Chapter 5 also examines whether the distinction between intelligence and EI may be due to the different systems in past research (standards-scoring for intelligence and consensus-scoring for EI). The distinction of consensus-scored EI from both standards-scored and consensus-scored intelligence is examined to determine whether consensus scoring is acting as a method effect. To provide further validation of the STEU and STEM tests and the EI construct generally, evidence of the incremental prediction of state stress, state loneliness, life satisfaction and alexithymia over intelligence and personality is collected

for both trait and performance-based EI. In addition, Chapter 5 briefly examines the structure of a measure of trait EI (the Schutte Self-Report Scale), and its distinction from personality at the facet level of the NEO-PI-R model of personality.

Chapter 6 investigates EI's distinction from intelligence with a similar latent variable approach to Chapter 5. However, EI is defined by indicators of emotion recognition ability in addition to the verbally-presented measures of Strategic EI used in previous chapters. This allows a test of whether Strategic EI and emotion recognition ability do form an EI factor separable from intelligence, as well as allowing EI to be assessed at a broader conceptual level. In addition, this chapter assesses whether relationships between EI tests and between EI and fluid and crystallized intelligence can be accounted for by reading comprehension ability.

The final chapter discusses recommendations for test construction approaches in the light of evidence from the four empirical chapters. The final question to be answered, based on the combined findings of Chapters 2 to 6 is: *Is there a thing called "emotional intelligence" that is separate from intelligence and personality, and what use is it?* Evidence that EI is somewhat distinct from existing constructs is discussed in terms of the structural equation models examining EI's distinction from fluid and crystallized intelligence. Evidence that EI may have some utility in applied settings such as selection for employment or educational opportunities is discussed in terms of the hierarchical regressions showing EI's unique prediction of criterion variables, as well as the group differences in EI compared to group differences in the cognitive ability tests currently used in such settings. The answers to these questions are cautiously positive. Strategic EI appears to be a new component within crystallized intelligence, but may predict the more emotionally-relevant aspects of performance, and may lower adverse impact on women when used for selection purposes. However, such distinctions are slight, indicating that EI tests capture a different angle in the cognitive ability testing picture, rather than painting a different picture altogether.